

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 6 Novembre 2024 par

Maxime DUMONT

Deep Learning pour le phasage de télescope segmenté.
Analyse de front d'onde en plan focal pour télescope spatial : application au
phasage d'un Cubesat deployable pour l'observation de la Terre.

Discipline

Physique et sciences de la matière

Spécialité

Optique, photonique et traitement d'image

École doctorale

ED 352

Laboratoire/Partenaires de recherche

Laboratoire d'Astrophysique de Marseille
ONERA
Faculdade de Engenharia da Universidade do Porto

.....
Composition du jury

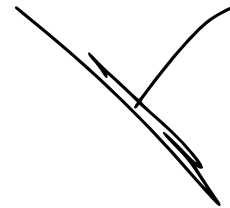
Remi SOUMMER STScI	Rapporteur
Esteban VERA Pontificia Universidad Católica de Valparaíso	Rapporteur
Charlotte BOND UK-ATC	Examinatrice
Marc FERRARI Observatoire de Haute Provence	Président
Jean-François SAUVAGE ONERA / LAM	Directeur
Jaime CARDOSO INESCTEC	Co-Directeur de thèse
Carlos CORREIA FEUP	Invité

Affidavit

Je soussigné, Maxime DUMONT, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de Jean-François SAUVAGE, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect à la fois de la charte nationale de déontologie des métiers de la recherche et de la charte d'Aix-Marseille Université relative à la lutte contre le plagiat.

Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Marseille le 03/09/2024



Cette œuvre est mise à disposition selon les termes de la [Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Liste de publications et participation aux conférences

Liste des publications et/ou brevets réalisées dans le cadre du projet de thèse :

1. Dumont, Maxime, et al. "Phasing segmented telescopes via deep learning methods : application to a deployable CubeSat." JOSA A 41.3 (2024) : 489-499.

Participation aux conférences et écoles d'été au cours de la période de thèse :

1. JRIOA22, Nice
2. SPIE 2022, Montreal
3. COSPAR 2023, Singapour
4. AO4ELT7, Avignon
5. JRIOA24, Palaiseau
6. SPIE 2024, Yokohama

Résumé et mots clés

Atteindre à la fois une haute résolution angulaire et des temps de revisite fréquents pour l'observation de la Terre ou d'autres planètes depuis une orbite basse pose plusieurs défis importants. Trouver un équilibre entre la nécessité d'une plus grande ouverture et les coûts accrus et contraintes associées nécessite des solutions innovantes. AZIMOV, un prototype de charge utile pour un télescope déployable segmenté sur un CubeSat 6U actuellement en développement, aborde ce problème en utilisant un télescope avec une ouverture de 30 cm. Ce grand miroir primaire permet d'obtenir une distance d'échantillonnage au sol de 1 mètre dans le spectre visible. Cependant, pour des performances optimales, un alignement précis (ou phasage) des segments du miroir primaire est crucial. En raison des contraintes des CubeSats, telles que le volume, la puissance et les ressources de calcul limitées, les méthodes conventionnelles de détection de front d'onde ne conviennent pas. L'analyse de front d'onde en plan-focal est la seule approche pratique pour ces petites plateformes, mais les techniques traditionnelles sont souvent itératives et nécessitent beaucoup de calculs en raison de la relation non linéaire entre la phase du champ électromagnétique et l'intensité de l'image.

Cette thèse explore l'utilisation du deep learning pour corriger les aberrations de piston et de tip-tilt sur les quatre segments du miroir primaire à partir d'une seule image du plan focal. Nous montrons que notre approche, utilisant des réseaux de neurones convolutifs, peut atteindre des performances à la limite de diffraction lors de l'observation d'une source ponctuelle. Cette méthode d'apprentissage profond est robuste face au bruit et aux aberrations de plus hauts ordres, et elle surpasse les méthodes itératives classiques en termes de rapidité, de précision et de robustesse. Des tests sur des données expérimentales confirment l'adaptabilité des réseaux de neurones entre les données simulées et expérimentales et montrent la faisabilité de notre méthode sur des données réelles. Dans le cadre de l'observation de la Terre depuis une orbite basse, ou lors de l'imagerie d'objets étendus inconnus à la surface terrestre, notre approche, améliorée pour l'analyse des objets étendus, montre des performances pouvant atteindre la limite de diffraction, tout en soulignant les limitations de l'analyse de front d'onde en plan focal à partir d'une seule image, dues à aux caractéristiques de l'objet observé.

Mots clés : Analyse de front d'onde plan-focal, Deep Learning, phasage, télescope segmenté spatial.

Abstract and keywords

Achieving both high angular resolution and frequent revisit times for Earth or other planetary observations from low Earth orbit poses several significant challenges. Balancing the need for a larger aperture with the increased costs and associated constraints requires innovative solutions. AZIMOV, a payload prototype for a segmented deployable telescope on a 6U CubeSat currently in development, addresses this issue by utilizing a telescope with a 30 cm aperture. This large primary mirror enables a ground sampling distance of 1 meter in the visible spectrum. However, for optimal performance, precise alignment (or phasing) of the primary mirror segments is crucial. Due to the constraints of CubeSats, such as limited volume, power, and computational resources, conventional wavefront sensing methods are not suitable. Focal-plane wavefront analysis is the only practical approach for these small platforms, but traditional techniques are often iterative and require significant computation due to the nonlinear relationship between the phase of the electromagnetic field and the image intensity.

This thesis explores the use of Deep Learning to correct piston and tip-tilt aberrations on the four segments of the primary mirror from a single focal plane image. We demonstrate that our approach, using convolutional neural networks, can achieve diffraction-limited performance when observing a point source. This deep learning method is robust against noise and higher-order aberrations and outperforms classical iterative methods in terms of speed, accuracy, and robustness. Tests on experimental data confirm the adaptability of neural networks between simulated and experimental data and demonstrate the feasibility of our method on real-world data. In the context of Earth observation from low Earth orbit or when imaging unknown extended objects on the Earth's surface, our approach, enhanced for extended object analysis, shows performance that can reach the diffraction limit while highlighting the limitations of single-image focal-plane wavefront analysis due to the characteristics of the observed object.

Keywords: Focal-plane wavefront analysis, Deep Learning, Phasing, Segmented Space Telescope

Remerciements

Afin de ne pas susciter une quelconque jalousie chez les lecteurs de ce manuscrit, je prierai le lecteur de ne pas lister cette liste de remerciements par ordre d'importance, mais plutôt je lui laisserai le soin d'observer une classification peut-être plus originale pour organiser mes pensées.

Prenons comme référence mon bureau au Laboratoire d'Astrophysique de Marseille.

À environ 5 pas de la porte d'entrée de mon bureau se trouve le bureau de Romain Fetick, la légende, sa porte est à l'image de son cœur, toujours ouvert et prêt à partager sa science. Rajoutons une dizaine de pas pour y trouver le bureau d'un duo sportif, Paul et Jordi. À commencer par Paul dont le calme et la folie m'auront donné des sourires et d'excellents moments, je ne saurai jamais assez te remercier pour le temps que tu as passé à m'aider à me préparer pour ma soutenance. Puis arrive Jordi-san, compagnon de voyage Japonais (avec une immense pensée pour des clefs de casier perdues à Tokyo), une enfance commune dans une fameuse station balnéaire méditerranéenne, un nombre incalculables de fous rires et avant tout un immense "copaing". En parlant de copaing, un même nombre de pas me sépare d'un second copaing à la capillarité douteuse, Arnaud. Une amitié qui a rapidement vu le jour d'abord co-bureau à écouter de la musique peut être trop fort, jusqu'à ce que la troisième année nous sépare. Ensuite, quelques pas de plus, changement de direction, et on retrouve un directeur de thèse d'exception, Jeff, ou "papa Jeff" pour les intimes. Je souhaite amplement te remercier pour tout; si la thèse est aboutie à ce jour, c'est grâce à toi. Je te remercie pour ces échanges scientifiques et amicaux dont tu as fait preuve; ça a été un réel et sincère plaisir de travailler à tes côtés. En s'éloignant un peu, on retrouve le bureau de la paire Benoit-Thierry, des puits de savoir scientifique qui ne rechignent pas à donner d'excellents conseils et des sourires malgré leur emploi du temps de ministre. On retrouve ensuite un certain Cedric Taïssir Héritier Salama, le chercheur Italiano-Egypto-Français. Merci Taï pour ton temps, tes conseils et la pertinence de tes remarques. Tu m'as fait découvrir comment apprécier les cafés, les pizzas, et malgré ton mental "de chips", j'ai passé d'excellents moments à tes côtés. On retrouve aussi un peu plus loin Mathieu le nordiste fan du LOSC, bien heureux en ligue des champions cette année! François et son bureau des stagiaires tant adoré, d'ici à ce que tu lises ces lignes tu aura peut être demenagé, mais sache que pendant tout mon parcours au LAM tu as fais partie intégrante du LAMpanada. Enfin au plus loin de mon bureau se trouve Eloise, peut être aurais-je du te mettre en numéro 1, au plus proche de mon bureau vu ces heures passées à la salle café..

Montons d'un étage pour retrouver Mahawa, ma jumelle de thèse dont les partages d'expérience et les tentatives pour se rassurer s'accumulaient au fur et à mesure des avancements de la thèse, aloha! On retrouve ensuite Morgan; je te remercie pour ces

heures à discuter de tout sauf de science, puis occasionnellement de Deep Learning. Un plaisir d'avoir travaillé avec toi depuis le stage de fin d'étude jusqu'au dernier moment du doctorat. Ce tour du LAM me fait penser aussi à Nicolas Levraud, toujours de bon conseil et d'expérience; je n'oublierai jamais cette fois où tu nous as fait découvrir Naguini entre deux chapitres de ta rédaction!

Prenons maintenant un peu de recul depuis le LAM et restons à Marseille pour remercier la famille du handball de l'ASPTT, pour ne pas tous les nommer : Samy, Clément, Flora, Pierre, Carole, Lucie, Nathan, Nathan, Jonathan, Pauline, Thomas, Micka, Kahina, Hugo, Christophe, Paul, Corentin, Philou, Stef et j'en passe. Une pensée forte aussi pour Assaf, dont le destin nous a mis tous les deux dans la cité phocéenne.

Petit crochet par le Var pour y remercier Elodie, Benjamin, Eloi, et Come, un quatuor riche en énergie, pas de tout repos mais toujours présent.

On peut s'éloigner une fois de plus pour arriver dans la plus belle ville du monde, Nîmes, où se trouve l'inévitable maman. C'est une évidence que sans toi je n'en serais pas là aujourd'hui; je t'ai certainement transmis un bon paquet de stress et d'émotions pour des projets de voyage à l'autre bout du monde et mes études mais tu m'as toujours soutenu et rendu ça possible, donc si le doctorat est possible, c'est aussi grâce à toi. Une grosse pensée pour Tinou et Pinpin, évidemment. Dans la périphérie nîmoise se trouve aussi mon plus grand soutien psychologique des derniers mois de la thèse : *la casa*. Un immense merci à Mathias et Loan, m'ayant réservé une chambre pour l'entièreté des week-ends de septembre à novembre; ces parenthèses le temps de plusieurs week-ends m'ont énormément servi à tenir le fil, je ne l'oublierai jamais. Une belle pensée pour les copains du sud aussi, Robin, Cedric, Jerem, Elise, Steph. Plus à l'Ouest, dans l'Herault c'est ma chère grand-mère vers laquelle mes pensées se tournent.

On se retrouve maintenant dans l'Isère pour remercier un compagnon de voyage, Serge, toujours présent, fidèle à lui-même; courage pour la fin de ta thèse! Une pensée toute particulière vient à une fleur autrefois grenobloise, une rose dont la dureté des piquants est masquée par la beauté de ses pétales. Merci d'avoir été une oreille attentive et d'avoir essayé tant bien que mal de m'écouter et de comprendre mes travaux.

Changement de direction pour la Suisse, où ce bon vieux Simon réside, expert du tour de magie au chapeau; une amitié unique et ancrée dans le temps, tu es bien trop loin aujourd'hui!

Plus à l'ouest, on va recentrer de longues amitiés sur Lyon où tout a commencé, le groupe des copains de CPE, qui ont toujours été là pour suivre les aventures de leurs doctorants du groupe. Merci à vous les loulous, on en a fait du chemin depuis les HHs. Merci Artur, Ben, Dim, Ed, Elo, Hélène, Hippo, Julie, Laure, Leslie, Lucas, Marion, Mathieu, Simon, Val, et Yoyo.

La capitale recèle aussi des personnes présentes pour moi lors de mes travaux, à l'image du trinôme 8, Lolo, Théo, et Popo. Mes chers compagnons de galère depuis l'époque d'Alzon, on voulait aller loin, nous y voilà! La team HRA ne sera pas oubliée évidemment dans ces remerciements, et les locaux exceptionnels de Chatillon, merci à tous pour ces bons moments sous le beau temps parisien.

Toujours plus au nord se retrouve ma famille; une pensée pour ma grand-mère, Sophie, Didier, Thomas, Camille, Gus, et toute votre fine équipe, vous faites partie de ce doctorat vous aussi!

Au Luxembourg se retrouve un ami qui m'est très cher, Nico, un compagnon d'aventure depuis des années, impossible de se perdre de vue, et ce n'est pas prévu; merci d'avoir été là même si les aléas de la vie nous envoient à des kilomètres l'un de l'autre.

Partons à l'international avec la ville de Porto où une partie du doctorat s'est retrouvée. Un immense merci à mes chères amies et coloc Valentina et Cecilia, un trio franco-italiano-indonésien qui nous aura fait voyager à travers le monde pour se retrouver. Une immense pensée et un souvenir chaleureux d'un mariage à Helsinki. I'll also never forget the famous *Loteria Nacional*. Sur le site de FEUP se trouve Carlos, un immense merci Carlos pour ton encadrement pendant ces 3 ans, ta rigueur ne cessera de m'impressionner et m'a fait énormément progresser. Travailler avec toi m'a beaucoup fait grandir dans ma manière d'aborder les problèmes et dans mes réflexions scientifiques.

Toujours plus loin, c'est en Écosse qu'on se retrouve pour remercier Noah, un réel plaisir une nouvelle fois de travailler à tes côtés. Ta bonne humeur, tes anecdotes, et ta tchatche rendent les moments et les échanges facilement vivables, merci Noah!

Enfin pour s'éloigner vers des lieux invisitables, une pensée forte pour ceux qui ne pourront voir l'accomplissement des travaux : mes grands-pères dans un premier temps, et celui qui prétendait que j'avais "la tête dans les étoiles", mon père.

Table des matières

Affidavit	2
Liste de publications et participation aux conférences	3
Résumé et mots clés	4
Abstract and keywords	5
Remerciements	6
Table des matières	9
Introduction	13
0.1. Imagerie à haute résolution angulaire	14
0.1.1. Télescopes au sol	14
0.1.2. Télescopes spatiaux	14
0.2. Aberrations et analyse du front d'onde	19
0.2.1. Origine et influence des aberrations	19
0.2.2. Impact sur la formation d'image	23
0.2.3. Analyse du front d'onde	24
0.2.4. Phasage des télescopes segmentés	31
0.2.5. Limite de l'analyse du front d'onde pour le phasage de télescope pour le NewSpace	35
0.3. Ouverture de la mesure du front d'onde par Machine Learning	37
0.3.1. Le ML comme reconstituteur à partir d'ASO	37
0.3.2. Analyse du front d'onde plan-focal par ML	37
0.4. Contribution des travaux de thèse	38
0.4.1. Deep Learning pour le phasage en plan-focal	39
0.4.2. Télescopes d'intérêt	39
0.4.3. Plan du manuscrit	45
1. Formation d'image	47
1.1. Décomposition de la phase sur pupille circulaire	48
1.2. Pupille d'un télescope	50
1.2.1. Forme de la pupille	50
1.2.2. Asymétrie	51
1.3. Aberrations d'intérêt : mesure du piston-tip-tilt	54
1.3.1. Phasage des piston-tip-tilt	54
1.3.2. Normalisation des aberrations sur une pupille carrée	55

1.3.3. Définition de l'erreur de front d'onde	56
1.4. Formation d'image d'un point source : la PSF d'un télescope segmenté	57
1.4.1. Sources de perturbations additionnelles	57
1.4.2. Génération d'images	60
1.5. Cas d'un objet étendu	66
1.5.1. Relation objet-image	66
1.5.2. Sélection d'images caractéristiques	69
1.5.3. Base de données d'objet	71
2. Methodologie	74
2.1. Introduction aux réseaux de neurones	75
2.1.1. Principe	76
2.1.2. Multi-Layer Perceptron	76
2.1.3. Réseaux de neurones à couches de convolution	79
2.1.4. Fonction d'activation	81
2.1.5. Loss	82
2.1.6. Optimiseurs et retropropagation des gradients	83
2.2. Stratégie d'apprentissage	84
2.3. Architecture	85
2.3.1. AlexNet & VGG-Net	85
2.3.2. ResNet	87
2.3.3. EfficientNetB0	88
2.3.4. Variational Auto-Encoder	90
2.4. Méthodologie pour le phasage	93
2.4.1. Simulation de données	93
2.4.2. Choix de la loss	93
2.4.3. Hyperparametres	94
2.4.4. Le phasage par Deep Learning	94
3. Étude d'une source ponctuelle	96
3.1. Étude d'une pupille simplifiée	98
3.1.1. 1 Dimension : Tip	98
3.1.2. 2 Dimensions - 1 Pupille	101
3.1.3. 1 Dimension - 2 Pupilles	103
3.1.4. 2 Dimensions - 2 Pupilles	106
3.2. Étude sur pupille complète	110
3.2.1. Optimisation de la surface d'asymétrie	111
3.2.2. Résultats	114
3.2.3. Choix de l'architecture	115
3.3. Limitation de l'estimation du piston	116
3.4. Propagation de bruit	119
3.4.1. Performance à SNR constant	119
3.4.2. Robustesse au SNR	120
3.5. Influence des aberrations de hauts ordres	121

3.6. Comparaison aux méthodes d'AFOPF applicables	123
3.6.1. Diversité de Phase	123
3.6.2. Image Sharpening	124
3.6.3. Comparaison et temps d'estimation	125
3.7. Scenario de phasage complet	128
3.8. Application à de plus grands télescopes	131
3.8.1. Réseaux de neurones	132
3.8.2. Résultats	132
3.9. Application au petaling de l'ELT (sans turbulence)	135
3.9.1. Formation d'image	135
3.9.2. Réseau de neurones	137
3.9.3. Résultats	137
4. Validation expérimentale	140
4.1. Acquisition de données	141
4.1.1. Banc optique	141
4.1.2. Génération des données expérimentales	142
4.2. Validation des données	144
4.2.1. Mesure de la fréquence de coupure	144
4.2.2. Ajustement des paramètres expérimentaux	147
4.3. Application aux réseaux de neurones	150
4.3.1. Validation simulée et expérimentale	150
4.3.2. Performance des R-Sim appliqué aux données expérimentale	153
4.3.3. Fine-tuning	154
4.4. Jeu de données optimal	157
5. Etude de l'objet étendu	160
5.1. Contexte	161
5.1.1. Observation de la Terre en orbite basse	161
5.1.2. Base et limite de la formation d'image sur objet étendu	163
5.2. Stratégie de phasage par réseaux de neurones	165
5.2.1. Simulation d'objet	165
5.2.2. Stratégie naïve d'estimation : méthode directe	169
5.2.3. Renforcement de la stratégie d'apprentissage : PINN	173
5.3. Estimation de la phase	174
5.3.1. Vérification de la méthode sur point source	174
5.3.2. Application aux objets simulés	176
5.3.3. Diversité d'objet connu	176
5.4. Limitation de l'analyse de front d'onde	179
5.4.1. Contraste de l'objet	179
5.4.2. Sensibilité de la phase au contraste	182
5.5. Phasage en boucle fermée	183
Conclusion et perspectives	187

Table des matières

Bibliographie	190
ANNEXES	197
A. Analyse des paramètres d'un réseau de neurones convolutif	198
Table des figures	205
Acronyms	211
Article	214

Introduction

Dans ce chapitre, je présenterai le contexte dense des travaux de la thèse. De la formation d'images au sein d'un télescope, en passant par les aberrations de la phase du champ électromagnétique et leur incidence sur l'image formée. Le phasage des télescopes sera notamment abordé, étant le sujet central des travaux de la thèse présentée. Enfin, les contributions de l'intelligence artificielle à l'instrumentation des télescopes seront présentées.

Sommaire

0.1. Imagerie à haute résolution angulaire	14
0.1.1. Télescopes au sol	14
0.1.2. Télescopes spatiaux	14
0.2. Aberrations et analyse du front d'onde	19
0.2.1. Origine et influence des aberrations	19
0.2.2. Impact sur la formation d'image	23
0.2.3. Analyse du front d'onde	24
0.2.4. Phasage des télescopes segmentés	31
0.2.5. Limite de l'analyse du front d'onde pour le phasage de télescope pour le NewSpace	35
0.3. Ouverture de la mesure du front d'onde par Machine Learning	37
0.3.1. Le ML comme reconstituteur à partir d'ASO	37
0.3.2. Analyse du front d'onde plan-focal par ML	37
0.4. Contribution des travaux de thèse	38
0.4.1. Deep Learning pour le phasage en plan-focal	39
0.4.2. Télescopes d'intérêt	39
0.4.3. Plan du manuscrit	45

0.1. Imagerie à haute résolution angulaire

L'[Imagerie à Haute Résolution Angulaire \(IHRA\)](#) est la capacité d'un instrument d'optique à discerner les détails fins d'une scène observée. Par exemple l'astronome cherche à discerner dans le ciel nocturne des détails de l'ordre du millième de seconde d'arc. "*Plus un télescope est long, plus il peut voir loin dans l'espace!*". C'est ce que devait se dire l'auteur de ce manuscrit étant jeune. Mais qu'en est-il vraiment? En réalité, il y a bien un lien entre la taille du télescope et la résolution de ce dernier, ou plus précisément sa résolution angulaire. Cependant ce n'est pas la longueur du télescope qui vient influencer sur cette résolution mais bien le diamètre de son miroir primaire. Cette résolution angulaire est régie par le rapport $\frac{\lambda}{D}$ où D est le diamètre du télescope et λ la longueur d'onde d'observation. Les plus grands télescopes construits aujourd'hui atteignent plus d'une dizaine de mètres de diamètre, et observent dans les longueurs d'onde optiques proche infra-rouge. Un télescope de 10 m observant à 1 μm atteint une résolution angulaire de 10 μrad , soit environ 2 mas. On peut donc directement jouer sur la longueur d'onde pour améliorer la résolution angulaire mais on se retrouve alors limité par la longueur d'onde d'émission de l'objet étudié. De plus, observer dans l'ultra-violet pourrait améliorer la résolution angulaire cependant cela complexifie les moyens optiques d'observations. Enfin l'atmosphère absorbe une grande partie de ces longueurs d'onde, rendant le ciel désespérément obscur. Il ne reste alors qu'à augmenter le diamètre du télescope.

0.1.1. Télescopes au sol

Sur Terre, la construction des télescopes est limitée par le coût de ces derniers mais aussi par les difficultés mécaniques et techniques liées à la construction des très gros design. La surface du miroir primaire peut donc être étendue autant que les ressources et les technologies le permettent, autorisant aujourd'hui des résolutions angulaires inédites : c'est un télescope de 39 mètres qui est en cours de construction au Chili, l'[Extremely Large Telescope \(ELT\)](#) ([[Tam+22](#)]), multipliant par 4 le diamètre des précédents télescopes au sol : le Keck de 10 m de diamètre ou encore le Very Large Telescope de 8.2 m de diamètre.

Cependant, ces télescopes au diamètres inédits ne peuvent être composés d'une pupille monolithique en raison des contraintes de fabrication. La surface collectrice du miroir primaire doit donc être segmentée afin d'atteindre des diamètres aussi conséquents et la résolution adéquate.

0.1.2. Télescopes spatiaux

A contrario, dans l'espace, le volume est contraint car directement limitant pour le transport en orbite. De plus l'espace ajoute de nombreuses contraintes, amenant à une augmentation conséquente du budget du télescope en plus du budget de mise en orbite. La contrainte étant plus forte, les diamètres sont plus petits comparés aux télescopes au sol. Radiations, gradient thermiques, vibration lors de l'orbite, ou

encore collision, les télescopes spatiaux se doivent d'être robustes pour évoluer dans un environnement non coopératif. On peut distinguer les télescopes spatiaux en deux catégories : grandes missions spatiales (dénomination M ou L selon l'[ESA](#)) avec de larges télescopes, ou les missions de type NewSpace impliquant la multiplication de petits télescopes.

De même que pour leur homologue au sol, les télescopes spatiaux peuvent bénéficier d'un miroir primaire segmenté : il permet alors une facilité d'embarquement par rapport au volume du télescope, et autorise un diamètre du primaire supérieur à la surface de la plateforme.

Les grands télescopes

Les grands télescopes spatiaux répondent à des missions spatiales à haut budget et de grande envergure en terme de performance, de fiabilité et de temporalité. Ces projet de télescopes spatiaux permettent de s'affranchir de la contrainte majeure des télescopes au sol : l'atmosphère terrestre; au prix d'un coût de production et de lancement bien plus conséquent ($\sim 1\text{k €}/\text{kg}$) et une operabilité des opérations de maintenance bien plus complexe voir impossible. Un des télescopes spatiaux les plus emblématique est le *Hubble space telescope*, lancé en 1990, ayant servi entre autres à la découverte du trou noir au centre de la galaxie M87. Ce télescope de 2.4 m de diamètre a vu son successeur décoller le 25 décembre 2021, le [James Webb Space Telescope \(JWST\)](#). En plus de tripler la taille du miroir primaire par rapport à Hubble à l'aide d'un miroir segmenté de 18 segments, le JWST a l'avantage d'observer dans l'infrarouge (dans le visible pour Hubble). Les observations dans l'infrarouge donnent accès aux objets très lointains, mais pas uniquement. Elles permettent aussi de détecter et d'étudier des objets relativement froids à travers leur émission thermique. En effet, d'après la loi de Planck, tout corps émet de la lumière dont le pic d'émission dépend de la température. Ainsi, un astre d'une température de 5500°C , comme le Soleil, rayonne en lumière visible. Avec une température de 2550°C , une petite étoile comme Trappist-1, émet principalement dans l'infrarouge proche, et une planète comme la Terre, avec une température de 15°C , dans l'infrarouge moyen (à $10\text{ }\mu\text{m}$ de longueur d'onde). La lumière infrarouge a également la propriété de traverser les nuages en étant bien moins absorbée que la lumière visible. On peut ainsi voir ce qui se passe derrière certains nuages, alors qu'ils sont opaques à la lumière visible, ou pénétrer dans des nuages de gaz et de poussière où se forment les étoiles.

D'autres grand télescopes spatiaux sont mis en orbite afin d'assurer l'observation de la Terre. On peut aussi citer certains de ces télescopes à très haute résolution angulaire :

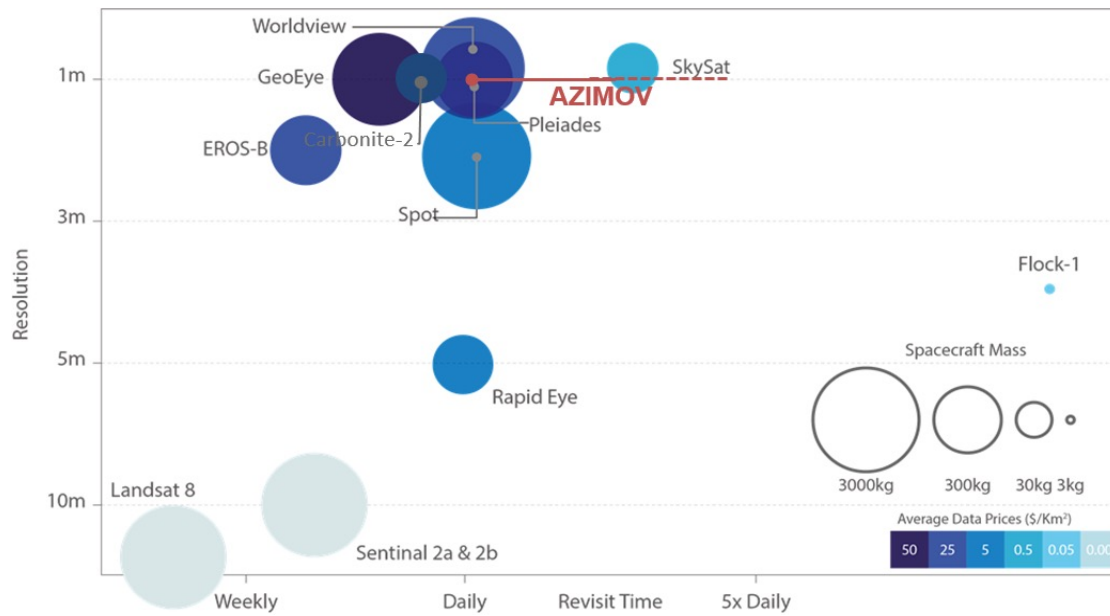


FIGURE 1. – Exemple de télescope spatiaux avec étude de leurs résolution par rapport à leur taux de revisite. Extrait de [ESA](#).

Les images d’observation de la Terre à haute résolution sont actuellement fournies par des systèmes tels que WorldView-4, Pléiades, SPOT 6/7, SkySat-3 et GeoEye-1 où leur résolution, leur poids et leurs taux de revisite peuvent être retrouvés figure 1. Ils ont des résolutions au sol comprises entre 0,31 m et 1,5 m, une masse allant de 120 à 2485 kg et une fauchée entre 1.1 km et 20 km. Ces systèmes sont très lourds et coûteux (par exemple, le coût de la mission WorldView-4 est de 835 millions de dollars), ce qui entraîne des coûts d’image élevés.

- **WorldView-4** : WorldView-4, lancé en 2016, fournit des images avec une résolution au sol de 0.31 m. Ce satellite a une masse d’environ 2485 kg et une fauchée de 13.1 km. Le coût de la mission est estimé à 835 millions de dollars ([AAD21]).
- **Pléiades** : Les satellites Pléiades offrent des images à haute résolution avec une résolution spatiale de 0.5 m. Chaque satellite a une masse d’environ 970 kg et une fauchée de 20 km. Ces satellites sont utilisés pour diverses applications, notamment la cartographie et la surveillance environnementale ([Leb+13]).
- **SPOT 6/7** : Les satellites SPOT 6 et SPOT 7 fournissent des images avec une résolution au sol de 1.5 m. Chacun a une masse de 720 kg et une fauchée de 60 km. Ces satellites sont conçus pour la continuité du service de la série SPOT ([Fra+19]).
- **SkySat-3** : SkySat-3, faisant partie de la constellation SkySat, offre des images à haute résolution avec une résolution spatiale de 0.9 m. Chaque satellite de la constellation a une masse d’environ 120 kg ([KGB18]).
- **GeoEye-1** : GeoEye-1 fournit des images avec une résolution au sol de 0.41 m en mode panchromatique. Le satellite a une masse de 1955 kg et une fauchée de

15.2 *km*. Il est utilisé pour la cartographie détaillée et la surveillance ([JWG13]).

Imagerie à haute résolution angulaire pour le NewSpace

Cependant, ces dernières années, l'accent s'est déplacé de l'augmentation de la résolution spatiale à l'augmentation de la couverture et du temps de revisite, visant à fournir une surveillance terrestre en quasi-temps réel. Le secteur de l'observation de la Terre depuis l'espace est passé de l'utilisation de quelques satellites grands, complexes et coûteux à l'établissement de constellations de nombreux petits satellites à faible coût : le *NewSpace*. Il vient inscrire une marque différente : une industrie spatiale à bas coût et accessible. De part la miniaturisation des composants électroniques, les nouveaux matériaux, et les algorithmes d'intelligence artificielle embarqués ; les nouveaux satellites sont petits, flexibles et bien moins chers que leurs prédécesseurs. Des « microsatellites » dont le poids dépasse à peine les 100 *kg*, nous sommes passés aujourd'hui aux nanosatellites (à peu près 10 *kg*) permettant l'émergence des constellations de satellites. Pour l'enveloppe budgétaire d'une grosse plate-forme (500M€), il est possible de placer sur orbite plusieurs dizaines (voire centaines) de petites plate-formes standardisées. Cette capacité nouvelle permet d'adresser une grandeur jusqu'alors peu accessible : la couverture temporelle, ou *revisit rate*. La standardisation des plateformes des satellites vient jouer un rôle majeur dans le *NewSpace*. En effet, le standard *CubeSat* (voir figure 2) est un format de satellite artificiel qui est le plus couramment utilisé pour les engins de très petite taille (moins de 20 kilogrammes). Son cahier des charges définit sa forme, sa masse maximale ainsi que certaines des caractéristiques de ses composants et de son fonctionnement. Dans sa forme basique, appelée 1U, c'est un cube (d'où son appellation) de 10 centimètres de côté et dont la masse ne doit pas excéder deux kilogrammes. La majorité des CubeSats sont formés par la juxtaposition de plusieurs cubes élémentaires (de 2U jusqu'à 24U).

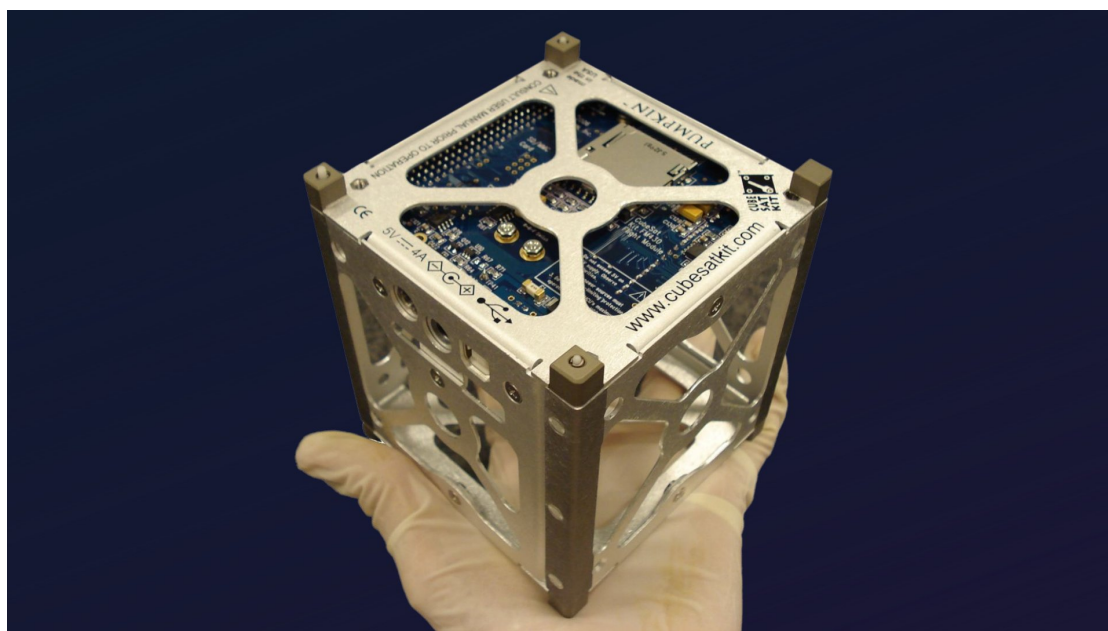


FIGURE 2. – Exemple de cubesat. Extrait de [Cubesatkit](http://Cubesatkit.com).

Les imageurs spatiaux s'inscrivent parfaitement dans ce domaine, car l'accessibilité en terme de coût de ces plateformes permet la multiplication de ces dernières, créant ainsi des constellations de satellites augmentant donc la fréquence temporelle d'observation. Cependant l'accès à la haute résolution se trouve compromise car la réduction du coût des satellites implique une réduction du volume de la plateforme, et donc du diamètre du miroir primaire du télescope. Une solution au volume restreint du *CubeSat* serait d'intégrer un miroir deployable au sein de la plateforme permettant non seulement d'optimiser le volume de lancement du télescope, mais aussi de conserver la haute résolution angulaire en multipliant la surface collectrice du télescope, liant ainsi haute résolution angulaire et haut taux de revisite.

Projet de l'IHRA pour le NewSpace

Nombre de projets d'imageurs émergents ont fait leur apparition s'inscrivant dans le NewSpace. L'imagerie à haute résolution angulaire pour l'observation de la Terre ou l'astronomie bénéficie aussi de cet essor technologique. Dolkens ([DKC18]) propose un télescope déployable d'une ouverture d' 1.5 m observant dans le visible, le Delft Space Telescope. Le satellite DeMi s'inscrit aussi dans le NewSpace, étant un télescope miniature avec un miroir déformable embarqué comme preuve de concept d'une technologie peu coûteuse et capable d'effectuer la détection d'exoplanètes à bas coût. On peut citer aussi GEO-HR, un projet de télescope large de l'ESO positionné en orbite géostationnaire possédant une ouverture de 3.5 mètres de diamètre. Ce projet vise une résolution au sol (*Ground Sampling Distance (GSD)*) de 4 m pour l'observation de la Terre pour des applications telles que la surveillance maritime ou la gestion des

inondations.

Quelques PME et spin-offs universitaires ont montré aussi de l'intérêt pour la technologie déployable, y compris SatRevolution, l'Université de Technologie de Delft, Wyvern, et Super-Sharp. Cependant, ces initiatives sont généralement orientées vers l'imagerie infrarouge thermique (notamment plus facile à réaliser en termes de qualité optique, par un facteur de x20) et pour des plateformes plus grandes (typiquement $\approx 100 + kg$). La principale lacune réside dans l'absence de démonstration de capacité et de piste éprouvée sur le contrôle précis de la position des miroirs, essentiel pour atteindre une haute résolution. En résumé, voici quelques acteurs de l'imagerie utilisant une petite-moyenne plateforme spatiale.

- **Delft Space Telescope** : Un télescope déployable d'une ouverture de 1.5 m observant dans le visible, proposé par [DKC18].
- **DeMi** : Un télescope miniature avec un miroir déformable embarqué, servant de preuve de concept pour une technologie peu coûteuse et capable d'effectuer la détection d'exoplanètes à bas coût.
- **GEO-HR** : Un projet de télescope large de l'ESO, positionné en orbite géostationnaire avec une ouverture de 3.5 mètres de diamètre. Ce projet vise une résolution au sol GSD de 4 m pour l'observation de la Terre, avec des applications telles que la surveillance maritime ou la gestion des inondations.
- **SatRevolution** : Une PME spécialisée dans les technologies spatiales déployables.
- **Wyvern** : Une entreprise focalisée sur l'imagerie infrarouge thermique pour des plateformes de taille moyenne.
- **Super-Sharp** : Un spin-off universitaire travaillant sur des technologies optiques avancées pour les satellites.

0.2. Aberrations et analyse du front d'onde

Cependant, pour une application sur un télescope réel, la résolution maximale atteignable ne dépend pas que du diamètre du télescope, mais également de sa qualité optique. La présence d'aberrations optiques qui perturbent la propagation de la lumière peut dégrader la résolution angulaire du télescope. Lorsque le niveau de qualité optique est suffisant pour qu'un système optique réel puisse délivrer une image à la résolution égale à la limite théorique du télescope on appelle alors cette limite la *limite de diffraction*, le télescope étant uniquement limité par la physique de la conception de ce dernier.

0.2.1. Origine et influence des aberrations

Je vais m'intéresser dans cette section à décrire la grandeur physique d'intérêt : le champ électromagnétique qui se propage dans un système optique. En particulier je vais m'intéresser à décrire l'effet des aberrations optiques sur ce champ, comment elles le modifient et comment elles dégradent la qualité de l'image qu'il fournit. On peut décrire une onde électromagnétique monochromatique dans le plan

pupille telle que :

$$\Psi_{pup}(x, y) = P(x, y)e^{j(\Phi(x, y) - \omega t)} \quad (0.1)$$

où $P(x, y)$ est la fonction pupille, $\Phi(x, y)$ la phase, (x, y) représentent les coordonnées spatiales de l'onde et ω représente la pulsation. Dans un système optique, on définit la relation entre plan pupille et plan-focal par une transformée de Fourier $\mathcal{F}$. On définit donc le champ électrique dans le plan-focal tel que

$$\Psi_{foc} = \mathcal{F}(\Psi_{pup}) \quad (0.2)$$

Dans le cas d'un imageur, le champ électrique en plan-focal est converti en figure d'intensité $U_{detecteur}$ par le détecteur qui filtre et échantillonne spatialement l'onde reçue. Le détecteur étant sensible à la moyenne temporelle du module carré du champ électromagnétique, cette réponse est calculée en effectuant la moyenne temporelle du module carré du champ électrique en plan-focal discrétisé tel que

$$U_{detecteur} = \left| \frac{1}{T} \int_0^T \Psi_{foc} dt \right|^2 \Pi_p \quad (0.3)$$

où T est le temps d'intégration du détecteur et Π_p le peigne de Dirac où p représente la taille d'un pixel du détecteur. On peut noter h la fonction qui transforme un champ électromagnétique en plan pupille (qui dépend directement de la phase Φ) en figure d'intensité sur le détecteur telle que $I = h(P, \Phi)$ où I est l'image sur le détecteur.

Plus précisément on appelle la réponse impulsionnelle du système optique au plan-focal la fonction d'étalement de point (*Point Spread Function (PSF)*) étant observé comme la réponse du détecteur (équation 0.3) à l'observation d'une source ponctuelle à l'infini d'énergie unitaire. Dans le cas d'un instrument parfait, $\Phi = 0$ et donc $\Psi_{parfait} = P(x, y)$, le télescope atteint alors la limite de diffraction (*Diffraction Limit (DL)*) qui résulte en une onde sphérique convergeant parfaitement sur le détecteur. On a alors

$$PSF_{DL} = |\mathcal{F}(P)|^2 \quad (0.4)$$

Dans le cas d'une pupille monolithique circulaire à la limite de diffraction, la PSF issue de l'équation 0.4 est la tache d'Airy présentée en figure 3.

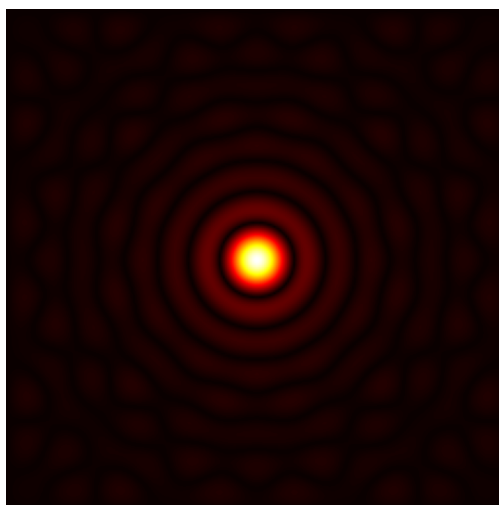


FIGURE 3. – Exemple de tâche d'Airy en échelle racine carrée.

Néanmoins, ce scénario idéal est compromis pour tout système optique réel, étant entaché d'aberrations optiques qui modifient ce résultat. En pratique, la limite de diffraction ne peut être qu'approchée.

En effet, pour tout système optique réel, des aberrations peuvent entraîner des fluctuations sur la phase et l'amplitude de l'onde incidente, qui ont pour effet de modifier la qualité d'image.

Phasage de télescope segmenté

L'image issue du télescope est impactée par des aberrations statiques et quasi-statiques, internes au télescope. La composante temporelle des aberrations est relative au temps de pause de détecteur. Une aberration dynamique réfère alors à une perturbation qui évolue plus rapidement que le temps de pause, une aberration quasi-statique correspond à une évolution plus lente que le temps de pause, enfin une aberration fixe ne varie pas dans le temps.

Dans le cas de télescope segmenté les défauts d'alignement du miroir primaire sont la source majeure de perturbation quasi-statiques. En particulier, les erreurs de positionnement et les vibrations des segments peuvent avoir un impact important sur l'image obtenue. Un bon contrôle de ces erreurs est crucial pour obtenir la qualité promise par la télescope afin que le miroir segmenté agisse de façons similaire à un miroir monolithique. Ces défauts introduisent des aberrations de faible fréquence spatiale. L'identification et la correction de ces aberrations sont appelées **phasage du télescope** et vont être étudiées dans ce manuscrit.

Aberration quasi-statiques et petaling

Dans un second temps, d'autres aberrations à l'évolution lente peuvent exister :

- Dans le cas de télescopes de grande taille, les araignées (plus communément appelé *spiders* en anglais), support mécanique de maintien du miroir secondaire, sont d'une importante largeur (jusqu'à 50 cm pour la pupille de 39 m de l'ELT). Ces araignées bloquant la propagation de la lumière créent une zone d'ombre sans signal qui impacte la formation d'image. Cette largeur vient virtuellement segmenter la pupille du télescope en différentes portions relativement indépendante les unes des autres impliquant une phase discontinue en plan pupille. Cette fragmentation cumulée aux aberrations dynamiques de l'atmosphère complique la mesure du front d'onde aberré et peut observer un défaut de phasage qui n'existe pas. La correction de ce "faux" défaut de phasage entraîne son introduction dans le système. C'est ce qu'on appelle l'effet de pétale ou encore la fragmentation de pupille. Cette fragmentation peut souffrir des mêmes aberrations qu'un télescope segmenté classique.
- Les défauts fixes d'alignement des optiques pouvant introduire des aberrations en plan-focal.
- Les défauts fixes de polissage des miroirs, faisant intervenir de très hautes fréquences spatiales.

Aberration dynamiques

Des aberrations optiques de plus haute fréquence peuvent aussi influencer sur la qualité des télescopes au sol. En effet, la lumière parvenant des sources astronomiques situées à de très grandes distances par rapport au diamètre du télescope peut être modélisée par une onde plane avant de traverser l'atmosphère terrestre. L'atmosphère, composée de couches turbulentes dues aux fluctuations de températures engendre une variation de l'indice de réfraction optique de l'air qui perturbent l'onde la traversant. Cependant, dans un cas astronomique, les perturbations d'amplitude sont négligeables par rapport aux fluctuations de phase ([Rod81]). Le paramètre r_0 , paramètre de Fried, permet de quantifier la force de la turbulence le long d'une colonne d'air. Un r_0 petit correspondant à une turbulence forte alors qu'un r_0 élevé, à une turbulence modérée. On définit l'influence de la perturbation atmosphérique sur le télescope par le rapport D/r_0 . Lorsque $D \ll r_0$ alors la résolution du télescope est limitée par la fréquence de coupure du télescope $f_c^{tel} = D/\lambda$. Dans le cas contraire elle est alors limitée par D/r_0 . Néanmoins, pour l'observation de la Terre depuis l'espace (sujet développé dans ce manuscrit) l'influence de l'atmosphère est négligeable car les couches turbulentes ne se situent pas à proximité du système imageur, mais à proximité de l'objet image :

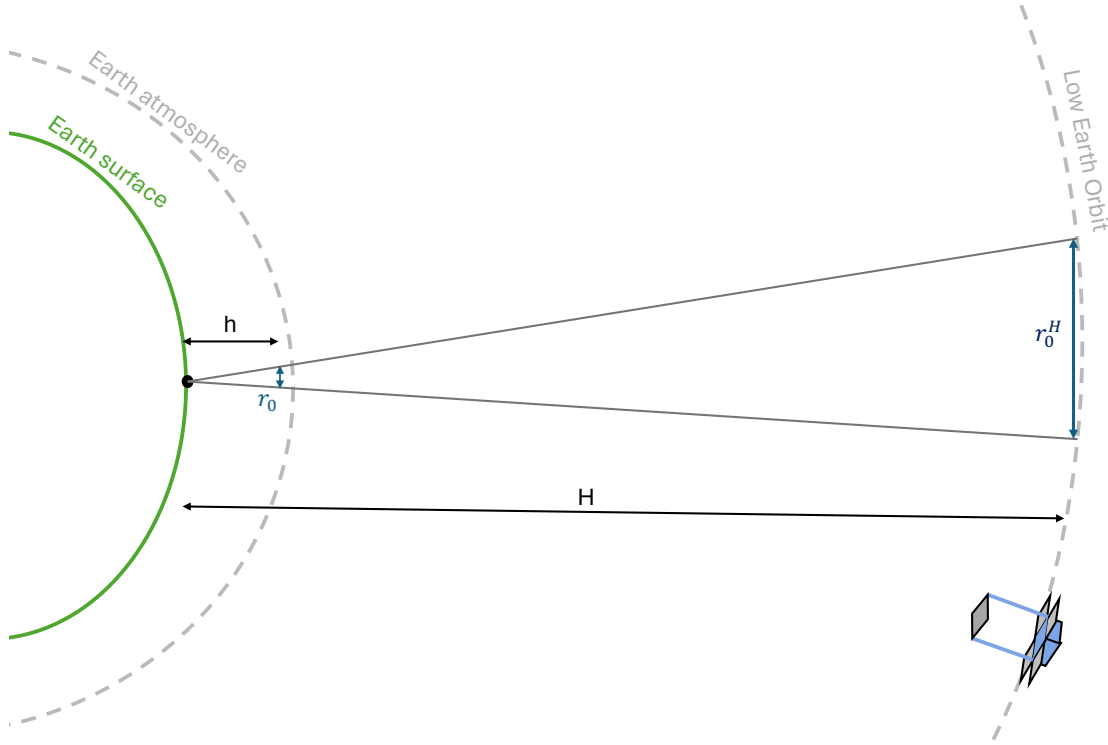


FIGURE 4. – Schéma de l'influence de l'atmosphère au niveau d'une orbite basse.

Supposons une couche turbulente r_0 à une altitude h par rapport à l'objet d'observation sur la surface de la Terre (schéma figure 4). Cette cellule r_0 se propage jusqu'à l'orbite basse terrestre H et atteint une taille caractéristique r_0^H . On peut alors facilement retrouver la taille r_0^H projetée en orbite basse telle que $r_0^H = r_0 \times \frac{H}{h}$. Dans un cas extrême, on peut supposer une couche turbulente virulente telle que $r_0 = 5 \text{ cm}$ à $h = 5 \text{ km}$. L'orbite basse se situant autour de $H = 550 \text{ km}$, on en déduit la projection $r_0^H = 5.5 \text{ m}$. On compare alors le r_0^H par rapport au diamètre D d'un télescope. Dans le cas d'un télescope de $D = 30 \text{ cm}$ de diamètre, alors $D \ll r_0^H$, l'influence de l'atmosphère terrestre sur les télescope en orbite basse est alors négligeable et le télescope est donc majoritairement limité par ses aberrations propres.

0.2.2. Impact sur la formation d'image

Les aberrations statiques, quasi-statiques et dynamiques ont un impact sur la phase du champ électrique incident : elles viennent dégrader la PSF du système optique. De façon plus générale, l'Eq.0.4 s'écrit :

$$PSF = |\mathcal{F}(\Psi_{pup})|^2 = |\mathcal{F}(Pe^{i\Phi})|^2 = h(P, \Phi) \quad (0.5)$$

On décrit la **Fonction de Transfert Optique** (FTO ou OTF) comme la transformée de Fourier de la PSF :

$$FTO = \mathcal{F}(PSF) = h(\tilde{P}, \Phi) \quad (0.6)$$

La FTO étant la transformée de Fourier (modélisée par $\tilde{h}$) d'un module carré d'une transformée de Fourier, elle peut alors s'écrire comme la fonction d'autocorrelation du champ électrique en plan pupille :

$$FTO = \mathcal{F}(|\mathcal{F}(\Psi_{pup})|^2) = \Psi_{pup} \otimes \Psi_{pup} \quad (0.7)$$

où $\otimes$ représente l'opération de convolution. Cette fonction de transfert est aussi impactée par les aberrations de la phase et se comporte alors comme un filtre passe bas avec une fréquence de coupure $f_c^{abb} < f_c^{tel}$. On peut observer la FTO théorique d'un télescope parfait et la FTO d'un télescope aberré en figure 5.

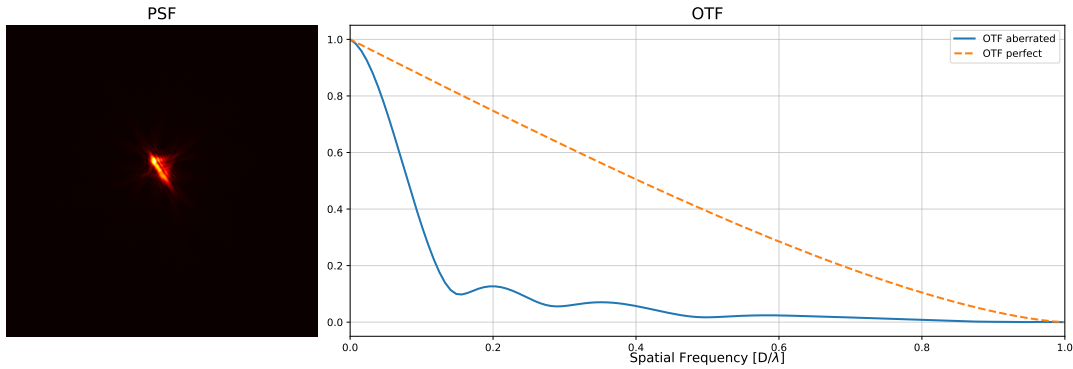


FIGURE 5. – Exemple de PSF aberrée (gauche) et de la FTO correspondante (figure de droite en trait plein), comparée à la FTO théorique du télescope parfait (figure de droite en pointillé).

On observe en figure 5 une PSF sur la figure de gauche et une coupe de la FTO à droite. On note une nette perte des fréquences spatiales dues aux aberrations de la PSF, dégradant ainsi la qualité d'image.

0.2.3. Analyse du front d'onde

L'impact des aberrations de la phase détériore la qualité de l'image produite par le télescope. On s'intéresse donc à quantifier ces aberrations afin de les corriger dans un second temps. Dans le cadre du phasage d'un télescope, ces aberrations en plan pupille proviennent du piston, du tip et du tilt, correspondant respectivement à une élévation du miroir par rapport à la position nominale, et une inclinaison horizontale et verticale par rapport à cette position. Dans cette section le focus est aussi mentionné, correspondant à un déplacement du détecteur dans le plan focal.

Observation du front d'onde

Le front d'onde perturbé étant limitant pour l'observation d'objet astronomique ou de la Terre à haute résolution angulaire, il se doit d'être analysé puis corrigé. L'observation du front d'onde peut reposer sur l'optique géométrique pour traduire le front d'onde en intensité ou alors observer directement les figures d'interférence sur le détecteur en plan-focal, mélangeant alors phase et intensité du champ électromagnétique avant de les convertir en amplitude. Indépendamment de la méthode choisie, les intensités reçues sont ensuite traitées et traduites pour reconstruire le front d'onde incident et le corriger.

Méthodes plan-focal

On propose d'étudier les méthodes en utilisant un analyseur de surface d'onde : le détecteur de la voie science. C'est une méthode plan-focal qui utilise directement l'image produite par le télescope pour remonter aux aberrations de la phase sans nécessiter de voie optique dédiée.

Le plan-focal : Information de la phase mêlée à l'objet et ambiguïté de signe

Les méthodes plan-focal consistent en la résolution du problème inverse de l'identification de la phase à partir d'une carte d'intensité issue du détecteur. La relation non linéaire entre phase et intensité (fonction h équation 0.5) rend le problème complexe : identifier la phase Φ à partir des intensités de l'image en plan-focal consiste en un problème inverse non linéaire pouvant être mal posé. En effet l'un des facteurs limitant de l'Analyse de Front d'Onde Plan-Focal (AFOPF) est la non-unicité de la solution due à la relation entre la FTO et la phase dans la pupille. Pour rappel l'équation 0.7 définit la FTO comme l'autocorrélation du champ électromagnétique dans le plan pupille. Si la pupille P est à symétrie de révolution, alors une phase $\Phi(x, y)$ et une seconde phase $\Phi'(x, y) = -\Phi(x, y)$ donnent rigoureusement la même FTO : $\tilde{h}(P, \Phi) = \tilde{h}(P, \Phi')$. La décomposition en partie paire et impaire de la phase indiquent une indétermination sur la partie paire de la phase (cette indétermination sera développée partie 1.2.2). Les aberrations de phase paire telle que le focus ne peuvent alors physiquement pas être identifiés à partir d'une unique image car ils forment une image identique en plan-focal.

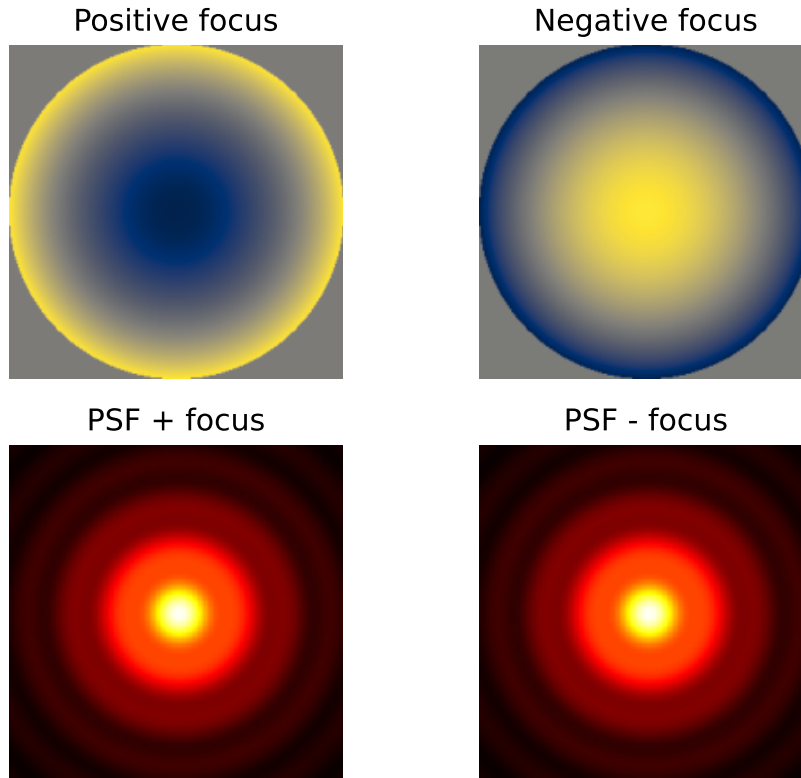


FIGURE 6. – Illustration de l'influence d'un focus positif et négatif sur l'image de la PSF sur le détecteur.

Par exemple, la figure 6 montre l'impact d'un focus positif ou négatif en plan pupille sur la PSF formée sur le détecteur en plan focal. Les 2 PSFs issues des aberrations de la pupille au dessus sont rigoureusement les mêmes alors que le focus est différent. Ce facteur extrêmement limitant contraint l'AFOPE. Il est donc nécessaire d'apporter de l'information supplémentaire au système de résolution du problème inverse afin de lever l'indétermination de signe et de retrouver la partie paire de la phase.

Diversité de phase

La méthode la plus répandue pour lever cette indétermination est la diversité de phase (*Phase Diversity* (PD)) ([Gon82]) : elle consiste à utiliser deux images, dont l'une diffère de l'autre d'une phase connue Φ_d . On utilise très généralement le defocus comme phase connue, pour sa facilité de mise en oeuvre (déplacer une caméra ou un détecteur). Un schéma du fonctionnement de la diversité de phase peut être trouvé en figure 7.

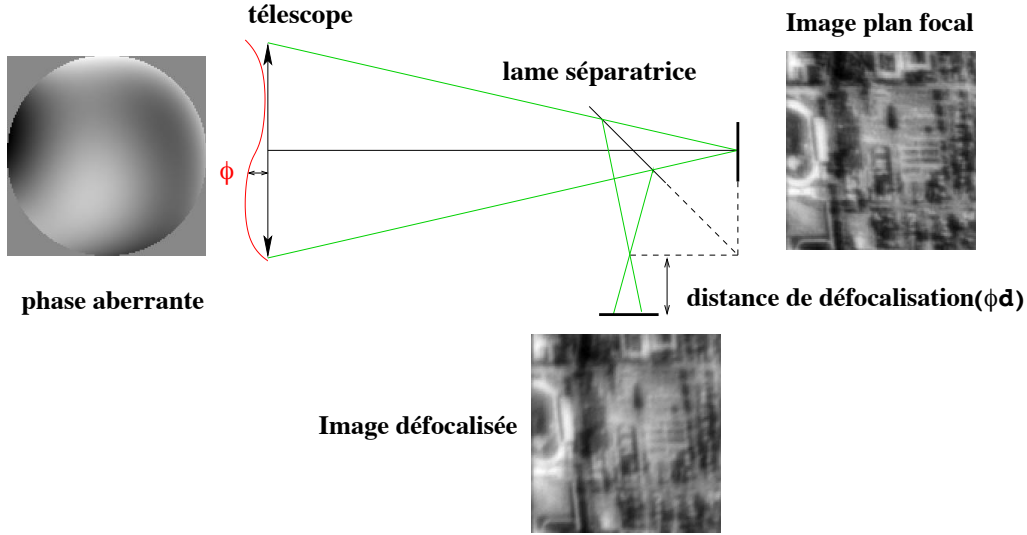


FIGURE 7. – Schéma de la mise en place de la diversité de phase. Extrait de la thèse d'Amandine Blanc.

Cette diversité va apporter une information supplémentaire sur la direction des aberrations. Dans le cas de l'observation d'un point source on a donc :

$$i_f = PSF_f = |\mathcal{F}(Pe^{j\Phi})|^2 \quad (0.8)$$

$$i_d = PSF_d = |\mathcal{F}(Pe^{j(\Phi+\Phi_d)})|^2 \quad (0.9)$$

Les deux images étant reliées par une relation non-linéaire, elles permettent de lever l'ambiguïté de signe. Le principe de la diversité de phase est alors de minimiser l'erreur quadratique entre un modèle de PSF et les images focalisée et défocalisée :

$$J(\Phi) = \|i_f - h(P, \Phi)\|^2 + \|i_d - h(P, \Phi + \Phi_d)\|^2 \quad (0.10)$$

On cherche alors la phase qui minimise la quantité J

$$\hat{\Phi} = \operatorname{argmin}_{\Phi} J(\Phi) \quad (0.11)$$

où $\hat{\Phi}$ représente la quantité estimée de Φ . Cependant, comme tout problème inverse non linéaire, la solution n'est pas évidente et nécessite une optimisation itérative de la métrique ci-dessus. C'est pourquoi la diversité de phase a été utilisée dans des cas où il n'était pas possible d'ajouter des éléments spécifiques pour effectuer la mesure du front d'onde, et donc pour caractériser des aberrations de front d'onde qui évoluent doucement dans le temps, notamment pour le phasage de télescopes segmentés ([Lof+98], [CTP00]) et la caractérisation d'aberrations lentes d'un système

optique ([LDS98], [KAD94], [Lee+97]) ou encore pour la restauration d'image avec une principale application pour l'imagerie solaire ([LS94]).

Les méthodes d'optimisation pour la minimisation du critère et d'utilisation de la diversité de phase utilisés dans ce manuscrit seront développées partie 3.6.1.

Image Sharpening

L'*Image Sharpening* (IS) consiste en l'optimisation d'une métrique directe de l'image à partir d'une séquence d'image ([MB74]). L'intensité totale ou encore la netteté de l'image est calculée et maximisée dans un processus itératif de correction du miroir primaire par optique active afin de trouver la correction optimale.

$$\hat{\Phi} = \operatorname{argmax}_{\Phi} C(h(\Phi)) \quad (0.12)$$

où $C(x)$ est une fonction représentant la métrique d'intensité ou de contraste à partir d'une image. Cette méthode implique des implémentations matérielles plus simples que la détection directe du front d'onde mais nécessite une correction active des composants de l'imageur à chaque itération. Le processus de mesure indirecte est par nature itératif et est généralement appliqué dans des domaines tels que la microscopie ou les communications lasers. Il n'est généralement pas applicable aux applications où les aberrations changent rapidement comme la correction de la turbulence atmosphérique mais peut être appliqué pour la correction d'aberrations quasi-statique et donc pour le phasage de télescope segmenté. De même que la diversité de phase, les méthodes d'optimisation pour la minimisation du critère et d'utilisation de la diversité de phase seront développée partie 3.6.2.

Ces approches proposent une mesure directe du front d'onde à partir de l'image produite par le télescope, n'impliquant donc que très peu de contrainte matérielle supplémentaire. Néanmoins, des systèmes à plus haute fréquence existent sur les télescopes actuels, se basant se l'optique géométrique et une voie optique dédiée.

Optique Adaptative L'Optique Adaptative (*Adaptive Optics* (AO)), système optique permettant d'analyser et de corriger le front d'onde en quasi temps réel est caractérisé par 3 composants principaux : Le miroir déformable (*Deformable Mirror* (DM)), l'*Analyseur de Surface d'Onde* (ASO) et le Calculateur Temps Réel (*Real Time Computer* (RTC)). Le schéma de principe de l'OA est présenté en figure 8. L'onde perturbée issue du télescope est réfléchi sur le DM dont la forme est adaptée à chaque instant. Après réflexion, idéalement, la surface d'onde corrigée est proche d'une onde plane. La lumière est ensuite séparée en deux voies, l'une pour former la grandeur d'intérêt (une image, un spectre, une mesure de flux...) et l'autre pour mesurer les perturbations de la surface d'onde par l'ASO.

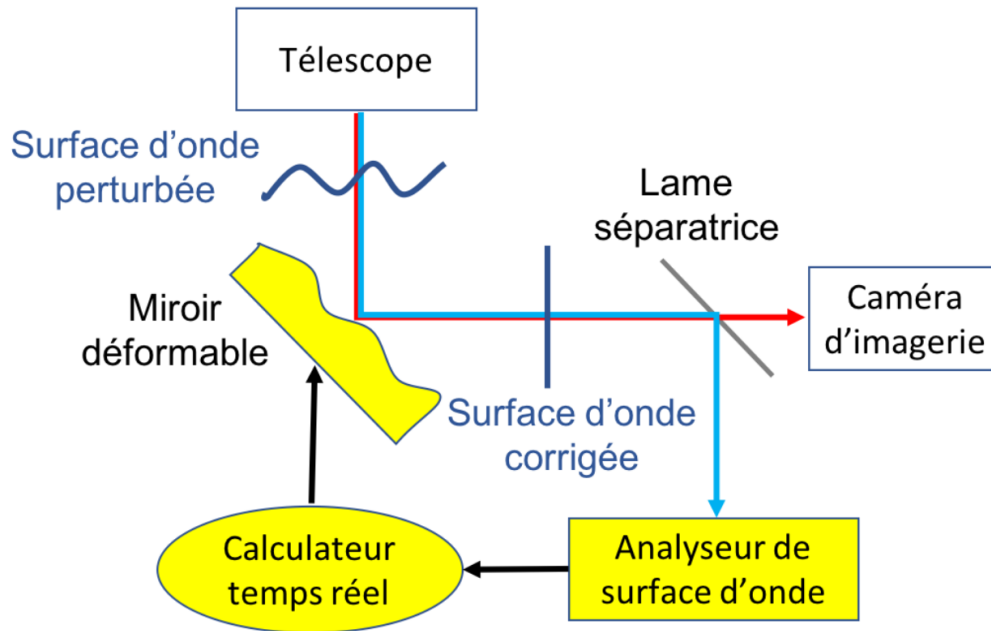


FIGURE 8. – Schéma du système d'optique adaptative. Extrait de [RF22].

- **ASO** : Permet de convertir la phase en figure d'intensité sur le détecteur. La mesure n'étant pas directe, un reconstituteur est nécessaire pour analyser ces mesures et les convertir en front d'onde. L'ASO permet donc, dans la limite de sa sensibilité, de linéariser autour de son point de fonctionnement la relation entre phase et intensité afin de gagner en fréquence de calcul. Deux senseurs de front d'onde vont être brièvement évoqués dans le paragraphe suivant.
- **RTC** : Le RTC comprend l'algorithme de reconstruction à partir du détecteur de l'ASO mais aussi le hardware permettant d'effectuer ce calcul. Le RTC se doit d'être performant afin de minimiser le temps de calcul pour réduire l'erreur temporelle (fréquence de calcul inférieure à la fréquence d'évolution des perturbations, impliquant une correction avec k -trames de retard).
- **DM** : Il permet de corriger le front d'onde à partir du front d'onde reconstruit en lui appliquant la déformation opposée des observations afin de minimiser les résidus de front d'onde.

On peut alors citer les deux ASOs les plus étudiés à ce jour :

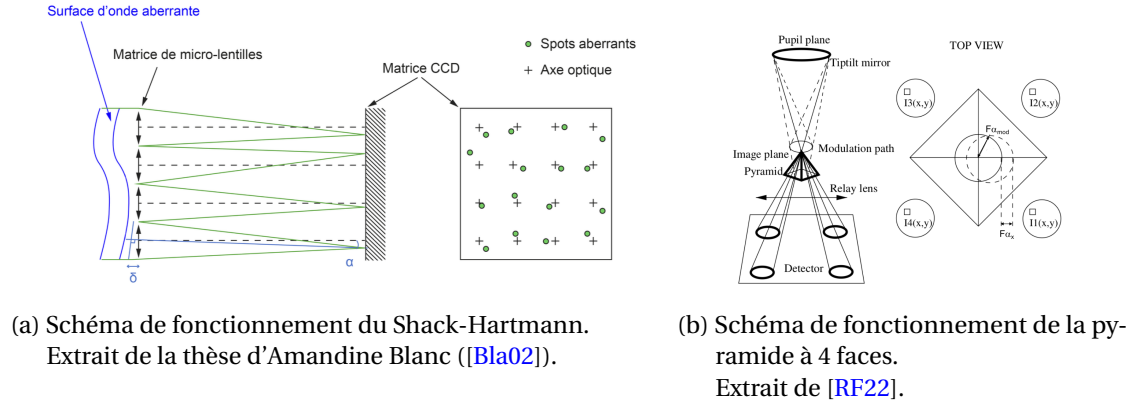


FIGURE 9. – Exemple d'ASOs. Extrait de [RF22].

Discrétiser spatialement le front d'onde dans la pupille en éléments suffisamment petits pour supposer que le front d'onde local n'est impacté uniquement que par des pentes de tip-tilts, et mesurer ces pentes par la position d'un spot local, c'est toute l'idée du **Shack-Hartmann (SH)** ([Sha71]), figure 9 (a)). Une matrice de micro-lentille discrétise spatialement et localement le front d'onde formant un champ de spots sur un détecteur CCD. Les pentes du front d'onde agissent localement comme un tip-tilt pour le spot et le dévie donc par rapport à une position de référence (pour un front d'onde plat). Pour chaque sous-pupille la mesure de la pente locale peut être obtenue à partir du centre de gravité du spot.

L'analyseur de surface d'onde Pyramide ([Rag96], figure 9 (b)), plus récent et utilisé dans les instruments dernière générations (HARMONI à l'ELT ([Tha+16])) est une pyramide en verre de 4 faces généralement qui vient séparer le champ électromagnétique dans le plan-focal. Chaque face de la pyramide renvoie 4 images séparées de la pupille, chaque image étant filtrée par une des faces. On peut ensuite remonter au front d'onde incident en multipliant l'image des sous-pupilles par la pseudo inverse de la matrice de calibration préalablement construite.

L'optique adaptative est la méthode connue et reconnue pour son efficacité sur ciel et permet de corriger les dégradations de la phase à très haute fréquence afin de se rapprocher de la limite de diffraction promise par le dimensionnement du télescope. Néanmoins elle requière une instrumentation supplémentaire et donc voie optique dédiée. Cette implémentation est nécessaire sur les télescopes au sol, mais plus contraignant pour les télescopes spatiaux. Néanmoins ces méthodes ont prouvées leur efficacité dans l'analyse du front d'onde de manière générale notamment dans la correction de la turbulence atmosphérique. On propose maintenant de faire l'état de l'art des implémentations de méthode d'analyse de surface d'onde pour le phasage des télescopes segmenté.

0.2.4. Phasage des télescopes segmentés

Le phasage de télescope segmenté consiste en l'identification et la correction du piston et du tip-tilt entre chaque segment d'un télescope. Il n'est généralement pas géré par les systèmes d'optique adaptative car les erreurs d'alignement peuvent être directement corrigé par le miroir primaire (optique active). Cela permet de ne pas saturer le DM avec des pentes trop importante afin lui assurer une plus grande dynamique. Par ailleurs les analyseurs de l'optique adaptative supposent souvent la continuité du front d'onde dans la mesure et la reconstruction, et sont peu sensibles à des défauts de front d'onde discontinus. Par exemple, plus la pyramide est modulée (afin d'augmenter sa dynamique) plus elle perd en sensibilité aux discontinuités de la phase (voir figure 10). A contrario, certaines solutions spécifiques utilisant une sous-pupille de SH situé à cheval sur deux segments permettent d'effectuer une mesure pertinente.

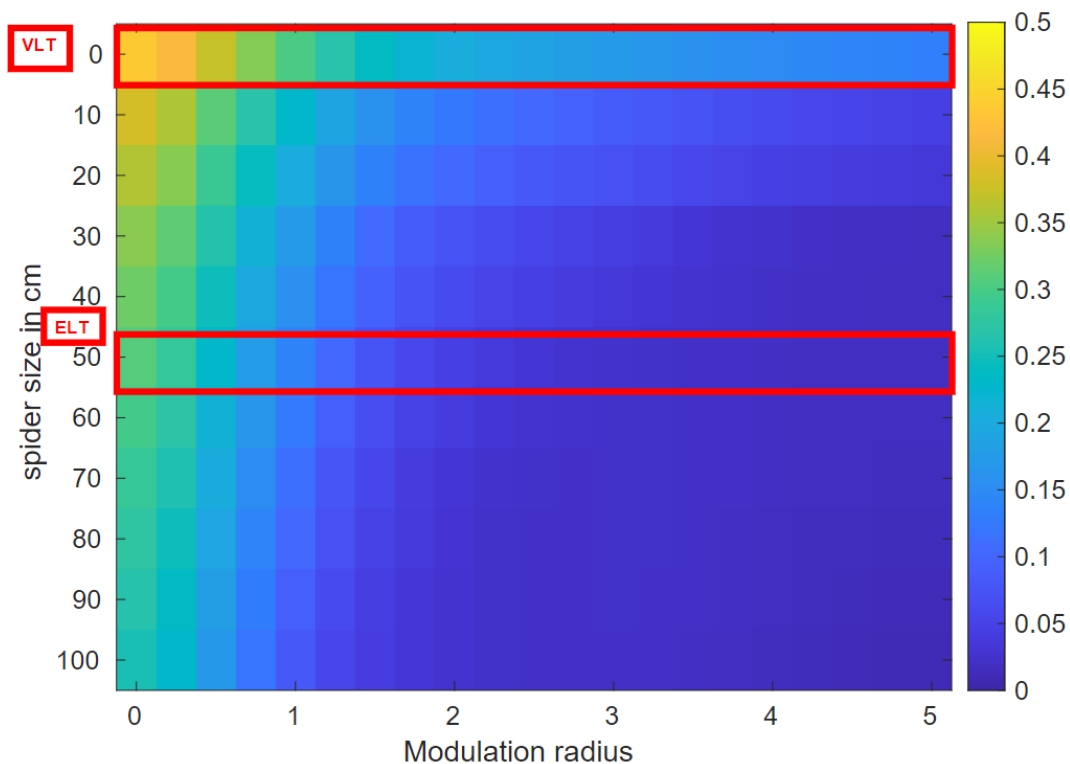


FIGURE 10. – Sensibilité de la Pyramide modulée à une discontinuité de phase (largeur du spider). Sans modulation la pyramide reste très peu sensible aux discontinuités. Extrait de la thèse de N.Levraud.

Phasage de télescope segmenté au sol

Les grands télescopes au sol segmentés doivent donc s'appuyer sur une autre approche pour identifier les défauts d'alignement de leurs miroirs primaire. Dans cette

section on propose d'expliciter les méthodes employées pour 4 télescopes au sol.

Keck Le miroir primaire du télescope Keck est constitué de 36 segments hexagonaux qui s'alignent pour former une surface optique de 10 mètres de diamètre. Chaque segment, de 90 cm de côté, est contrôlé par trois actionneurs mécaniques situés à l'arrière du miroir. Chacun des 84 bords inter-segment comptent 2 capteurs capacitifs, qui détectent les changements dans les hauteurs relatives locales des segments. Une fois le miroir principal phasé à l'aide des capteurs capacitif, la caméra de phasage prend le relais. Cette caméra de phasage fonctionne comme un Shack-Hartmann positionné aux bords inter-segment des miroirs et peut être appelée capteur de bord (*Edge Sensors* (ES)). Les détails de la figure de diffraction formée par les sous-ouvertures étant sensibles à la hauteur du pas entre deux segments, on peut alors en extraire les informations pour le phasage des segments à partir des intensités de l'image. Le phasage du miroir primaire utilisant la caméra de phasage permet une stabilisation du phasage avec un pas inférieur à 100 nm.

ELT L'Extremely Large Telescope, le plus gros oeil sur l'univers jamais construit avec un diamètre de presque 40 mètres. Un tel diamètre oblige un miroir primaire segmenté, avec plus de 790 segments d'1.4 mètre de diamètre. Ces segments hexagonaux viennent bâtir le miroir primaire du télescope et représentent un réel défi pour le phasage. La stratégie de phasage repose sur le même concept que le Keck. 2 Edges sensors sont positionnés par arrête de chacun des 6 bords (le miroir est sur-échantillonné), ce qui représente + de 4000 edges sensors au total!

LBT Le Large Binocular Telescope est un télescope composé de deux miroirs primaires de 8.4 mètres de diamètre produisant une surface collectrice équivalente à celle d'un télescope de 11.8 mètres. L'alignement des deux miroirs se fait à l'aide d'une PHASEcam ([Mai+20]). La PHASEcam, composée d'un détecteur PICNIC ([Koz+00]), réimage la pupille du LBT. Elle fonctionne alors comme un détecteur de frange : en effectuant la transformée de Fourier de l'image au plan pupille (figure de frange), on peut repérer alors le piston et le tip-tilt entre les miroirs : l'amplitude de la transformée de Fourier donne une mesure du tip-tilt, tandis que la phase permet d'identifier le piston.

GMT Le *Giant Magellan Telescope* (GMT), un télescope de 25 mètres de diamètre est composé de 7 miroirs primaires de 8.4 mètres, propose une stratégie de co-phasage qui repose en plusieurs étapes. La particularité de ce dernier est que le miroir primaire et le miroir secondaire sont segmentés.

- **Le phasage initial.** Le phasage du GMT diffère des précédentes méthodes car la taille des segments et de l'espace inter-segment est plus élevée que dans les télescopes segmentés ci-dessus. Les edges sensors ne permettent pas une stabilité du phasage face aux perturbations extérieures (vent, gravité, effet thermiques).

2 capteurs de bords sont quand bien même utilisés afin de comparer les erreurs de phasage avec la méthode utilisée. Le phasage est finalement géré par le système *Acquisition Guiding and Wavefront Sensing System* ([Qui+18]). Il est composé d'un **Dispersed Fringe Sensor** (DFS) et d'un Shack-Hartmann pouvant corriger les erreurs de pistons jusqu'à $40\mu m$. Au cours d'une observation, il maintient le télescope en phase à une fréquence relativement faible les erreurs étant quasi-statiques.

- **Le phasage résiduel et correction des aberrations** se fait par deux modes d'optique adaptative et d'un ASO pyramide (*Natural Guide-star AO* ([Hin+10]) et *Laser Tomography AO* ([Dam+12])).

Phasage du James Webb Space Telescope

Pour les télescopes spatiaux tels que le JWST, la stratégie de mise en phase est divisée en plusieurs étapes ([Act+12], [Act+22]). Dans un premier temps, 1 semaine a été nécessaire pour le déploiement du miroir primaire. Ensuite viennent les étapes de phasage :

- **L'identification de l'emplacement du segment** pour identifier la position grossière de chaque segment. L'ajustement du miroir secondaire permet d'amener le télescope à un focus maîtrisé.

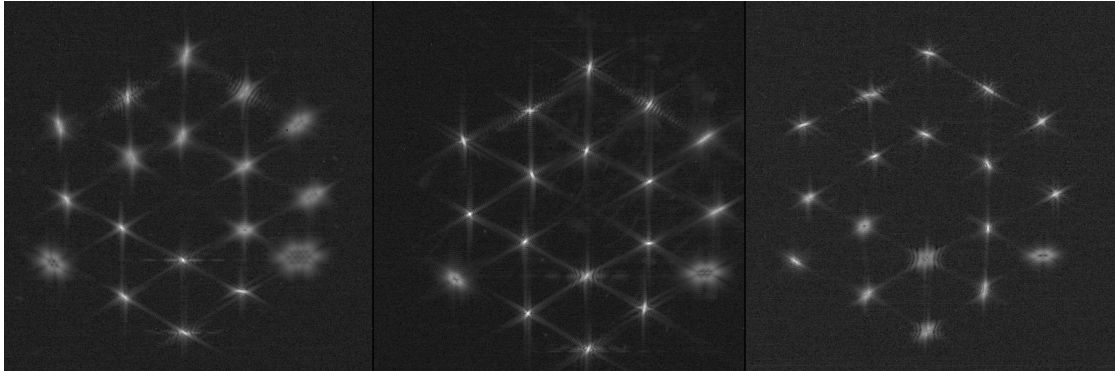


FIGURE 11. – Illustration des simulations de l'impact sur les sous-PSFs de l'ajustement du focus du JWST. Extrait de Acton [Act+22].

Les segments sont ensuite alignés par une méthode d'identification assez simple : en appliquant une inclinaison sur un des segment du miroir, on repère la différence entre l'image initiale acquise sur la NIRcam (instrument capturant des image dans le proche infrarouge entre $0,6\mu m$ et $5\mu m$) et celle une fois l'inclinaison effectuée.

- **Le co-phasage grossier : superposition des sous-PSFs.** L'intérêt maintenant est de faire superposer toutes les sous-PSFs sur le détecteur, une étape simple une fois les PSFs identifiées mais qui comporte des limites entre la précision des déplacements large des miroirs et le re-centrage des actionneurs de précisions.

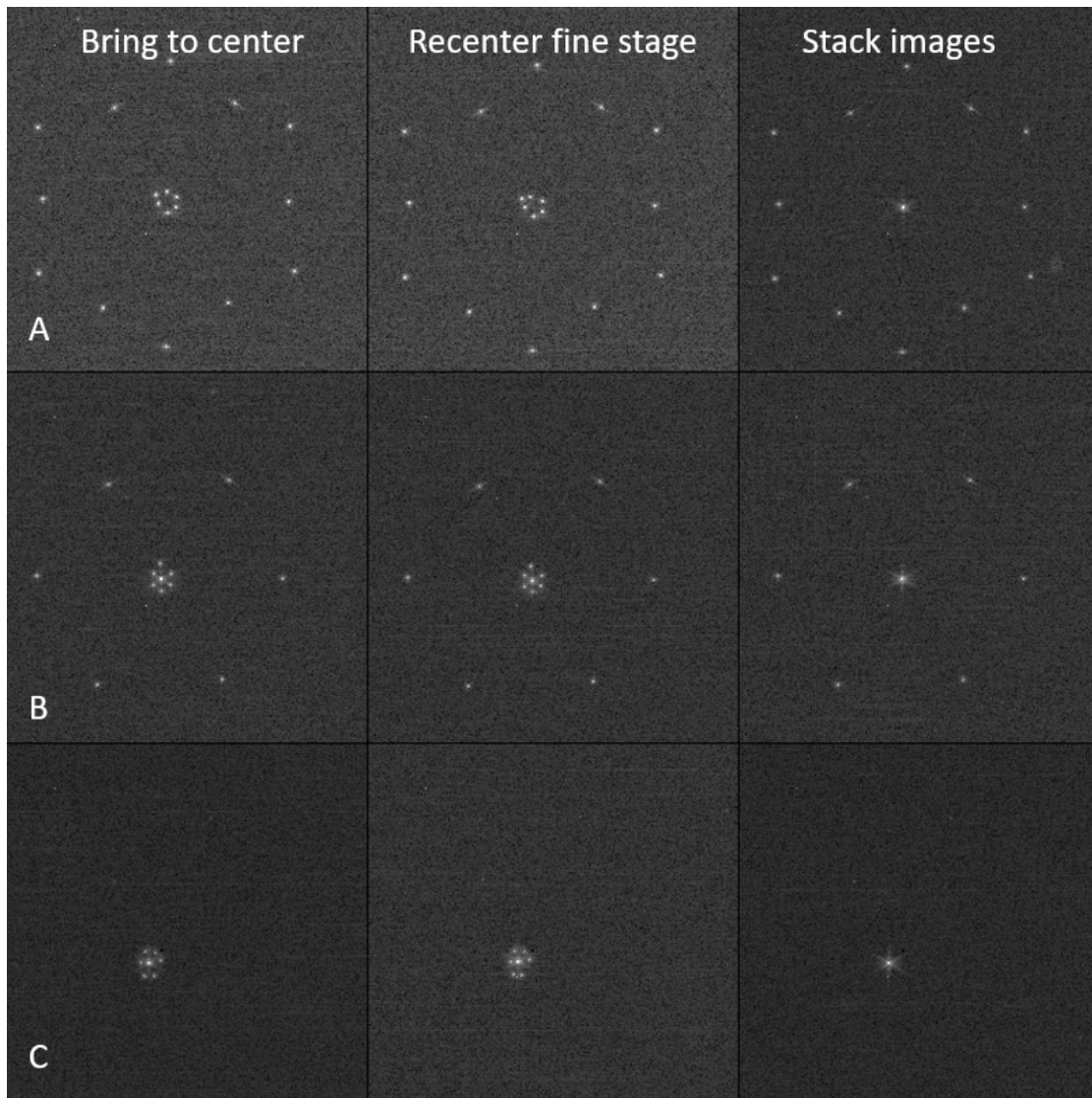


FIGURE 12. – Illustration du centrage itératif des sous-PSFs du JWST, corrigeant le tip et le tilt de chacun des segments. Extrait de Acton [[Act+22](#)].

Les erreurs de piston importantes (supérieure à 300 nm) entre les segments sont ensuite estimées par la détection de franges dispersées. Le DFS vient analyser les franges de la structure du spectre de l’image.

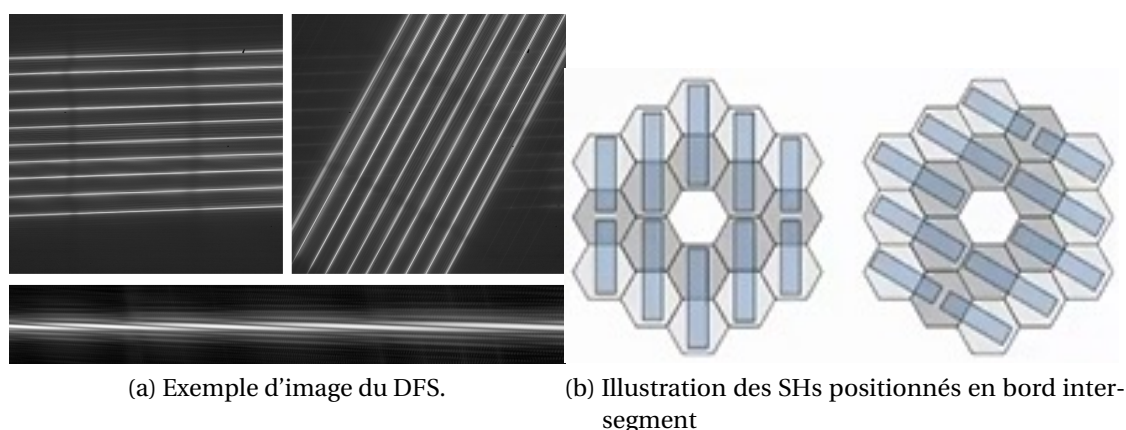


FIGURE 13. – Illustration des méthodes de correction des pistons du JWST. Extrait de Acton [Act+22].

En sélectionnant plusieurs longueurs d'onde le piston peut être retrouvé de manière très précise. Il est utilisé dans la NIRcam et imagé à l'aide d'un un SH entre chaque bord inter-segments du miroir primaire, permettant de mesurer les erreurs de piston allant jusqu'à $350\ \mu m$.

- **Correction de l'erreur de mise en phase fine des 18 segments.** Enfin, la mise en phase fine du télescope est réalisée à l'aide d'un algorithme de diversité de phase à différents defocus ([Dea+06]).

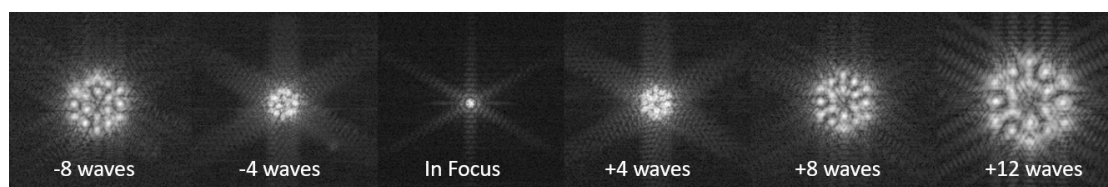


FIGURE 14. – Illustration de la diversité de phase à 6 longueurs d'onde appliquée au JWST. Extrait de Acton [Act+22].

À l'issue de ce long processus, le miroir du JWST est finement co-phasé avec un front d'onde résiduel de l'ordre de $50\ nm\ RMS$.

Ces étapes de phasage forment un processus extrêmement long, particulièrement lorsqu'il s'agit d'un télescope situé à des millions de kilomètre de la Terre. Scott Acton (responsable du phasage du JWST) donne l'allégorie que le télescope a été phasé à la vitesse ou l'herbe pousse.

0.2.5. Limite de l'analyse du front d'onde pour le phasage de télescope pour le NewSpace

Cependant, le point commun entre toutes ces méthodes de phasage est qu'elles requièrent toutes une instrumentation supplémentaire : caméra de pupille, Shack-Hartmann ou encore senseur capacitif.

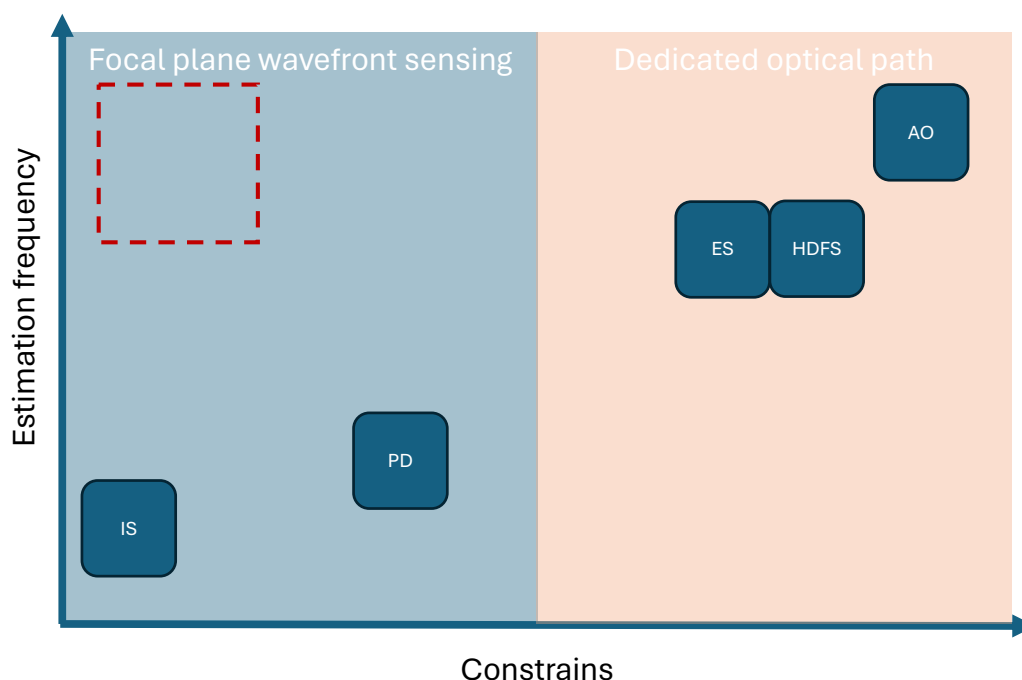


FIGURE 15. – Illustration des différentes méthodes d’analyse du front d’onde (IS pour *Image Sharpening*, PD pour *Phase Diversity*, ES pour *Edge Sensor*, HDFS pour *Holographic Dispersed Fringe Sensor*, AO pour *Adaptive Optics*).

Le schéma présenté en figure 15 permet de comparer, à une échelle arbitraire, les méthodes d’analyse de surface d’onde citées ci dessus par le rapport entre leur fréquence de calcul et leurs contraintes d’implémentation, critère majeur dans le cas des télescopes spatiaux. De plus le phasage post-déploiement de ces derniers est une étape très chronophage qui demande ensuite un processus actif de maintien en temps quasi-réel. On voit ici alors se dessiner une contrainte majeure : qu’en est il du NewSpace? Chercher à réduire le volume des plateformes et le coût des satellites revient à minimiser tout système additionnel embarqué, et donc, de phasage. L’idéal serait d’employer un méthode promettant des performances correspondant à la zone encadrée en rouge sur le schéma ci-dessus : une méthode avec un minimum de contrainte supplémentaire garantissant un phasage relativement rapide. On doit se tourner vers les méthodes plan-focal, prenant directement comme entrée l’intensité de l’image sur le détecteur, mais alors la contrainte de l’ambiguïté de signe et la non-linéarité entre intensité et phase rend le problème inverse itératif et chronophage. Ce qui rendrait la tâche plus aisée pour s’affranchir de tout calcul supplémentaire à bord de la plateforme du satellite et d’effectuer cette optimisation au préalable pour n’effectuer que les calculs opérationnels une fois en orbite. Cette solution, l’Intelligence Artificielle et notamment le *Machine Learning* (ML) permet de l’exploiter. En déplaçant la charge de calcul de l’entraînement des algorithmes au sol, le ML une fois embarqué effectue uniquement l’inférence des différents paramètres

de phasage à partir de l’image plan-focal. C’est cette application qui est présentée dans ce manuscrit.

0.3. Ouverture de la mesure du front d’onde par Machine Learning

Dans un premier temps on s’intéresse de manière plus générale à l’application du Machine Learning pour la mesure du front d’onde. Plus spécifiquement, un développement sur les algorithmes de Deep Learning sera effectué partie 2.

0.3.1. Le ML comme reconstituteur à partir d’ASO

Le Machine Learning a été particulièrement étudié ces dernières années dans le domaine de l’optique adaptative, afin d’estimer ses performances par rapport aux analyses de front d’onde classique. Les réseaux de neurones peuvent par exemple être utilisés en sortie d’ASO pour reconstruire le front d’onde. L’une des premières applications de le ML sur ciel a été explorée par Osborn ([[Os14](#)]), implémentant une architecture appelée *Multi-Layer Perceptron* (MLP) pour de la reconstruction tomographique de front d’onde. Ce même type d’architecture a été utilisé en 2006 par Guo ([[Guo06](#)]) et en 2019 par Xu ([[Xu19](#)]). Guo utilisant l’architecture MLP pour reconstruire le front d’onde aberrant à partir des pentes en sortie du Shack-Hartmann pour l’analyse des aberrations de l’oeil humain et Xu pour compenser les erreurs d’alignement entre le SH et le DM en plus des perturbations atmosphériques. Plus tard en 2018, Swanson ([[Swa18](#)]) propose aussi de reconstruire le front d’onde à partir des pentes du SH mais cette fois à partir de plusieurs frames de pentes. Il utilise une architecture de réseau de neurones convolutif pour reconstruire et prédire le front d’onde. Pour l’analyseur pyramide, on peut citer Alvarez ([[ASB08](#)]) qui reconstruit le front d’onde à partir d’une pyramide à 3 faces et la reconstruction de caractéristiques des images d’entrée (ΔI de la sortie de pyramide). Les travaux d’ Escobar ([[EV21](#)]) montrent une comparaison détaillée entre les différents réseaux pour la reconstruction du front d’onde avec SH et Pyramide.

0.3.2. Analyse du front d’onde plan-focal par ML

Pour entrer plus dans le sujet des travaux de la thèse, on propose ici de recenser des travaux pour l’analyse de front d’onde en plan focal par Deep Learning. L’utilisation de techniques de Deep Learning a été proposée dans des études antérieures comme un outil puissant pour l’AFOPE, visant souvent des niveaux de performance similaires ou supérieurs à l’état de l’art ([[Ros22](#); [Her18](#); [PF18](#)]). La majorité des méthodes de Deep Learning pour l’AFOPE utilise une seule image défocalisée, comme le démontrent des études telles que [[Ros22](#); [Wan21](#); [RTF22](#); [AOE20](#); [PF18](#); [Orb21](#)]. D’autres ont utilisé un préconditionneur optique (surexposition, défocalisation, diffusion) pour améliorer les performances ([[Nis19](#)]). Enfin, d’autres encore ont utilisé une paire

d’images in et out-of-focus à la manière d’une diversité de phase pour identifier les aberrations de hauts ordres à partir d’une paire de PSF ([Que+22]).

Ces travaux ont montré d’excellentes performances pour l’analyse de front d’onde à l’aide d’architectures de Deep Learning profondes. Des architectures complexes et profondes sont alors utilisées telles que Xception et Inceptionv3. Ces architectures ont par exemple été employées par [PF18], [RTF22], [AOE20], [Nis+19], et [Orb+21]. D’autres architectures plus populaires, telles que ResNet-50 et VGG-16, ont également été testées par [PF18] et [Orb+21]. Des architectures telles que EfficientNet-B4 (utilisée par [Que+22]), Bi-GRU pour les données séquentielles (utilisée par [Wan+21]), ou encore des MLPs (utilisés par [Ros+22]) ont également démontré leur efficacité pour la détection de front d’onde dans le plan-focal.

Dans ces études, deux types de détection de front d’onde dans le plan-focal ont été identifiées. La majorité d’entre elles se concentrent sur l’identification aberrations de hauts ordres pour un imageur à pupille monolithique ([PF18], [AOE20], [Nis+19], [Orb+21], et [Que+22]), tandis que d’autres visent à phaser une pupille segmentée en identifiant les modes de bas ordres ([Wan+21] se concentre sur le Piston et le Tip-Tilt, [Ros+22] et [RTF22] se concentrent sur le Piston uniquement).

Article	Output	Method	Pupil	Architecture
[Ros+22]	Piston	1 defocused PSF	Segmented	Fully connected
[Her+18]	PSF parameters	PSF	Monolithic	CNN
[PF18]	Haut ordres	1 defocused PSF	Monolithic	Xception, ResNet-50
[Wan+21]	Piston and Tip-Tilt	1 defocused PSF	Segmented	Bi-GRU
[RTF22]	Piston	1 defocused PSF	Segmented	Xception
[AOE20]	Haut ordres	1 defocused PSF	Monolithic	Xception
[Orb+21]	Haut ordres	1 defocused PSF	Monolithic	Xception, VGG-16
[Nis+19]	Haut ordres	Préconditioning	Monolithic	Xception
[Que+22]	Haut ordres	in and out-of-focus PSF	Monolithic	EfficientNet-B4

TABLEAU 1. – Résumé des études sur l’utilisation des techniques de Deep Learning pour l’AFOPE.

Le principal défaut de ces études est que l’architecture des réseaux de neurones de l’état de l’art nécessite des centaines de milliers voir des centaines de millions de paramètres (par exemple, ResNet-50 et VGG-16 nécessitent respectivement 23 millions et 138 millions de paramètres). Notre objectif, outre le traitement d’erreurs de segmentation spécifiques, est de minimiser le nombre de paramètres afin d’optimiser le temps d’estimation. C’est pourquoi nous devons explorer de nouvelles architectures adaptées qui réduiront les dimensions des architectures.

0.4. Contribution des travaux de thèse

Les travaux de cette thèse s’appuient donc sur l’utilisation du Deep Learning pour l’analyse front d’onde plan-focal appliqué au phasage de télescope spatial segmenté.

0.4.1. Deep Learning pour le phasage en plan-focal

On propose d'explorer les performances du Deep Learning pour le phasage complet d'un télescope segmenté. En s'inscrivant principalement dans le contexte d'un télescope segmenté spatial pour le NewSpace, on souhaite donc mettre l'accent sur l'accès rentable à l'espace et introduire de nouvelles contraintes telles que l'allocation d'un volume limité. Il est donc nécessaire de développer des méthodologies innovantes, telles que celles présentées dans ce manuscrit. Le Deep Learning s'inscrit idéalement dans ce contexte, permettant de lier à la fois efficacité calculatoire et analyse de front d'onde non invasive et sans chemin optique dédié : en tirant parti d'une seule image du plan-focal, nous estimons et corrigeons (dans un scénario de correction idéale) par la suite les aberrations de phase de bas ordre, le piston et le tip-tilt des segments du miroir primaire en utilisant des réseaux de neurones. Le phasage de télescope comprend aussi bien l'étape de phasage grossier d'un déploiement d'une pupille segmentée que le phasage fin. Le NewSpace limitant non seulement le volume mais aussi la puissance de calcul, la complexité du modèle sera prise en compte dans ce manuscrit, car elle joue un rôle crucial dans la minimisation du contrôle de l'aberration optique en vue d'une mise en œuvre petits satellites.

0.4.2. Télescopes d'intérêt

On propose d'étudier l'identification des erreurs de phasage de 3 télescopes : le scénario complet de phasage du télescope spatial AZIMOV [Sch+20], [Sch+22], le phasage fin du JWST, et l'identification du mode pétale de l'ELT.

AZIMOV

L'utilisation d'une plateforme de petite taille pour un télescope spatial, comme les normes CubeSat de 10 cm, entraîne naturellement une limitation de l'ouverture optique maximale et une baisse de la résolution spatiale des images. Par exemple, un télescope de 10 cm embarqué à bord d'un CubeSat ne peut fournir que des images avec une résolution de 3 mètres depuis une orbite de 500 km (orbite basse) dans le spectre visible, en raison de la limite de diffraction. Une solution consiste à déployer un télescope avec une ouverture de 30 cm dans l'espace, offrant ainsi une résolution métrique (pour une orbite de 500 km et une longueur d'onde de 500 nm), en plus d'améliorer les capacités photométriques naturelles et le rapport signal/bruit grâce à une plus grande ouverture. D'autre part, cette approche repose sur l'utilisation de normes CubeSat peu coûteuses pour former une constellation, assurant ainsi un taux de revisite élevé. Un satellite de 6 unités (6U) pesant environ 10 kg est suffisant pour déployer un télescope de 30 cm dans l'espace. C'est le principe du télescope déployable *AutomatiZation of MOVable optics* (AZIMOV), voir figure 16.

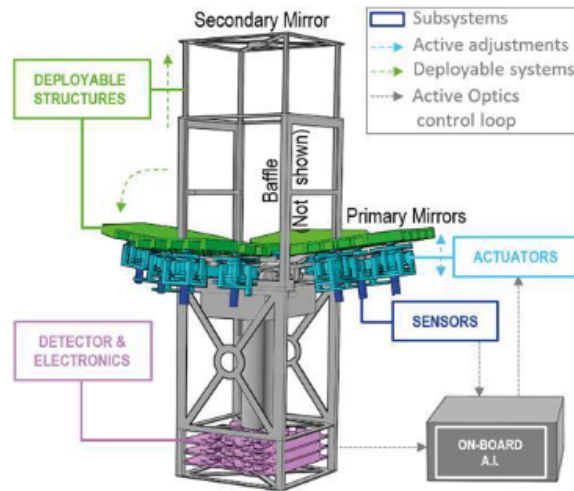


FIGURE 16. – Télescope spatial deployable AZIMOV. Extrait de Schwartz [Sch+20].

Le schéma du chemin optique du télescope est présenté en figure 17.

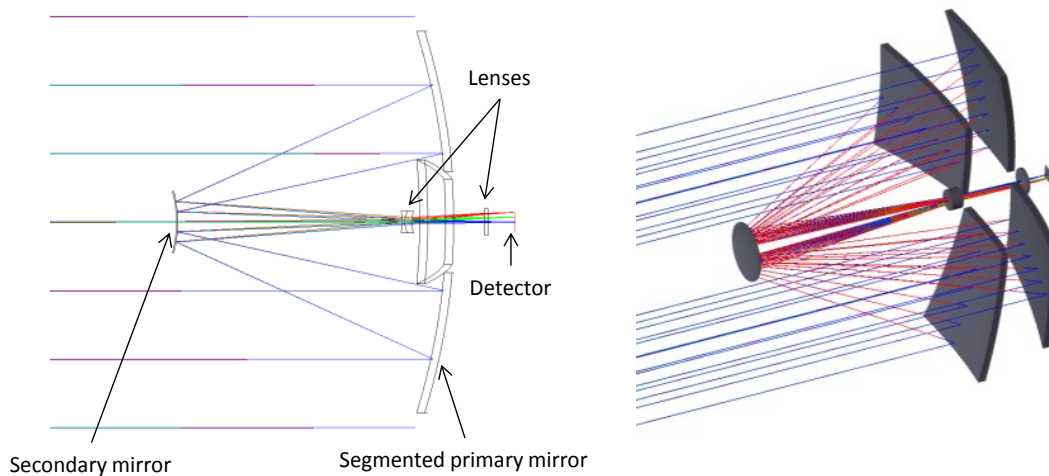


FIGURE 17. – Schéma optique du télescope spatial deployable AZIMOV. Extrait de Schwartz [Sch+20] et [Sch+22].

Le système d'imagerie du Cubesat déployable AZIMOV est basé sur un télescope de type Cassegrain avec un ensemble de lentilles pour ajuster la longueur focale et corriger les aberrations sur l'ensemble du champ de vision. La plus longue distance entre deux segments de notre miroir primaire segmenté est de 300 mm , apportant une résolution de $0,46\text{ arcsec}$ à 550 nm à la limite de diffraction. Le télescope est

composé d'un détecteur d'imagerie CMOS dont les pixels sont typiquement de $5.5\ \mu m$. Pour tirer pleinement parti de la résolution limite de diffraction, la longueur focale du télescope doit être réglée de manière à ce que la limite de diffraction corresponde à 2 pixels sur le détecteur, ce qui permet d'obtenir des images échantillonnées à Nyquist. Ces paramètres impliquent une longueur focale de $4900\ mm$ (rapport focal de $f/16.4$). À une altitude de $400\ km$, cela correspond à une distance d'échantillonnage au sol de $0,45\ m/pixel$ et une résolution d'image de $0.9\ m$ au sol. Un détecteur CMOS 2000×2000 pixels donnera alors un champ de vue de $0,9\ km^2$. En comparaison, les satellites SPOT6/7 pèsent $714\ kg$ pour obtenir une résolution au sol de $1,5$ mètre.

Cependant, des contraintes sur l'alignement de ce miroir primaire peuvent venir dégrader la qualité d'image. Dans les spécifications d'AZIMOV, le front d'onde résiduel WFE ne doit pas dépasser $70\ nm\ RMS$ sur l'ensemble de la pupille. Le front d'onde résiduel résultant des erreurs de déploiement, des erreurs d'analyse de front d'onde, des erreurs de correction, des évolutions temporelle du front d'onde entre l'analyse et la correction et des aberrations de haut ordre non corrigibles.

Category	Definition	Source	Comment
Intrinsic Optical Quality	Optics Quality	Manufacturing	Low-order residual due to the manufacturing process of the segments and system optics
		Polishing	Polishing errors, mainly high-order
	Initial alignment (in the laboratory)		Optical residuals of the optical elements after system alignment in the laboratory (excluding segment phasing)
	Ground-to-space degradation	Launch impact	Deformation of intrinsic optical quality due to launch
		Space environment	Deformation of intrinsic optical quality due to the difference between ground and space environments
	Observation environment	Thermal deformation	Effect of thermal deformation on the system's optics
		Line-of-sight deviations	Effect of the system's pointing error
Segment-related Optical Quality		Residual differential piston between segments	Phasing residual after active optics correction
		Residual differential tip-tilt between segments	Orientation residual after active optics correction
		Scalloping effect	Effect of the mismatch between the curvature of the segments and the curvature of the telescope, in terms of wavefront error
Deployment-related Optical Quality		Secondary mirror deployment accuracy	Effect of secondary mirror deployment on wavefront error

TABLEAU 2. – Résumé des perturbations de la qualité optique d'AZIMOV. Extrait de Schwartz [Sch+20].

Le tableau-ci dessus illustre les nombreuses origines pouvant dégrader la résolution du télescope AZIMOV. En l'absence de toute analyse système, nous proposons de réduire le budget d'erreur de l'analyse de front d'onde uniquement à environ **15 nm RMS** sur l'ensemble de la pupille. La valeur RMS du front d'onde résiduel est détaillée partie 1.3.2.

Propriété

La précision de l'analyse de front d'onde ne doit pas excéder 15 nm RMS .

Le phasage du télescope doit alors s'effectuer après déploiement et après pre-centrage des segments sur le détecteur à l'aide de la méthode ELASTIC ([VCM17]). Ensuite, on peut proposer 2 stratégies de phasage. La première en utilisant d'abord un point source connu, la seconde directement en observant la scène étendue (ici la surface de la Terre). On peut diviser le phasage en 2 étapes. La première serait la phasage initiale ou les aberrations sont plus importantes, le second serait le phasage fin, lorsque les aberrations sont petites et évoluent plus lentement dans le temps.

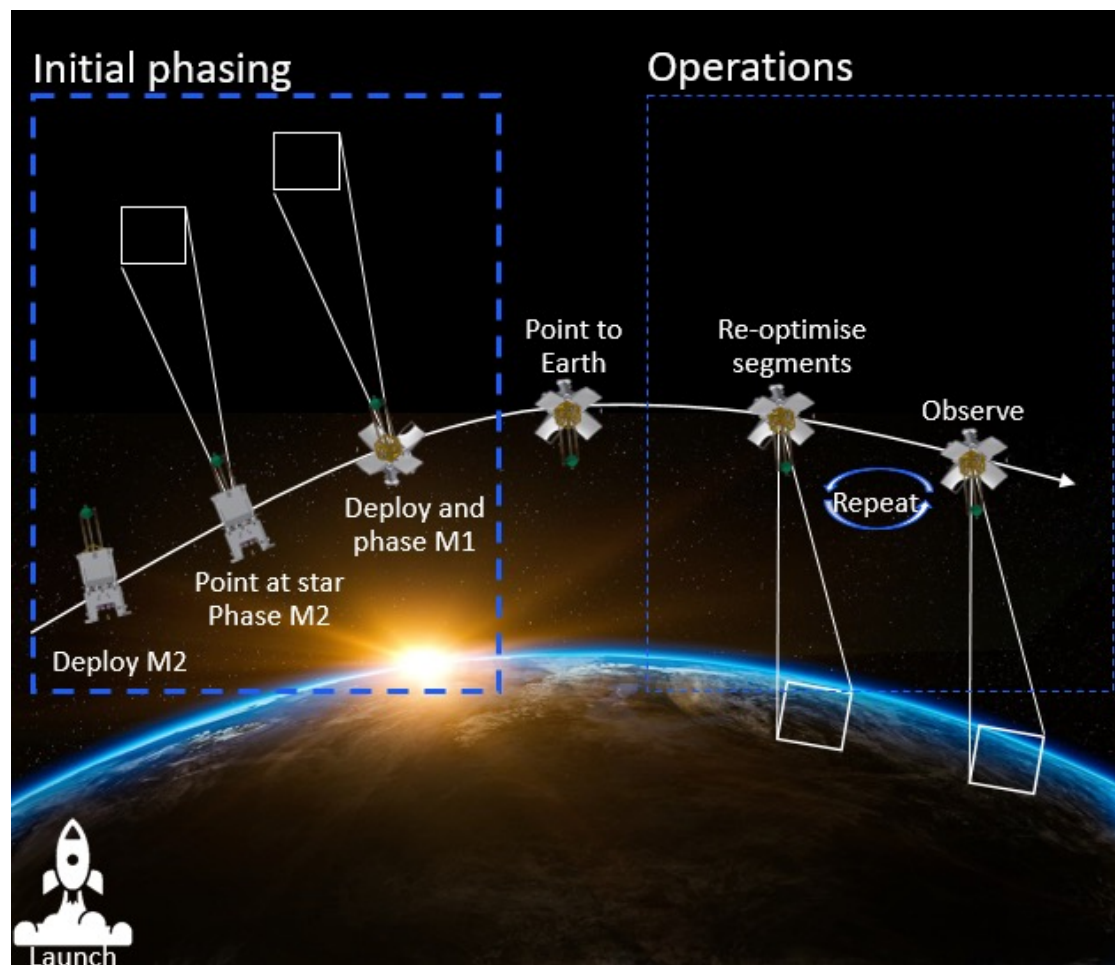


FIGURE 18. – Exemple de la stratégie de phasage d’AZIMOV. Extrait de Schwartz [Sch+22].

On peut observer un exemple d’une stratégie de phasage d’AZIMOV en figure 18. Dans ce manuscrit nous allons alors nous intéresser aux méthodes permettant de remonter aux erreurs de phasage du miroir primaire à partir de l’image plan-focal

du télescope en utilisant des méthodes de Deep Learning. Indépendamment de la stratégie de phasage choisie, nous allons étudier le phasage dans le cas d'observation d'un point source et d'un objet étendu.

Application à un plus grand nombre de segments

On propose d'étendre nos travaux au phasage d'un télescope à plus grand nombre de segments afin d'identifier les limites de notre méthode. Cette application s'apparente au phasage du télescope JWST avec ses dix-huit segments. La méthodologie développée pour AZIMOV, qui s'appuie sur le Deep Learning pour l'analyse du front d'onde en plan-focal, peut être adaptée pour gérer la tâche plus complexe de l'alignement d'un plus grand nombre de segments de miroirs. Bien que l'échelle du JWST s'accompagne de défis supplémentaires, notre méthodologie reste intacte.

Ouverture sur l'ELT et le mode pétale

Pour que l'ELT atteigne sa résolution angulaire théorique de 4 mas dans le domaine des longueurs d'onde visibles, ou 12 mas dans la bande K, il doit être équipé de systèmes d'optique adaptative performant et rapide, capables de corriger les aberrations de front d'onde jusqu'à quelques dizaines de nm. En raison de ses six grandes araignées, d'une largeur de 50 cm chacune, et de la surface segmentée de son miroir principal, la résolution angulaire de l'ELT est affectée par la présence de pistons qui provoquent des erreurs de phasage : en raison de la complexité de la forme de leur pupille (segmentation, grandes araignées...), certains pistons différentiels peuvent apparaître entre les parties isolées de la pupille complète pendant les observations. Bien que le système d'optique adaptative classique soit capable de corriger les effets de la turbulence, il sera aveugle à ces perturbations spécifiques induites par le télescope. Par conséquent, mesurer les erreurs de mise en phase de la pupille de l'ELT équivaut à mesurer ces différences de piston, également appelées pistons différentiels.

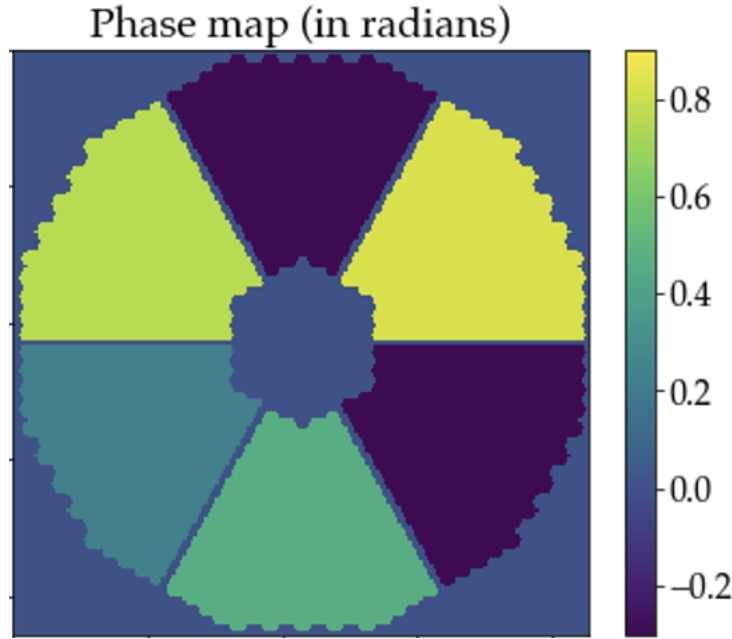


FIGURE 19. – Exemple du petaling de l’ELT. Les zones grisées représentent les araignées de support du M2. Extrait de la thèse de N.Levraud.

Ainsi, ce piston différentiel (illustré en figure 19), également appelé mode pétale, empêchera d’atteindre la limite de diffraction du télescope et représentera finalement la principale limitation de l’observation assistée par optique adaptative. De nombreux travaux proposent de corriger ce défaut des déphasage des sous-parties de la pupille ([Lev+24], [Ber+22]). Néanmoins, on propose d’étudier cette configuration et d’y apporter notre méthodologie en tant que solution potentielle et novatrice à la résolution des erreurs de pétaling du télescope. On propose d’appliquer notre méthode développée dans le cadre AZIMOV à un contexte presque similaire : mesurer les erreurs de piston de l’ELT (entre les 6 zones définies par les araignées) en utilisant les images issues d’un SH 2×2 fournies par l’un des multiples senseurs disponibles dans l’instrument. Ce cas applicatif est similaire puisqu’il s’agit de retrouver la quantité de pistons à partir d’images plan-focal (du SH) sans la présence d’atmosphère.

0.4.3. Plan du manuscrit

Le premier chapitre de cette thèse reprendra les principes de la formation d’image au sein d’un imageur, appliqué aux particularités de notre contexte. Dans ce chapitre, j’étudierai les aberrations optiques d’un télescope de manière générale, puis j’analyserai la particularité de la pupille d’AZIMOV ainsi que la formation d’image pour une source ponctuelle et un objet étendu. Le second chapitre portera l’accent sur la méthode employée, explicitant les rouages du Deep Learning. Une première introduction sera présentée, ainsi que les méthodes d’apprentissage, les architectures

et la méthodologie pour le phasage de télescope abordées dans ce manuscrit. Le troisième chapitre sera consacré aux résultats obtenus avec ces méthodes sur point source, leur robustesse et leur limite pour les différents télescopes. Je développerai une étude sur l'explication des paramètres du réseau de neurones dans le cas d'une pupille simplifiée. Je vérifierai ensuite la méthode sur une pupille complète, en évaluant sa robustesse, ses limites et sa comparaison avec l'état de l'art, ainsi qu'un scénario de phasage complet. Une ouverture sera finalement présentée sur une potentielle application au problème de petaling de l'ELT, et aux plus grands télescopes, à l'image du JWST. Une application expérimentale d'AZIMOV sera étudiée en chapitre 4. Dans ce chapitre, je validerai la méthode de Deep Learning pour des données expérimentales, puis j'étudierai le choix optimal du nombre de données expérimentales pour notre approche, à l'aide du *fine-tuning*. Le cinquième chapitre portera sur l'analyse de front d'onde sur objet étendu où une brève introduction à la formation d'image sera rappelée, suivie par une étude de notre méthode directe. Je présenterai ensuite une méthode pour intégrer la physique de la formation d'image dans l'entraînement d'un réseau de neurones, puis j'exposerai les résultats obtenus. Les limites seront présentées en fin de chapitre.

1. Formation d'image

Dans ce chapitre, le formalisme de la formation d'image en présence d'aberrations sera détaillé, en faisant apparaître les notions importantes pour nous : les aberrations au sens large, la forme de la pupille, la normalisation des aberrations, la longueur d'onde, le bruit et les perturbations. L'objectif de ce chapitre est d'explicitier notations et formalisme qui seront employés durant l'ensemble de la thèse et de les appliquer à notre cas précis d'étude.

Sommaire

1.1. Décomposition de la phase sur pupille circulaire	48
1.2. Pupille d'un télescope	50
1.2.1. Forme de la pupille	50
1.2.2. Asymétrie	51
1.3. Aberrations d'intérêt : mesure du piston-tip-tilt	54
1.3.1. Phasage des piston-tip-tilt	54
1.3.2. Normalisation des aberrations sur une pupille carrée	55
1.3.3. Définition de l'erreur de front d'onde	56
1.4. Formation d'image d'un point source : la PSF d'un télescope segmenté	57
1.4.1. Sources de perturbations additionnelles	57
1.4.2. Génération d'images	60
1.5. Cas d'un objet étendu	66
1.5.1. Relation objet-image	66
1.5.2. Sélection d'images caractéristiques	69
1.5.3. Base de données d'objet	71

1.1. Décomposition de la phase sur pupille circulaire

Décomposition modale de la phase : les modes de Zernike

Le front d'onde est une grandeur physique qui s'exprime comme une fonction réelle à deux dimensions de l'espace, continue ou non. A ce titre, on peut utiliser différentes bases modales pour la décrire. Afin d'analyser au mieux le front d'onde, on peut, comme pour tout espace vectoriel, le décomposer sur une base orthonormée afin de générer n'importe quelle phase à partir de cette base. Une des bases les plus courantes pour la décomposition de la phase est la base des polynômes de Zernike. Cette base, orthonormée sur un support circulaire, est composée de polynômes qui s'expriment comme le produit d'une fonction trigonométrique et d'une fonction polynomiale :

$$Z_i(r) = R_n^m(r)\Theta_n^m(\theta) \quad (1.1)$$

avec :

$$R_n^m(r) = \sum_{s=0}^{\frac{(n-m)}{2}} \frac{(-1)^s (n-s)!}{s! [\frac{(n+m)}{2} - s]! [\frac{(n-m)}{2} - s]!} r^{n-2s} \quad (1.2)$$

et

$$\Theta_n^m(\theta) = \begin{cases} \sqrt{n+1} & \text{si } m = 0 \\ \sqrt{2(n+1)} \cos(m\theta) & \text{si } m \neq 0 \text{ et } i \text{ pair} \\ \sqrt{2(n+1)} \sin(m\theta) & \text{si } m \neq 0 \text{ et } i \text{ impair} \end{cases} \quad (1.3)$$

On a alors une représentation des modes de Zernike sur une pupille circulaire en fonction du degrés radial et du degrés azimuthal.

1. Formation d'image – 1.1. Décomposition de la phase sur pupille circulaire

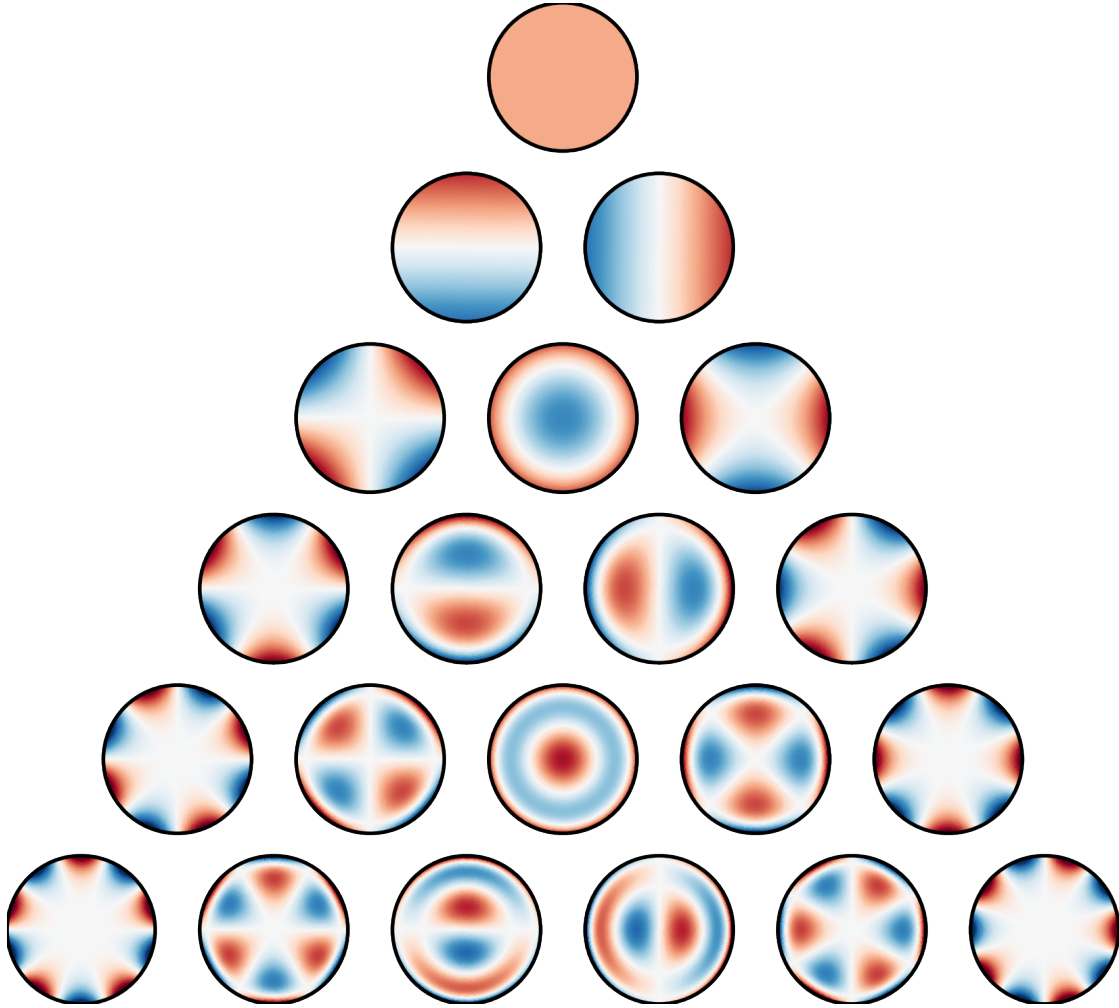


FIGURE 1.1. – Représentation des 21 premiers modes de Zernike. Extrait de [Polynôme de Zernike](#).

Cette base permet donc de décomposer la phase :

$$\Phi = \sum_{i=1}^{\infty} a_i Z_i \quad (1.4)$$

L'orthonormalisation de la base permet de mesurer simplement la variance spatiale de la phase en fonctions des coefficients d'amplitude a_i :

$$\sigma_{\Phi}^2 = \sum_{i=1}^{\infty} a_i^2 \quad (1.5)$$

Les trois premiers modes de Zernike représentés figure 1.1 représentent respectivement le *Piston, Tip, Tilt* (PTT). Ce sont les modes que l'on cherche à corriger sur chacun des segments lors du phasage d'un télescope. Les modes de plus hauts ordres,

quand a eux constituent des perturbations qui ne pourront pas être corrigée par AZIMOV.

1.2. Pupille d'un télescope

On s'intéresse ensuite aux formes de pupille des télescopes et à leur incidence sur la formation d'image.

1.2.1. Forme de la pupille

Un des éléments clef de la formation d'image est la forme de la pupille (ou miroir primaire) où vient se réfléchir l'onde incidente. Elle est généralement circulaire ce qui limite l'extension spatiale de l'onde électromagnétique se propageant dans l'instrument. Une obstruction centrale peut être observée au centre de la pupille due au positionnement du miroir secondaire du télescope. Le télescope principalement étudié dans ce manuscrit, AZIMOV, étant un satellite spatial compact, la forme de la pupille s'en trouve affectée. La pupille est composée de quatre segments distincts carrés. Elle n'est donc pas circulaire et la pupille est très diluée ce qui impacte la formation d'image et le flux collecté.

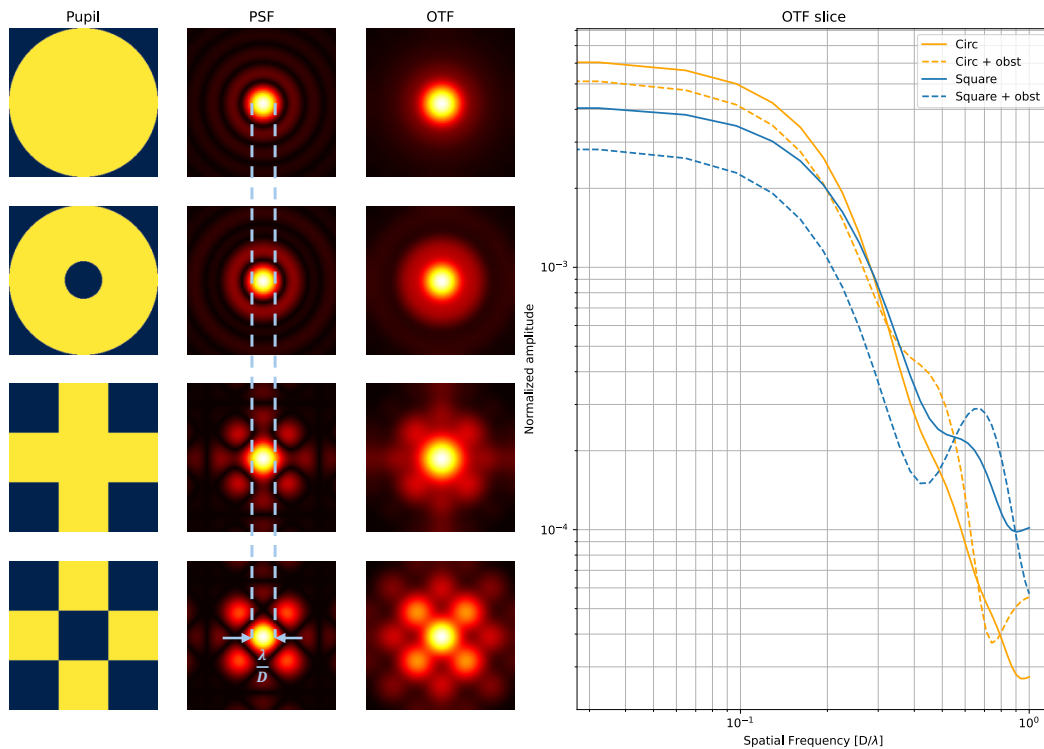


FIGURE 1.2. – Illustration de pupille circulaire et de forme d'AZIMOV, avec et sans obstruction centrale avec leurs PSF et FTO correspondantes.

La figure 1.2 permet d'apprécier l'impact de la forme de la pupille sur la fonction de transfert optique du système imageur. On observe tout d'abord que la largeur à mi-hauteur ($\frac{\lambda}{D}$) caractéristique de la PSF reste la même indépendamment de la pupille étant donné que le diamètre de ces dernières est identique. Ensuite les FTO parfaites sont montrées : ici dans ce cas idéal les FTO ne sont pas détériorées et montrent les limites physique des fréquences spatiales pouvant être imagées due à la géométrie de la pupille. Comparées aux pupilles non diluées, les FTO des pupilles diluées comportent des amplitudes plus faible sur la majorité des fréquences, traduisant ainsi une perte de transmission lors de l'imagerie des fréquences correspondantes.

1.2.2. Asymétrie

La pupille conçue par UK-ATC pour AZIMOV propose un design asymétrique : le **Miroir primaire (M1)** se compose de quatre segments de miroir individuels, qui doivent être alignés ensemble pour agir comme un seul miroir primaire. La figure 1.3 est une carte de front d'onde à l'alignement nominal, montrant la forme de la pupille d'entrée actuelle, avec indication de ses pétales individuels. Le pétale 1 est en position fixe, tandis que les pétales 2, 3 et 4 sont déployables et peuvent être ajustés en vol. Le diamètre de M1 aligné est de 300 mm. Les coins des pétales 3 et 4 sont coupés pour permettre aux pétales de s'adapter au volume du CubeSat replié. Cependant, ces coins se trouvent en dehors du diamètre de pupille requis de 300 mm et sont coupés par la taille du miroir secondaire.

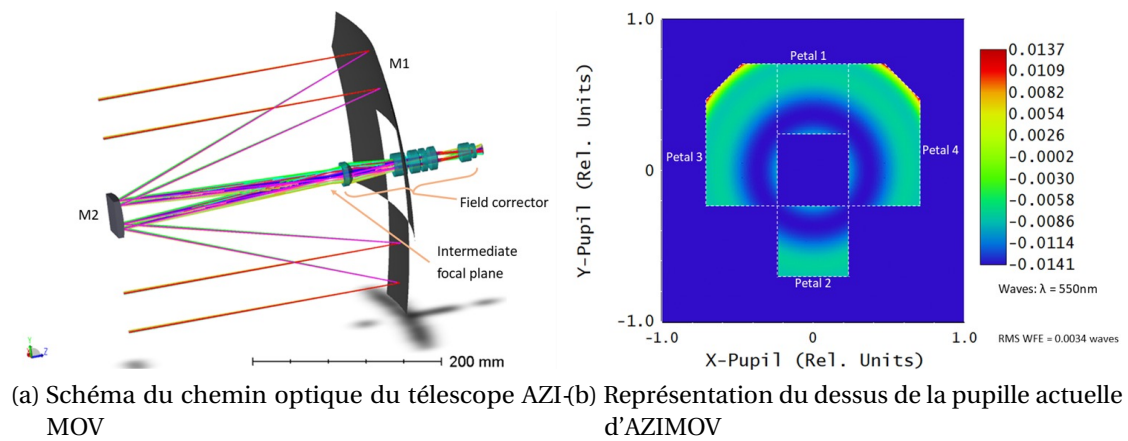


FIGURE 1.3. – Schéma du système optique d'AZIMOV.

Cette particularité géométrique permet une meilleure identification de la phase à partir de l'image plan-focal. Cependant le concept d'AZIMOV n'étant pas encore terminé à ce jour, la pupille est sujet à des modifications. On propose donc de travailler avec un modèle de pupille fixé symétrique et asymétrique utilisant des segments carrés dont la géométrie est présentée en figure 1.4. Cette forme de pupille, qui sera utilisée dans l'ensemble du manuscrit, correspond à la forme de la pupille telle que définie au

commencement des travaux de recherche. Les avancées des travaux sur le télescope AZIMOV ont modifiées à ce jour la forme réelle de la pupille telle que présentée en figure 1.3, n'influençant pas les solutions et les concepts étudiés ici.

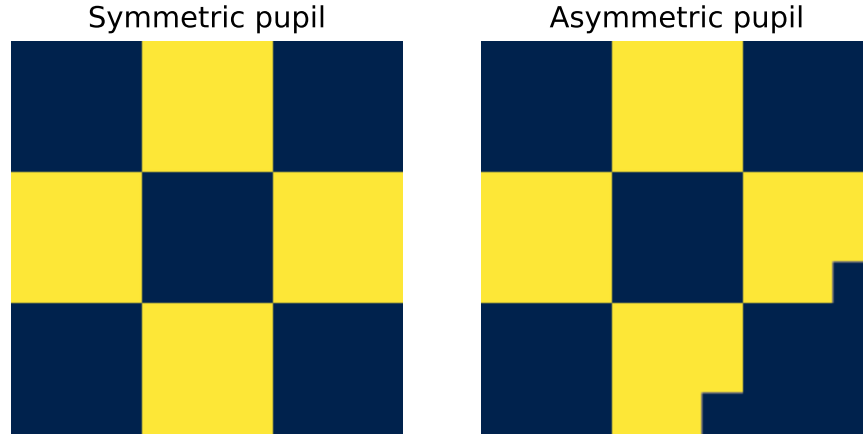


FIGURE 1.4. – Illustration du modèle choisi de la pupille AZIMOV symétrique et asymétrique.

On remarque en figure 1.4 que sur la pupille asymétrique, la partie en bas à droite de deux segments est coupée afin de briser la centro-symétrie de la pupille. Une partie de notre étude portera sur l'incidence de cette asymétrie sur l'AFOPF (développé partie 3.2.1).

Formalisme de l'asymétrie

Dans ses travaux, Martinache ([Mar13]) montre que l'asymétrie de pupille permet d'identifier la partie paire de la phase. En effet dans le cas de petites phases, on peut alors effectuer le développement limité à l'ordre 1 du champ électromagnétique dans le plan pupille tel que

$$Pe^{i\Phi} = P(1 + i\Phi) \quad (1.6)$$

Puis en réinjectant dans la fonction d'autocorrelation équation 0.7 puis en linéarisant on obtient alors l'expression de la FTO :

$$FTO = \int_{-\infty}^{+\infty} P(x)P(x+u) dx + i \int_{-\infty}^{+\infty} P(x)P(x+u)\Phi(x+u) dx + i \int_{-\infty}^{+\infty} P(x)P(x+u)\Phi(x) dx \quad (1.7)$$

La partie réelle dépendant uniquement de la pupille ne comporte alors pas d'information sur la phase à identifier. On s'intéresse donc uniquement à la partie imaginaire

de la FTO :

$$Im(FTO) = Im(\mathcal{F}(|\mathcal{F}(Pe^{i\Phi})|^2)) = \int_{-\infty}^{+\infty} P(x)P(x+u)\Phi(x+u) dx + \int_{-\infty}^{+\infty} P(x)P(x+u)\Phi(x) dx \quad (1.8)$$

On pose $A(x) = P(x)\Phi(x)$, un changement de variable $x' = x + u$ dans la seconde intégrale et en séparant le domaine d'intégration donne alors

$$Im(FTO) = \int_0^{-\infty} P(x)A(x+u) dx + \int_0^{+\infty} P(x)A(x+u) dx + \int_0^{-\infty} P(x)A(x-u) dx + \int_0^{+\infty} P(x)A(x-u) dx \quad (1.9)$$

En posant enfin $x' = -x$ on obtient alors

$$Im(FTO) = \int_0^{+\infty} P(x)A(x+u) dx - \int_0^{+\infty} P(-x)A(-x+u) dx + \int_0^{+\infty} P(x)A(x-u) dx - \int_0^{+\infty} P(-x)A(-x-u) dx \quad (1.10)$$

L'hypothèse de symétrie de la pupille implique que $P(-x) = P(x)$. On propose alors d'étudier la FTO de la partie paire de la phase, impliquant alors $\Phi(x) = \Phi(-x)$ et donc donc $A(x) = A(-x)$. La partie imaginaire de la FTO devient alors

$$Im(FTO) = \int_0^{+\infty} P(x)A(x+u) dx - \int_0^{+\infty} P(x)A(x+u) dx + \int_0^{+\infty} P(x)A(x-u) dx - \int_0^{+\infty} P(x)A(x-u) dx \quad (1.11)$$

les termes de la partie imaginaire s'annulent, elle ne peut donc pas être identifiée directement.

A contrario, briser la centro-symétrie de la pupille permet d'accéder à cette partie imaginaire de la FTO et donc la partie paire de la PSF. A partir de l'équation 1.10 en intégrant uniquement sur une surface d'asymétrie de la pupille S (n'impliquant plus les conditions $P(-x) = P(x)$ et $A(x) = A(-x)$), on obtient alors plus d'information sur la partie imaginaire de la FTO.

$$Im_s(FTO) = \int_S P(x)A(x+u) dx + \int_S P(x)A(x-u) dx \quad (1.12)$$

Les deux termes ne s'annulent alors plus, et donc la partie paire de la partie imaginaire de la FTO porteuse d'information sur la phase peut être identifiée. Projeté sur les modes de Zernike, cela signifie que le signe des modes paires de la base de Zernike peut être identifié directement à partir de la seule PSF en plan-focal ; étape physiquement et mathématiquement impossible lorsque la pupille est parfaitement centro-symétrique.

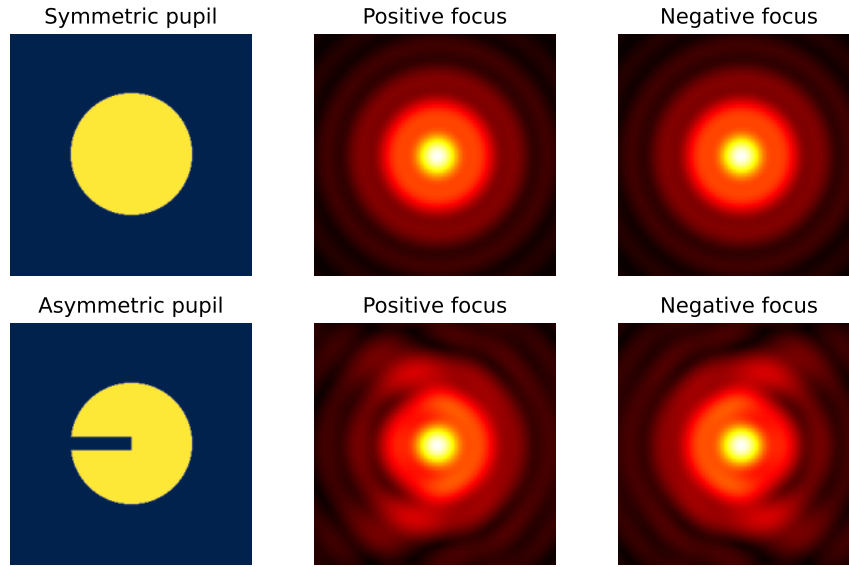


FIGURE 1.5. – Comparaison de l'impact d'un focus sur une pupille symétrique et asymétrique sur les PSFs (en échelle racine).

Pour exemple, la figure 1.5 montre l'impact d'un focus sur la PSF d'une pupille symétrique par rapport à une pupille asymétrique : dans le cas asymétrique, le signe du focus peut être déterminé à partir de la PSF seule.

Choix de l'asymétrie

Cependant le choix de la surface d'asymétrie n'est pas aisé car il constitue une perte du flux total réfléchi par la surface collectrice. Il faut alors effectuer un compromis entre la surface d'asymétrie et qualité de l'image observée. Dans le cadre du télescope AZIMOV, le choix de la surface d'asymétrie sera développé partie 3.2.1.

1.3. Aberrations d'intérêt : mesure du piston-tip-tilt

1.3.1. Phasage des piston-tip-tilt

L'objectif des travaux de la thèse est d'identifier les erreurs de phasage à partir d'une seule image plan-focal. Les 4 segments constituant la pupille pouvant chacun être désaligné par rapport à l'autre, on cherche alors à identifier le Piston et le Tip-Tilt pour chacun des 4 segments. 12 coefficients de la phase sont à estimer.

1.3.2. Normalisation des aberrations sur une pupille carrée

En introduction, nous avons évoqué la base des modes de Zernike, une base orthonormée sur pupille circulaire reconnue pour son utilité dans l'analyse de front d'onde. Cependant son orthonormalisation est limitée à une pupille circulaire et non aux pupilles carrées comme pour notre application. La normalisation des modes doit donc être recalculée. Dans les travaux décrit par ce manuscrit on propose de renormaliser la base des modes de Zernike, en conséquence de la forme de la pupille utilisée. Cette normalisation définit alors une base de mode orthonormée sur une pupille carrée B_i^k (i représentant le mode, k le segment associé) telle que

$$\langle B_i, B_j \rangle = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases} \quad (1.13)$$

et

$$\frac{1}{4} \sqrt{\frac{1}{4} \int_{-1}^1 \int_{-1}^1 M_i M_j dx dy} = 1 \quad \text{if } i = j \quad (1.14)$$

D'après la normalisation de Noll ([Nol76]), en coordonnée sphérique sur une pupille circulaire on a

$$\sqrt{\frac{1}{A} \int_0^1 \int_0^{2\pi} Z_k^2 d\rho d\theta} = 1 \quad (1.15)$$

avec A l'air de la pupille et Z_k le mode k de Zernike.

Cependant les normalisations doivent être effectuées sur pupille carrée. Commençons par le piston. Le piston étant une élévation constante unitaire sur la pupille qui en varie pas autour de sa moyenne, on peut aisément calculer sa valeurs RMS :

$$|p|_{RMS}^{\square} = \sqrt{\frac{1}{4} \int_{-1}^1 \int_{-1}^1 1^2 dx dy} = 1 \quad (1.16)$$

Le piston est donc normalisé aussi sur pupille carrée. On détermine maintenant la valeur RMS de la rampe de Tip et Tilt sur une pupille carrée telle que :

$$|tip|_{RMS}^{\square} = \sqrt{\frac{1}{4} \int_{-1}^1 \int_{-1}^1 x^2 dx dy} = \frac{1}{\sqrt{3}} \quad (1.17)$$

et

$$|tilt|_{RMS}^{\square} = \sqrt{\frac{1}{4} \int_{-1}^1 \int_{-1}^1 y^2 dx dy} = \frac{1}{\sqrt{3}} \quad (1.18)$$

Afin de normaliser les modes de tip et de tilt, ces derniers doivent aussi être ajustés afin que leurs valeur RMS soit unitaire. Donc imposer, pour une pupille carrée une normalisation des rampes par $b_{tip}^{\square} = b_{tilt}^{\square} = \sqrt{3}$. Cependant, dans le cas d'AZIMOV, un segment représente $\frac{1}{4}$ de la surface totale, le coefficient ponderateur b est alors ajusté en conséquence :

$$b_{piston}^{\star} = 4, b_{tip}^{\star} = 4\sqrt{3} \text{ et } b_{tilt}^{\star} = 4\sqrt{3} \quad (1.19)$$

finalemeht, pour la pupille AZIMOV on obtient les trois premiers termes de la base B_i^k :

Propriété
$\begin{aligned} B_1^k(x, y) &= 4 \\ B_2^k(x, y) &= 4\sqrt{3}x \\ B_3^k(x, y) &= 4\sqrt{3}y \end{aligned} \quad (1.20)$

La base B_i^k est donc orthonormalisée sur une pupille carrée. Néanmoins elle sera considérée sans modification dans l'ensemble du manuscrit alors qu'une pupille asymétrie sera utilisée pour une partie des résultats présenté.

On peut donc maintenant établir la phase à partir de notre base B_i^k telle que énoncée équation 1.4 :

$$\Phi = \sum_{i=1}^3 \sum_{k=1}^4 c_i^k B_i^k \quad (1.21)$$

où c_i^k représente l'amplitude du mode i du segment k.

1.3.3. Définition de l'erreur de front d'onde

L'erreur de front d'onde (*WaveFront Error (WFE)*) définie la quantité d'erreur de front d'onde sur l'ensemble de la pupille. On choisit de définir le WFE comme étant la somme sur les segments de la somme quadratique des 3 modes :

Propriété
$WFE(\{c_i^k\}) = \sum_{k=1}^4 \sqrt{\sum_{i=1}^3 c_i^{k2}} \quad (1.22)$

C'est cette métrique qui va dicter la définition de l'erreur de front d'onde tout au long du manuscrit. Notre métrique du WFE se différencie de la valeur RMS classique, correspondant à $WFE_{RMS} = \sqrt{\sum_{k=1}^4 \sum_{i=1}^3 c_i^{k2}}$. Notre définition du WFE 1.22 considère une contribution linéaire de chacun des segments, une caractéristique importante lorsque nous étudierons le front d'onde sur un nombre variable de segments. Ce faisant, nous choisissons une spécification plus contraignante, puisque notre définition du WFE est statistiquement 2 fois plus élevée que celle du WFE_{RMS} . La spécification de la précision de mesure d'un $WFE_{RMS} < 15 \text{ nm RMS}$ correspondrait à une valeur RMS d'une erreur de front d'onde $WFE < 30 \text{ nm}$. Néanmoins, on choisit d'appliquer la spécification d'un $WFE < 15 \text{ nm}$ défini à l'équation 1.22.

1.4. Formation d'image d'un point source : la PSF d'un télescope segmenté

Le formalisme du front d'onde étant établi, intéressons-nous maintenant au formalisme de la formation d'image à proprement parler en présence d'aberrations. La réponse impulsionnelle du système en plan-focal est l'image d'un point source à l'infini. Si l'on applique l'équation 0.5 à notre étude on écrit alors :

$$PSF(\mathbf{r}) = |\mathcal{F}(Pe^{i(\Phi)})(\mathbf{r})|^2 = |\mathcal{F}(Pe^{i(\Phi_{PTT} + \Phi_{HO})})(\mathbf{r})|^2 \quad (1.23)$$

Où P correspond à la pupille vue partie 1.2, Φ représente la phase du champ électromagnétique au sens strict, Φ_{PTT} représente la phase correspondantes aux aberrations de Piston et tip-tilt et Φ_{HO} la phase correspondante aux aberrations de plus hauts ordres qui seront développées par la suite. On rappelle aussi que la phase est chromatique et dépend de la longueur d'onde λ : $\Phi = \frac{2\pi}{\lambda} OPD$, ou l'OPD correspond à l'*Optical Path Difference (OPD)*, la carte des aberrations en plan pupille. De plus, le projet d'AZIMOV propose des observations dans le visible en panchromatique entre 500 et 900 nm. Dans notre étude, une seule longueur d'onde à $\lambda = 800 \text{ nm}$ est choisie.

1.4.1. Sources de perturbations additionnelles

Aberrations de plus hauts ordres

Origines

Les aberrations de plus haut ordres (*Higher Order aberrations (HO)*) vont aussi affecter le miroir primaire et l'ensemble des autres optiques. Cependant le design d'AZIMOV ne permettant pas leur correction, on s'affranchit de leur estimation tout en s'assurant que leur présence ne perturbe pas la mesure des PTT. Néanmoins ils interviennent dans la formation d'image et dégradent cette dernière. 3 causes principales sont à l'origine de ces perturbations :

- **Les gradients thermiques** : La source majeure de perturbation additionnelle est due aux variations de température. La principale source de déformation est due aux gradients thermique entre la surface optique du pétale et le support à l'arrière de ce dernier : dans un scénario extrême, une surface est au soleil, l'autre à l'ombre. Les gradients déforment le miroir, plus particulièrement il intervient principalement sous forme de focus additionnel à la phase. Plus le gradient de température est élevé plus le focus est conséquent.

1. Formation d'image – 1.4. Formation d'image d'un point source : la PSF d'un télescope segmenté

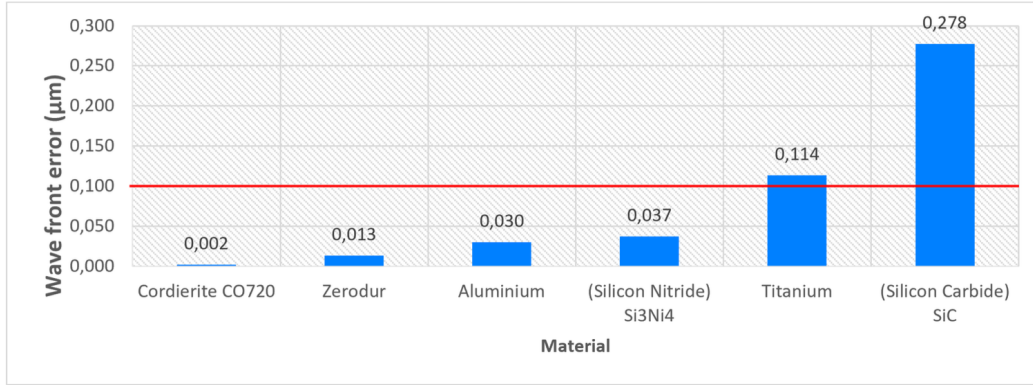


FIGURE 1.6. – Illustration de l'impact du choix des matériaux sur la déformation du miroir par des gradients thermiques. Extrait du rapport de stage de L. Bouron (2021).

La déformation minimum induite par les gradients thermiques est de l'ordre de 30 nm pour l'aluminium (voir figure 1.6), matériau choisi actuellement pour le design d'AZIMOV. Il correspond à la quantité de haut ordre maximum afin d'atteindre un système à la limite de la diffraction.

- **Les défauts de polissage de miroir** : Ils sont statiques et inhérents à la qualité de fabrication du miroir, étant de très haute fréquence spatiale, de faible intensité et statique, ils ne représentent pas une limite pour atteindre la qualité d'image promise par le télescope.
- **Les défauts d'alignement des optiques** : Ces erreurs à évolution très lente voire fixes sont aussi inhérentes à la conception du télescope, cependant les optiques peuvent bouger durant le lancement du satellites ou durant des vibrations pendant l'orbite, détériorant ainsi le front d'onde.

Distribution

Pour l'étude de l'impact des hauts ordres, on choisi alors de repartir l'amplitude ces derniers selon une distribution suivant une loi décroissante i^{-2} en fonction du mode i . On fixe la valeur RMS sur l'ensemble de la pupille des hauts ordres telle que :

$$WFE_{HO} = \sqrt{\sum_{i=4}^{100} \frac{1}{i^2}} = 30 \text{ nm RMS} \quad (1.24)$$

Les phases sont alors générées sur une pupille circulaire et mises à l'échelle afin d'obtenir un front d'onde d'exactement 30 nm sur l'ensemble de la pupille. Ce front d'onde aberrant n'est pas corrigé et n'intervient pas dans le calcul du front d'onde résiduel, il apporte donc simplement un bruit additionnel à la mesure. Comme l'indique l'équation 1.13, la génération des modes de plus haut ordres doit être orthogonale afin éviter la projection d'un mode de haut ordre sur les autres. On peut trouver un exemple d'une carte d'OPD composée des HO évoqués en figure 1.7

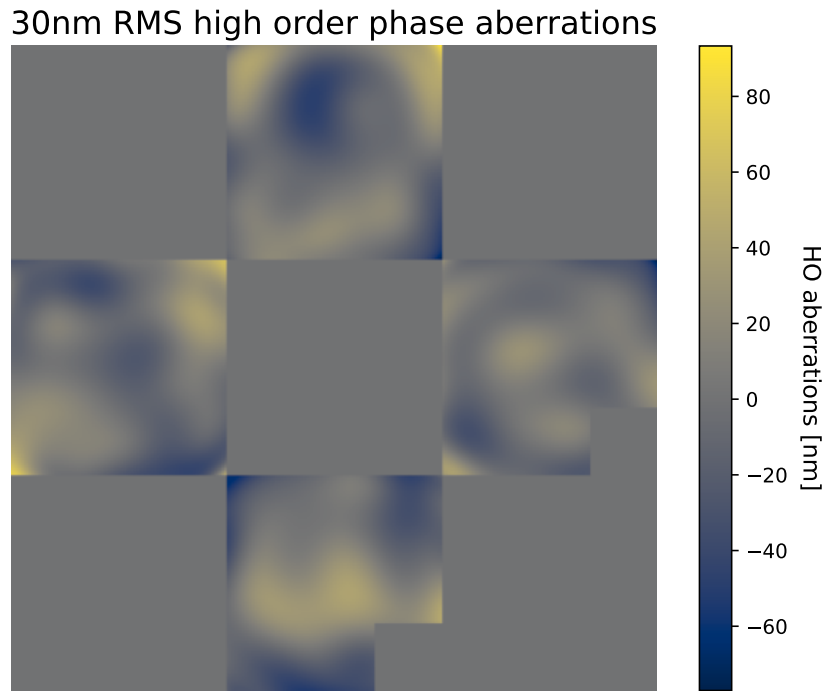


FIGURE 1.7. – Exemple d'aberration de haut ordres, additionnelle au Piston-Tip-Tilt sur la pupille AZIMOV asymétrique.

Bruit

Dans un régime de faible flux, l'image souffre principalement du bruit de lecture, estimé à 2.1 électrons (pour un détecteur CMOS Sony Pregius-S IMX530 avec 5320×4600 pixels, détecteur candidat pour AZIMOV), tandis que le bruit de photon domine dans les cas de flux élevé. Le bruit de photon suit une loi de Poisson dépendante du flux incident $\mathcal{P}(f)$ alors que le bruit de détecteur suit une distribution gaussienne de moyenne nulle et d'écart-type $\sigma_{ron} = 2.1e^-$. Ce sont les deux bruits pris en compte dans l'étude, les autres sources de bruit (bruit de fond ou encore bruit thermique) ayant été choisies d'être négligées. On peut alors définir le rapport signal à bruit (*Signal to Noise Ratio* (SNR)) permettant d'apprécier la quantité de bruit par rapport à la quantité d'information utile :

Propriété

$$SNR = \frac{f}{\sqrt{f + \sigma_{ron}^2}} \quad (1.25)$$

ici f représente le flux de l'objet étudié en photon.

Source d'erreurs dynamiques : Vibration et pointage

Les erreurs dues aux vibrations durant l'orbite ou les erreurs de pointage peuvent être divisées en 2 catégories :

- Les perturbations plus rapides que le temps d'exposition : dans ce cas le satellite n'est pas stable durant la collecte de photons. L'objet est alors flouté par une gaussienne représentative de la vibration. Alors, même dans le cas de l'observation d'un point source, l'image est alors plus étendue et non nécessairement centrosymétrique. Cette élongation de l'image causée par la vibration peut quelque fois être confondue avec un objet étendu.
- Les perturbations plus lentes que le temps d'exposition : dans ce cas l'erreur de pointage se comporte comme un tip-tilt global sur la pupille. L'objet est décalé par rapport au centre de l'image, et le tip-tilt peut être corrigé. Cependant il faut être attentif aux erreurs de pointages car la course des actionneurs est limitée. Corriger le tip-tilt en direct sur le miroir primaire est une solution envisageable pour les erreurs de pointage mais elle ne peut être appliquée en boucle fermée. Cependant, un miroir de pointage en avant peut être considéré pour compenser ces erreurs et décharger les courses des actionneurs.

1.4.2. Génération d'images

Les méthodes de deep learning et notamment les méthodes supervisées nécessitent une grande quantité de donnée afin d'en apprendre les caractéristiques intrinsèques. Malheureusement, une telle quantité de données réelles n'est pas disponible actuellement, encore moins avec les labels correspondants. Pour palier ce défaut, on propose alors de simuler numériquement les données utilisées. Cela permet non seulement d'accéder à un nombre illimité de données, mais aussi de maîtriser leurs caractéristiques et labels (ici, les coefficients de la phase) correspondants. Dans un premier temps nous nous intéressons aux données issues de l'observation d'un point source, soit la PSF.

Caractéristiques

Les données simulées présentent un champ de vue de 64×64 pixels avec une résolution d'environ 0.5 m/pix vue partie 0.4.2, cela représente un champ de vue de $32 \times 32 \text{ m}^2$ soit un champ de 13.2 arcsec à LEO. Les PSFs sont échantillonnées à Nyquist afin d'obtenir exactement l'information nécessaire à l'identification de la phase, et l'intensité globale de l'image est normalisée à 1. Les PSFs représentent donc la distribution d'un seul photon dans le plan-focal et sont encodées sur 32 bits.

Distribution des paramètres

La distribution des paramètres de la phase se doit d'être représentatives du scénario de phasage du télescope. Les tests de répétabilité de déploiement mécanique des segments indiquent une même distribution normale des erreurs de piston et tip-tilt.

- **Le régime de déploiement ou phasage grossier.** Dans ce régime, les sous-PSFs ne se recouvrent pas (séparation de l'ordre de plusieurs λ/d avec d la taille d'un segment) dans le champ mais apparaissent toutes sur le détecteur. On propose alors d'effectuer des tirages selon une gaussienne de moyenne nulle et d'écart-type $\sigma_{dep} = 70 \text{ nm RMS}$ pour le tip-tilt. Le piston lui est tiré selon la même loi de probabilité que pour le phasage fin : $\sigma_{dep}^{piston} = \sigma_{fin}^{piston}$. Ce choix est développé au prochain paragraphe.
- **Régime de phasage fin.** C'est le régime où les sous-PSFs issues des segments (ou sous-pupille) se superposent et interfèrent, les erreurs de front d'onde sont faibles. Le tirage des coefficients de la phase suit une loi normale de moyenne nulle et d'écart-type $\sigma_{fin} = 10 \text{ nm RMS}$ par coefficient.

1. Formation d'image – 1.4. Formation d'image d'un point source : la PSF d'un télescope segmenté

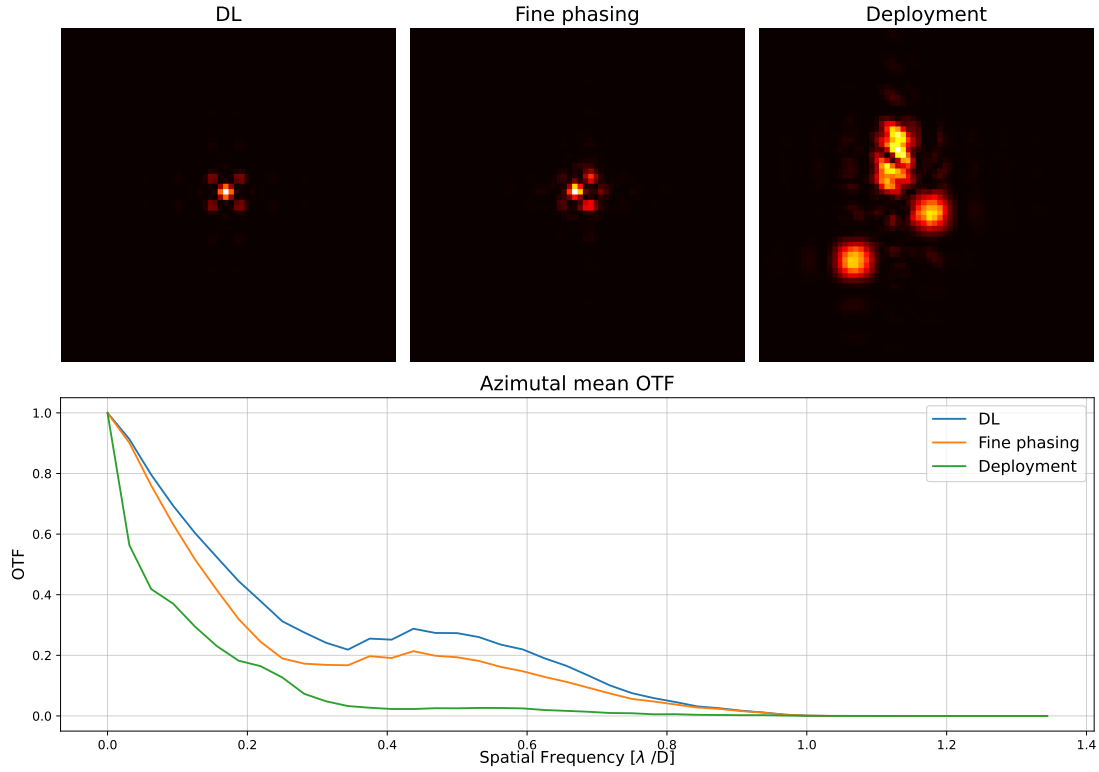


FIGURE 1.8. – Exemple de PSFs typique d’AZIMOV, pour les différents scenario évoqués et leur FTO correspondante. La première image (haut gauche) représente la PSF parfaite du télescope, la PSF à la limite de la diffraction. La seconde image représente la PSF dans le cas du phasage fin. La figure en haut à droite représente la PSF soumise aux erreurs de phasage de déploiement. La figure en dessous représente les moyennes azimutales des FTOs correspondantes.

On peut observer les PSFs correspondantes aux différents tirages figure 1.8. On observe alors la superposition des 4 sous-PSFs de largeur $\frac{\lambda}{D}$ pour la PSF *fine phasing* correspondant au régime de phasage fin. La troisième image représente les sous-PSFs réparties ou superposées dans le plan-focal. Les sous-PSFs sont alors de largeur $3\frac{\lambda}{D}$ car elles correspondent à la PSF produite par la sous pupille qui représente $\frac{1}{3}$ de la largeur totale de la pupille. L’échantillonnage est donc 3 fois plus grand pour les sous-pupilles. Les FTOs sont représentées sur la graphique du dessous, la courbe orange représente une optique de qualité légèrement dégradée jouant le rôle d’un filtre pour les fréquences spatiales de l’objet, et atténuant certaines transmission sans nécessairement les annuler. Pour la FTO du régime de déploiement (courbe verte), la transmission s’approche dangereusement de zéro mais ne va sans doute pas atteindre exactement zéro, cependant la transmission de certaines fréquences spatiales est tellement faible, que le SNR de ces fréquences spatiales dans l’image est très petit devant 1. Donc l’image ne possède plus ces fréquences spatiales. On peut remarquer

1. Formation d'image – 1.4. Formation d'image d'un point source : la PSF d'un télescope segmenté

que la fréquence de coupure du télescope est atteinte à $\frac{1}{3}$ de $\frac{\lambda}{D}$ correspondant donc à la fréquence de coupure d'un seul des segments de la pupille. A partir des tirages des coefficients PTT, on peut alors établir la statistique du front d'onde incident du jeu de données ainsi généré en calculant l'erreur de front d'onde pour chaque réalisation à l'aide de l'équation 1.22.

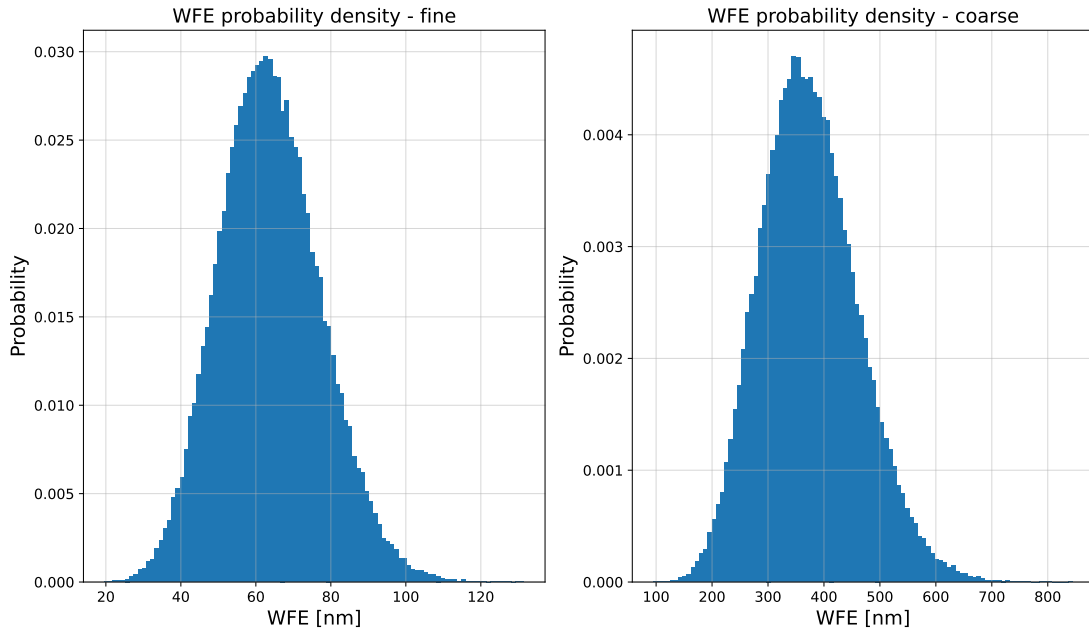


FIGURE 1.9. – Exemple de distribution du WFE sur un dataset de 10^5 PSFs. A gauche la distribution du régime de phasage fin, avec un WFE moyen de 64 nm. A droite la distribution pour le régime de déploiement avec un WFE moyen de 320 nm.

La figure 1.9 présente la distribution des WFEs sur l'ensemble de la pupille AZIMOV pour le régime de phasage fin et le régime de déploiement. Ces statistiques sur le front d'onde total de la pupille permettent donc d'identifier les deux régimes de phasage qui seront étudiés dans ce manuscrit. Cependant, la particularité d'utiliser une unique image monochromatique en plan-focal vient limiter la plage d'identification d'un des modes de la phase : le piston.

Sensibilité aux modes

Piston global

Dans un premier temps le piston global : l'image est insensible à l'erreur de piston global de la pupille. On choisit alors de le filtrer en soustrayant le piston moyen aux

coefficients de la phase :

$$p' = (I - \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}) \quad (1.26)$$

Forcer le piston moyen à une moyenne nulle permet d'éviter les divergences des valeurs du piston lors de l'estimation.

Phasage grossier et sensibilité au piston

A l'issue du déploiement du télescope, l'amplitude des tips et tilts est conséquente (pouvant atteindre quelques microns!), ce qui a pour effet de déplacer les sous-PSFs dans le champ, en les séparant les unes des autres (distance inter-PSF bien supérieure à λ/D). Ces dernières n'interfèrent donc que très peu entre elles. Par conséquent, chaque segment agit comme une pupille indépendante, et le piston de cette dernière ne peut être identifié. Dans notre contexte l'unique moyen de mesurer les erreurs de piston est lorsque les sous-PSFs se superposent dans le plan-focal.

Limitation de l'identification du piston

Cependant l'estimation de ce piston est limitée par l'ambiguïté de λ : en effet le piston à la particularité d'avoir une périodicité de λ empêchant la bonne mesure de la valeur du piston. C'est à dire que les intensités de l'image sont égales pour une valeur de piston $c_{p1} = p1$ et $c'_{p1} = p1 + \delta$ à $\delta = \lambda$ la différence de marche. Voir figure 1.10.

1. Formation d'image – 1.4. Formation d'image d'un point source : la PSF d'un télescope segmenté

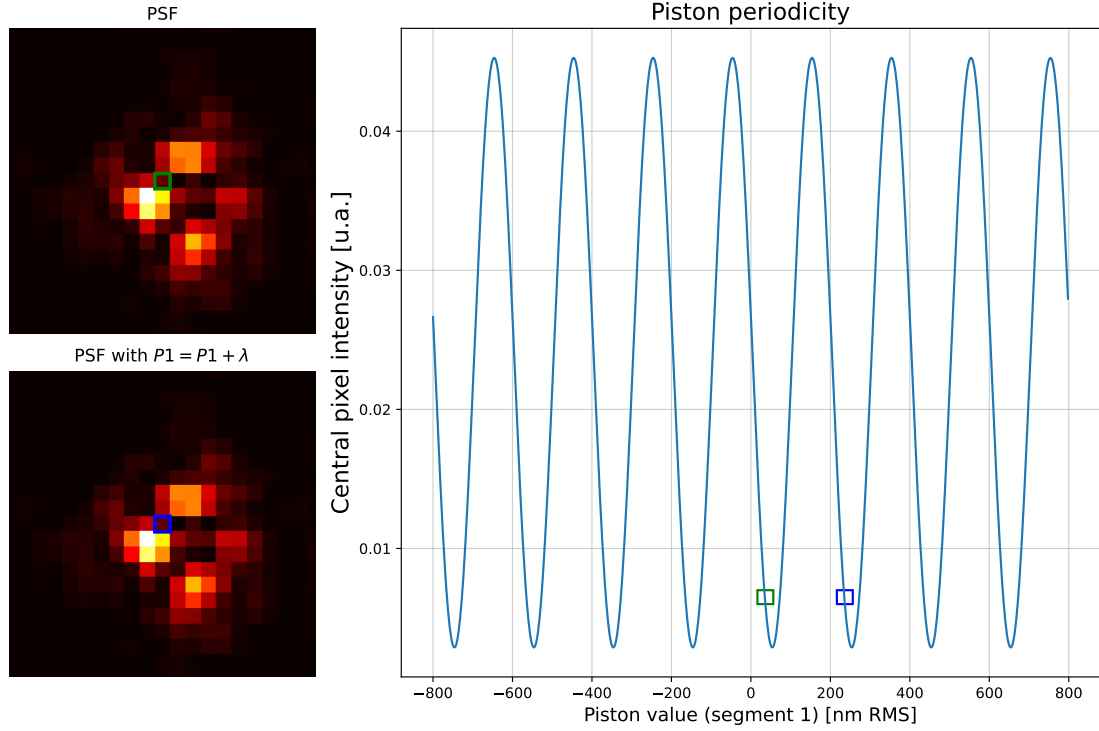


FIGURE 1.10. – Exemple de la relation entre piston et intensité des pixels (ici le pixel central de l'image est choisi). A gauche 2 PSF similaires sont présentées. La PSF du bas représente une PSF générée avec le piston du segment 1 à une valeur de $+\lambda \text{ nm PV}$ par rapport au piston 1 de la PSF du haut. Le graphique de droite représente l'évolution de l'intensité du pixel central de la PSF en fonction du piston 1.

On peut observer schéma de droite de la figure 1.10 une périodicité de 200 nm RMS de l'intensité du pixel central encadré sur les PSFs présentées à gauche. D'après la base présentée équation 1.20, la périodicité équivaut à 800 nm PV , soit la longueur d'onde λ . L'identification du piston en plan-focal mono-image est donc limitée : $\hat{p} \equiv p \pmod{\lambda}$. Cette ambiguïté de λ s'applique dans n'importe quel cas d'imagerie. Pour l'imagerie utilisant des pupilles segmentées les mesures des pistons par segments ou inter-segment seront faites modulo λ . On cherche alors à identifier un intervalle de piston pour lequel l'identification en plan-focal ne souffrira pas d'ambiguïté de lambda. En supposant une ambiguïté de λ durant une estimation, les coefficients de pistons estimés sont tels que :

$$\hat{c}_p = c_p + [\delta_1, \delta_2, \delta_3, \delta_4] \text{ tq } \begin{cases} PSF(c) = PSF(\hat{c}) \\ \sum \delta_k = 0 \end{cases} \quad (1.27)$$

ou c_p correspond aux coefficients de piston réels, $\hat{c}_p$ les coefficients de piston estimés et δ_i le décalage des valeurs. La première condition impose que $\forall i \neq j, |\delta_i - \delta_j| = k\lambda$, et la seconde condition est dictée par la mise à 0 de la moyenne des pistons.

1. Formation d'image – 1.5. Cas d'un objet étendu

On se retrouve dans deux cas de figure :

$$\begin{aligned} & \delta_i = \delta_j = -\delta_k = -\delta_l \text{ avec } i \neq j \neq k \neq l \\ \text{ou} \quad & \delta_i = -3\delta_{j \neq i} \end{aligned} \quad (1.28)$$

En cherchant alors l'intervalle minimum pour lequel des sauts de lambda peuvent survenir lors de l'estimation des pistons, on développe le premier cas de l'équation 1.28.

$$\begin{cases} \delta_i = -\delta_j \\ \delta_i - \delta_j = \lambda \end{cases} \quad (1.29)$$

et donc

$$\begin{cases} \delta_j = \delta_k = -\frac{\lambda}{2} \\ \delta_i = \delta_l = \frac{\lambda}{2} \end{cases} \quad (1.30)$$

On cherche donc un intervalle qui assure que l'estimation se fera sans ambiguïté de lambda, empêchant les différence de marche δ_i minimum de $\frac{\lambda}{2}$, soit un intervalle

$$p \in]-\frac{\lambda}{4}, \frac{\lambda}{4}[\quad (1.31)$$

Cet intervalle permet d'éviter tout saut de $-\frac{\lambda}{2}$ pour un segment et $+\frac{\lambda}{2}$ pour un autre. Des résultats concernant cette identification seront présentés partie 3.3.

N.B : la seconde équation de équation 1.28 mène à des sauts δ_i tels que :

$$\begin{cases} \delta_i &= \frac{\lambda}{4} \\ \delta_{k \neq i} &= -\frac{3\lambda}{4} \end{cases} \quad (1.32)$$

Ce cas de saut de λ ne devrait pas affecter notre étude étant donné les distribution relativement petite de piston pour le régime de déploiement et le régime de phasage moins. Néanmoins afin de s'assurer la bonne distribution des coefficient du jeu de données, les valeurs de piston sont tronquées afin d'assurer $|p| < \frac{\lambda}{4} \text{ nm RMS}$.

1.5. Cas d'un objet étendu

Dans le cas d'un scénario de phasage d'un télescope spatial pour l'observation de la Terre, un point source ne serait peut être pas accessible. Effectivement, un des scénario de phasage complet d'AZIMOV consiste à effectuer le phasage grossier sur un point source, puis, une fois en opération effectuer une boucle de phasage fin face à la Terre. L'objet d'étude n'est donc plus un point source mais un objet étendu très varié et non maîtrisé.

1.5.1. Relation objet-image

La modélisation de l'image d'un objet étendu est assez simple. On considère dans un premier temps que chaque points de l'objet o émet une onde électromagnétique

1. Formation d'image – 1.5. Cas d'un objet étendu

sphérique incohérente avec les autres points de l'objet. C'est une très bonne approximation dans notre cas. Ensuite on suppose que chaque point de l'objet va être vu de la même façon par l'instrument optique. La répartition de lumière dans le plan-focal sera donc la somme incohérente des réponses individuelles de chaque point de l'objet et s'écrit alors de la façon suivante :

$$I(\mathbf{r}) = \int \int o(\boldsymbol{\rho}) PSF(\mathbf{r}, \boldsymbol{\rho}) d\boldsymbol{\rho} \quad (1.33)$$

où o représente la distribution d'intensité de l'objet, $\mathbf{r}$ est la variable de position dans le plan image (sur le détecteur) et $\boldsymbol{\rho}$ est la variable de position angulaire dans le plan de l'objet et de la PSF. On peut noter la dépendance en $\mathbf{r}$ et $\boldsymbol{\rho}$ de la fonction PSF. C'est l'anisoplanétisme : c'est à dire que la PSF varie dans le champ due à une différence de chemin optique en fonction de l'angle de vue. C'est notamment le cas dans le cas d'étude d'objets spatiaux à partir de télescopes au sol : la lumière provenant de deux angles de vue différents ne traverse pas le même volume turbulent.

Dans le cas de l'observation de la Terre depuis l'espace, les aberrations perturbants la formation d'image au sein du télescope sont situées à proximité de l'instrument, dans des plans proches de la pupille et non tout le long de l'axe de visée, on peut alors considérer le domaine isoplanétique pour la formation d'image :

$$I(\mathbf{r}) = \int \int o(\boldsymbol{\rho}) PSF(\mathbf{r} - \boldsymbol{\rho}) d\boldsymbol{\rho} \quad (1.34)$$

On reconnaît alors l'opération de convolution entre la figure d'intensité de l'objet et la réponse impulsionnelle du système optique.

$$I(\mathbf{r}) = \{o \otimes PSF\}(\mathbf{r}) \quad (1.35)$$

On peut alors étendre l'expression ci-dessus au domaine de Fourier on obtient alors

$$\tilde{I}(\mathbf{r}) = \{\tilde{o} \times OTF\}(\mathbf{r}) \quad (1.36)$$

On comprend ici que l'image résultante se résume à l'opération d'un filtre passe bas, soit une multiplication fréquence spatiale à fréquence spatiale entre transformée de Fourier de l'objet et OTF. Cette de formation d'image impose un problème mal posé dans le but de retrouver la phase à partir de l'image. En effet, les variations de l'image sont alors soit dues aux variations de la PSF, soit à la géométrie de l'objet d'étude, les deux étant inconnues. À partir d'une seule image en plan focal, il faut alors dissocier l'information de l'objet de celle de la phase. Plus de détail sur la formation d'image sur objet étendu seront abordés au chapitre 5. Un exemple d'image est donnée ci dessous.

1. Formation d'image – 1.5. Cas d'un objet étendu

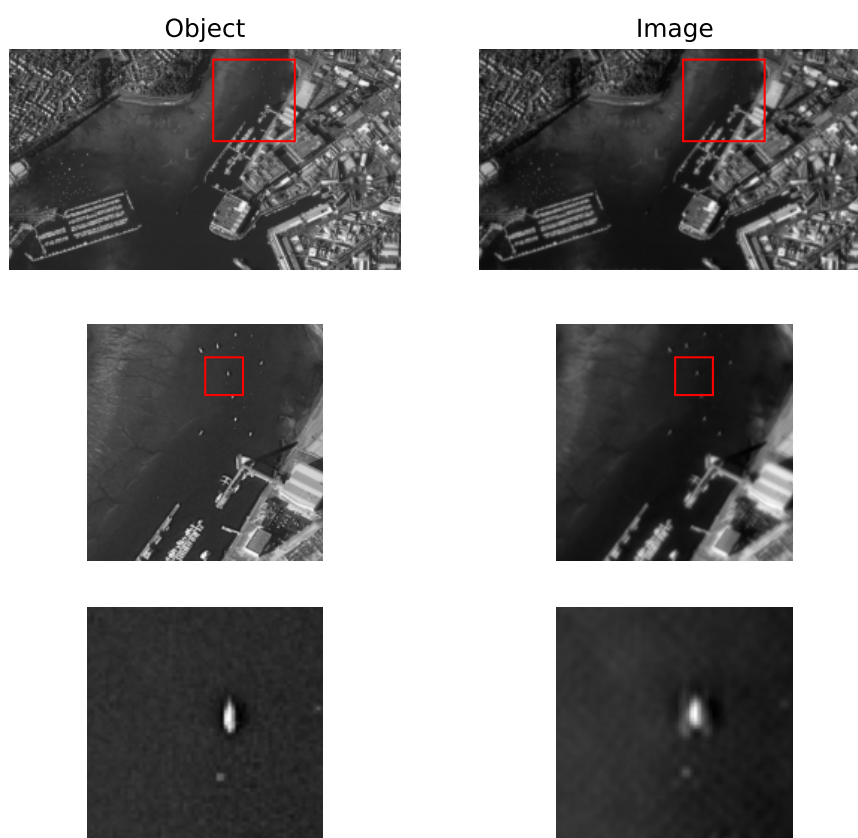


FIGURE 1.11. – Exemple d'objet (colonne de gauche) et d'images correspondante (colonne de droite).

On observe figure 1.11 une scène étendue d'un port. La convolution de l'objet par la PSF ne semble dans un premier temps pas dégrader l'image à vue d'oeil. Cependant c'est en s'intéressant aux petits objets presque ponctuel de l'image comme les bateaux par exemple que l'on peut observer alors dégradations de l'image due à la PSF. De plus, les objets défilant lors de l'observation de la Terre depuis l'orbite basse d'un télescope sont non répétitifs, étendus et inconnus. Le choix du détecteur étant de plusieurs milliers de pixels (une fauchée approximative de 5 km et une résolution de 1 m par pixel représente alors au minimum un détecteur d'environ 5000×5000 pixels), il n'est pas raisonnable d'imaginer une telle image en entrée d'un réseau de neurones. Cela nécessiterait une architecture beaucoup plus conséquente et moins embarquable à bord d'une petite plateforme. On cherche alors à extraire de petites images caractéristiques et porteuses d'informations utiles sur la phase. Un des avantages majeurs

de l'isoplanétisme est de pouvoir sélectionner indépendamment de la région dans le champ les images caractéristiques sans se soucier des variations de la PSF.

1.5.2. Sélection d'images caractéristiques

D'après l'isoplanétisme, on peut alors sélectionner une portion de l'image indépendamment de sa position, la PSF du système optique sera la même. Cela permet de se concentrer sur des régions finement choisies d'une image étendue pour effectuer l'analyse du front d'onde. Prenons un exemple d'images 64×64 extraite à partir d'une image. Tout d'abord à partir des coupes de la FTO de la figure 1.8, on observe que, notamment dans le cas du phasage fin, les variations dues aux dégradations de la PSF apparaissent principalement au niveau des hautes fréquences spatiales. En terme d'objet cela signifie qu'un objet contenant ces hautes fréquences spatiales sera donc plus influencé par les variations de la PSF.

1. Formation d'image – 1.5. Cas d'un objet étendu

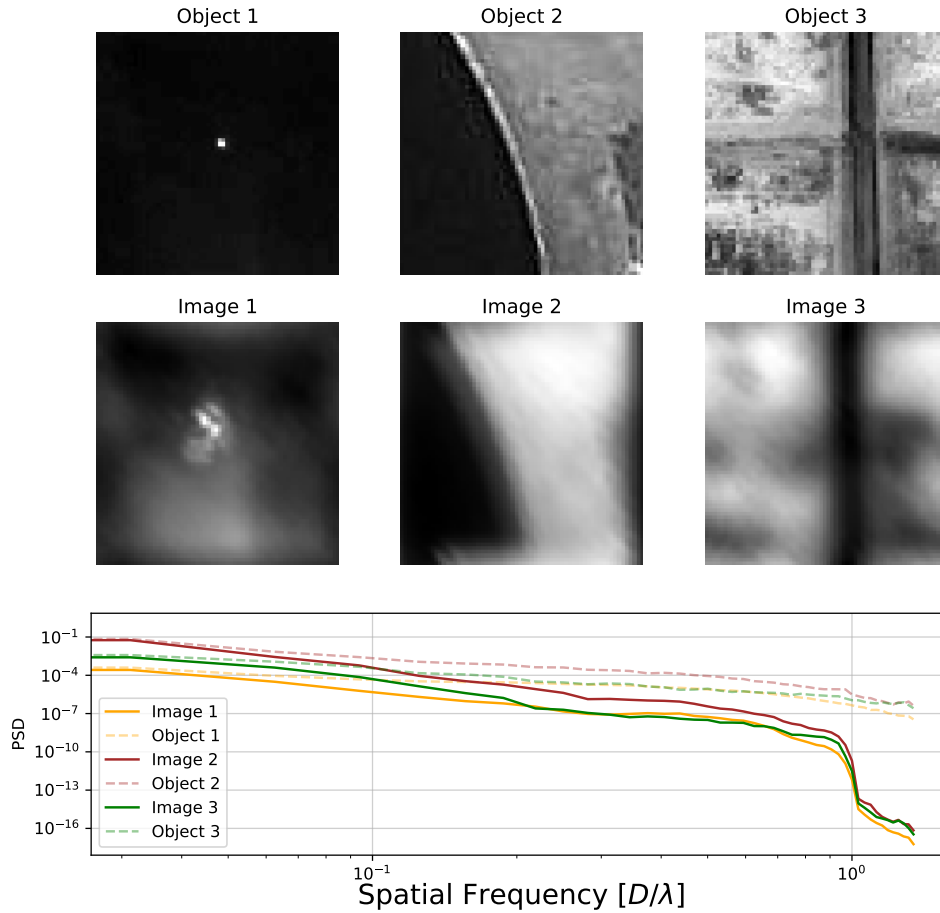


FIGURE 1.12. – Exemple d'objet observable et du contenu spectral une fois imagé par le télescope non phasé.

La figure 1.12 illustre les objets étudiés en première ligne, les images formées par le télescope non phasé en dessous, puis le contenu fréquentiel (**Densité Spectrale d'Energie, Power Spectral Density (PSD)**) correspondant à la dernière ligne. On observe nettement l'influence de la formation d'image sur le contenu spectral des images. La forme caractéristique de la FTO de la pupille couplée aux aberrations de phasage engendre une perte conséquente du contenu spectral, notamment autour des hautes fréquences. De plus, les fréquences au-delà de D/λ sont perdues. C'est intrinsèque au télescope : au-delà d'une certaine fréquence spatiale, le contenu ne peut être imagé car la FTO est définie sur un support limité. La signature des aberrations de la PSF se retrouve sur toute l'image. Néanmoins, la signature des aberrations est d'avantage observée lorsque l'objet étendu observé est très petit (inférieur ou égal à la largeur à mi hauteur de la PSF). En effet, les objets 2 et 3 étant très étendus, l'image issue de la

convolution de ces derniers par la PSF mixe l'information de la PSF en plan-focal et peut alors perturber l'analyse de front d'onde : il est alors impossible, avec une seule image, de dissocier si les amplitudes des images sont issues de l'objet ou de la PSF.

Le point source étendu On propose donc dans un premier temps d'étudier une forme d'objet récurrente dans l'imagerie de la surface de la Terre, les *points source étendus*. Ils peuvent représenter des objets brillants à la surface de la Terre comme des bouées en mer, une vague ou encore des voitures sur une route. Ils permettent de valider les résultats vu pour un point source dans un cas similaire. Des exemples de ce types d'objets peuvent être trouvés en figure 1.13.



FIGURE 1.13. – Exemple d'objet point source étendu.

1.5.3. Base de données d'objet

Afin de constituer une base de données à partir des points sources étendus, ces derniers vont être présélectionnés à l'aide d'un algorithme se basant sur le gradient local de l'image et une sélection sur l'intensité des pixels. Les imagerie candidates seront ensuite manuellement sélectionnées et ajoutées à la base de données.

Pléiade On se base d'une part sur les données acquises par le satellite Pléiade-1B (voir [CNES](#)), un satellite imageur en orbite basse avec une pupille d'un peu plus de 60 cm d'ouverture. Le satellite Pléiade-1B propose des images panchromatiques à une résolution de 50 cm à Nadir. On propose donc de biner les images afin de retrouver la résolution métrique de notre télescope. Cette base de données permet une pré-sélection de plus de 200 images de point sources étendus pour notre étude.

Planet La société Planet (voir [planet](#)) proposent aussi des données d'observation de la Terre. Avec leur constellation de 200 satellites, dont 21 satellites SKYSAT ([[Mur+14](#)]), autorisant tout comme Pléiades-1B, une résolution au sol de 50 cm après post-processing. Les données *planet* permettent l'extraction de plus de 200 points sources étendus manuellement sélectionnés.

Simulation Le nombre de données accessible étant limité, une section portant sur la simulation d'objet étendu sera développée partie 5. En somme, les objets utilisés

1. Formation d'image – 1.5. Cas d'un objet étendu

proviennent en partie d'image réelle de la surface de la Terre. Ils constituent une base de données limitées, permettant néanmoins une première analyse sur l'AFOPF par intelligence artificielle sur objet étendu.

1. Formation d'image – 1.5. Cas d'un objet étendu

Ce chapitre présente le modèle physique utilisé dans notre application de l'intelligence artificielle. Le formalisme de la formation d'image sur point source et sur objet étendu a été présenté, ainsi que les problématiques et limites de l'analyse de front d'onde en plan-focal mono-image. Les distributions des paramètres et des objets d'étude ont également été présentées afin de fournir au lecteur une vue d'ensemble précise des problématiques de phasage de télescope. Le formalisme de la méthode de Deep Learning utilisée va être abordé dans le prochain chapitre.

2. Methodologie

Dans ce chapitre, je vais étudier la méthodologie employée pour identifier les erreurs de phasage à l'aide de réseaux de neurones. Tout d'abord, je présenterai les bases du Deep Learning et des réseaux de neurones afin d'en expliquer les concepts fondamentaux. Ensuite, j'aborderai la stratégie d'apprentissage des réseaux. Par la suite, j'examinerai les différentes architectures utilisées dans ces travaux de thèse. Enfin, je décrirai la méthodologie finale appliquée au phasage.

Sommaire

2.1. Introduction aux réseaux de neurones	75
2.1.1. Principe	76
2.1.2. Multi-Layer Perceptron	76
2.1.3. Réseaux de neurones à couches de convolution	79
2.1.4. Fonction d'activation	81
2.1.5. Loss	82
2.1.6. Optimiseurs et retropropagation des gradients	83
2.2. Stratégie d'apprentissage	84
2.3. Architecture	85
2.3.1. AlexNet & VGG-Net	85
2.3.2. ResNet	87
2.3.3. EfficientNetB0	88
2.3.4. Variational Auto-Encoder	90
2.4. Méthodologie pour le phasage	93
2.4.1. Simulation de données	93
2.4.2. Choix de la loss	93
2.4.3. Hyperparametres	94
2.4.4. Le phasage par Deep Learning	94

2.1. Introduction aux réseaux de neurones

La recherche de la machine intelligente a poussé les chercheurs à imaginer des logiciels et algorithmes automatiques pour remplacer l'expertise humaine. Le Machine Learning vient apporter un changement de paradigme dans la programmation, permettant la résolution de modèles complexes par une approche inductive basée sur les données plutôt qu'une approche classiquement déductive. L'apprentissage successif sur ces données permet la génération et l'optimisation de ces algorithmes. Mais comment sont-ils réalisés? Ce chapitre est dédié au développement des méthodes de Deep Learning, un sous-ensemble du Machine Learning.

L'inspiration par la nature en technologie a permis un développement dans les techniques actuelles. Cependant, elle se retrouve vite limitée par nos modélisations : à l'image des oiseaux qui nous ont inspirés pour voler, les avions ne battent pas des ailes pour nous transporter.

Pour cette raison, l'application du fonctionnement des neurones biologiques aux réseaux neurones en programmation présente dans un premier temps une similarité nette avec leurs cousins biologiques, puis a tendu à s'en éloigner peu à peu pour devenir une technologie plus compétitive.

"You can get the best of both world by being open to biological inspirations without being afraid to create biologically unrealistic models, as long as they work well."

2.1.1. Principe

Le Deep Learning est une sous-discipline de l'apprentissage automatique (*machine learning*), étant lui-même une branche de l'intelligence artificielle. Dans le cadre du machine learning, on cherche à concevoir des modèles capables de tirer des enseignements d'une large quantité de données. Le Deep Learning utilise des architectures complexes, appelées réseaux de neurones, constituées de nombreuses couches de neurones artificiels. Ces réseaux peuvent ainsi traiter des volumes de données bien plus grands et modéliser des relations complexes entre les variables.

Le principe fondamental est d'apprendre les paramètres du réseau (poids, biais et filtres) pour chaque connexion entre les neurones, afin de modéliser efficacement les données. L'entraînement consiste à ajuster ces paramètres du réseau pour minimiser l'erreur entre les prédictions du modèle et les résultats réels. En Deep Learning, cette erreur est appelée fonction de perte (*loss function* en anglais), et l'optimisation de cette fonction est cruciale pour améliorer les performances du réseau. Par exemple, une technique populaire pour minimiser la perte est la descente de gradient, qui ajuste les poids progressivement en fonction de la direction qui réduit le plus la perte.

L'apprentissage en Deep Learning peut se faire de trois manières :

- **Supervisé**, où chaque donnée d'entraînement est labelisée, ce qui facilite la tâche du modèle en lui fournissant une référence.
- **Non supervisé**, où les données ne sont pas labelisées, et le modèle doit trouver lui-même des structures et des similarités.
- **Semi-supervisé**, une combinaison des deux méthodes, où une partie des données est labelisée, mais une grande quantité reste sans label.

Dans notre étude, c'est un apprentissage supervisé qui sera choisi. En effet les données étant simulées ou expérimentales, les images d'entrée sont générées avec leur labels correspondants. Les réseaux de neurones se distinguent également par leur non-linéarité, introduite par des fonctions d'activation, qui permettent de modéliser des relations complexes et non linéaires entre les données d'entrée et la sortie. Grâce à cette non-linéarité, les réseaux de neurones peuvent capturer des schémas plus profonds et plus abstraits, ce qui est essentiel pour des applications telles que la reconnaissance d'image, le traitement du langage naturel, ou encore la conduite autonome. Une fois entraîné, les paramètres du réseaux sont fixés et ce dernier fonctionne alors comme un estimateur déterministe.

2.1.2. Multi-Layer Perceptron

Inspiré par les neurones biologiques, les premiers réseaux de neurones artificiels sont composés d'une multitude de neurones connectés entre eux.

Définition d'un neurone

Un premier modèle simpliste est présenté comme un neurone artificiel par Frank Rosenblatt : le *Threshold Logic Unit* (TLU) ou *perceptron* dont l'entrée x et la sortie $\hat{y}$ sont des scalaires dont chaque connexion entre neurones est associée à un poids.

2. Methodologie – 2.1. Introduction aux réseaux de neurones

Le perceptron réalise alors la somme de ces entrées, ajoute un biais et applique une fonction non linéaire à cette somme appelée *fonction d'activation*. Par conséquent, le perceptron réalise l'opération mathématique suivante :

$$\hat{y} = a(z) \text{ avec } z = \sum_{i=1}^p w_i x_i + b = \mathbf{x}^T \mathbf{w} + b \quad (2.1)$$

avec a la fonction d'activation, p le nombre de neurones, w_i les poids des connections, et b le biais associé. On dissocie alors la partie transformation affine en scalaire z de la partie transformation non linéaire par la fonction d'activation a . Une représentation d'un neurone artificiel et de son homologue biologique est présentée en figure 2.1.

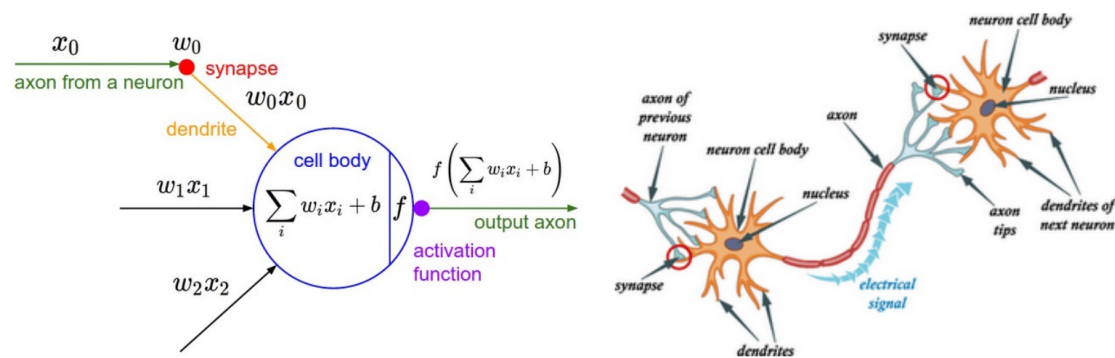


FIGURE 2.1. – Illustration d'un perceptron (gauche) et comparaison avec un neurone biologique (droite). Extrait des cours de M.Gray.

Mise en relation de neurones : couches denses

On peut maintenant facilement imaginer une succession de neurones mis en connexion les uns aux autres à très grande échelle afin de multiplier les paramètres internes du réseau (les paramètres définissant un neurone : poids et biais) à ajuster lors de l'apprentissage. On parle alors de couches, plus spécifiquement de couches entièrement connectées. On peut reconnaître dans une architecture entièrement connectée la couche d'entrée, les couches cachées et la couche de sortie. Une architecture entièrement connectée de la sorte est aussi appelée *Multi-Layer Perceptron (MLP)*.

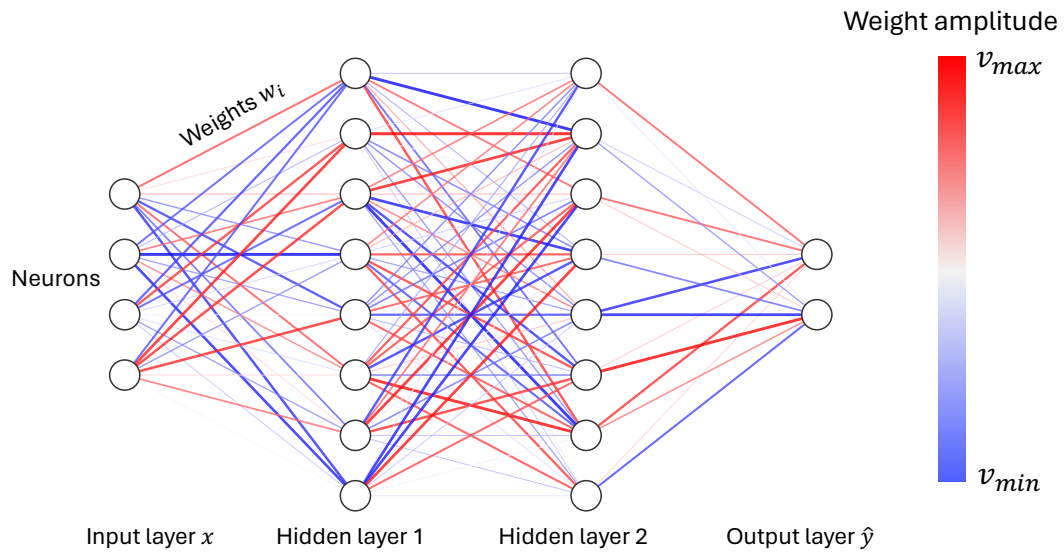


FIGURE 2.2. – Illustration d’un MLP. Dans cet exemple, les poids sont tirés aléatoirement et sont représentés en couleur et en intensité en fonction de leur amplitude. De gauche à droite : couche d’entrée, couche cachée, couche cachée, couche de sortie.

La figure 2.2 présente un exemple d’un MLP. En particulier, on peut remarquer particulièrement que le nombre de poids de $\mathbf{w}$ se détermine par toutes les connexions en amont d’un neurone : chaque neurone de la première couche cachée comptabilise ici 4 poids en amont correspondant au nombre de connexion avec la couche d’entrée du réseau, tandis que les neurones de la couche de sortie comptabilisent eux 8 poids. Le nombre de neurones d’entrée et de sortie est déterminé par la dimension du problème à résoudre. Dans le cas de la régression sur l’estimation de n coefficients à partir d’une image de taille $H \times W$, le réseau comprendra alors $H \times W$ neurones d’entrée et n neurones de sortie. Le nombre de neurones des couches cachées est plus complexe à déterminer, certains préfèrent une forme pyramidale des réseaux, mais cette pratique a été abandonnée en faveur de l’utilisation de réseaux dont les couches entièrement connectées comportent le même nombre de neurones (comme montré ici), réduisant ainsi le nombre d’hyperparamètres du système à ajuster.

Les MLPs ont prouvé leur efficacité dans le domaine de l’intelligence artificielle depuis des années. Cependant, ils se retrouvent rapidement limités dans l’analyse d’images 2D : la couche d’entrée n’est composée que d’un vecteur $\mathbf{x}$. Cela correspond à la linéarisation d’une image 2D et donc à une perte conséquente d’information spatiale. De plus, pour des images de grandes dimensions, les réseaux entièrement connectés voient leur nombre de paramètres exploser.

2.1.3. Réseaux de neurones à couches de convolution

Une nouvelle fois, l'inspiration par la nature vient pallier ce défaut. Cette fois-ci la structure du cortex visuel des chats étudiée par David H. Hubel et Torsten Wiesel montre des régions de récepteurs de champ local, discrétisant la vision en sous-champs locaux, pouvant se superposer et qui, une fois mis en relation, décrivent toute la scène d'observation du champ de vue. Ces récepteurs, directement connectés au champ de vue occupent une fonction distincte : certains vont identifier les zones de traits horizontaux, d'autres les structures courbées, ou encore les zones ponctuelles de la scène. En 1980, Fukushima ([Fuk80]) imagine une analogie en traitement d'image : utiliser des filtres de convolution afin d'extraire les informations et les structures porteuses d'information d'une image.

L'opération de convolution

Dans le domaine du Machine Learning, l'opération de convolution peut se dissocier légèrement de l'opération de convolution classique vue précédemment équation 1.36. Ici on fait référence à une image d'entrée I convoluée à un kernel K de taille $(h \times k)$ et dont la sortie est une carte de caractéristiques C . On peut alors mathématiquement définir cette opération telle que la multiplication pixel à pixel en superposant le kernel à l'image :

$$C(i, j) = a((I \otimes K)(i, j)) = a\left(\sum_{m=0}^h \sum_{n=0}^k I(i-m, j-n)K(m, n)\right) \quad (2.2)$$

Cette opération de convolution par un kernel est alors effectuée sur l'ensemble des pixels (m, n) de l'image I de taille $H \times W$ avec un décalage du filtre par un pas appelé *stride*, puis on applique éventuellement une fonction d'activation a .

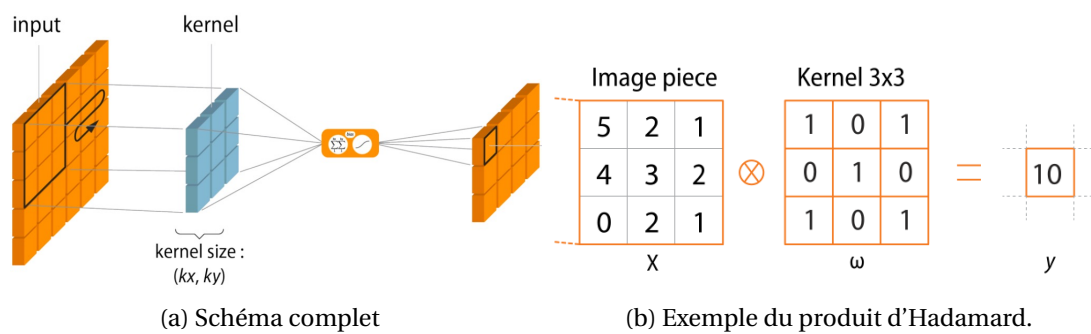


FIGURE 2.3. – Illustration de la convolution (produit d'Hadamard) d'une image par un kernel. Extrait de FIDLE.

C'est donc le kernel, dont l'exemple est donné au milieu de la figure 2.3b, qui va être appris au cours de l'apprentissage du réseau. Ses coefficients vont être ajustés au fur et à mesure de l'entraînement du réseau afin qu'un filtre donné permette

d'extraire une caractéristique spécifique. Dans un cas simple, le filtre se contente d'identifier des figures interprétables comme des bords, des courbes, des cercles ou figures géométriques diverses, mais lorsque la tâche devient plus complexe, alors il est plus difficile d'interpréter les composants du filtre et les caractéristiques qui découlent de la convolution.

La couche de convolution

Appliquer cette opération de convolution à une image permet de révéler certaines caractéristiques de l'image pour une échelle spatiale donnée, suivant la profondeur dans les couches du réseau de neurones où se trouve le filtre de convolution. L'idée d'une couche convolutive est donc d'apprendre une multitude de filtres de convolution par couche afin d'extraire un maximum d'informations utiles et différentes de l'image. Chaque filtre, une fois convolué par l'image d'entrée génère une carte de caractéristiques C différente. Le nombre de cartes de caractéristiques en sortie d'une couche est donc égal au nombre de filtres choisis.

Pooling layers

Une fois l'ensemble des convolutions d'une couche effectuée et les fonctions d'activation correspondantes calculées, une dernière étape de pooling est effectuée. L'objectif des couches de convolution étant d'identifier et de garder uniquement l'information utile d'une image, réduire la dimensionnalité des images de sortie permet de filtrer l'information utile. Cela permet au réseau de se concentrer sur les éléments les plus importants de l'entrée et d'ignorer les informations redondantes ou non pertinentes. Cette sélection s'appelle *pooling* et est généralement effectuée par un filtre 2 par 2 gardant uniquement le pixel de plus haute intensité (max pooling) de la fenêtre et retournant alors une image deux fois plus petite en sortie. Le pixel de plus haute intensité est généralement choisi car il représente la caractéristique la plus importante de la fenêtre de sélection et permet un filtrage du bruit. Le schéma de l'opération de maxpooling peut être retrouvé figure 2.4.

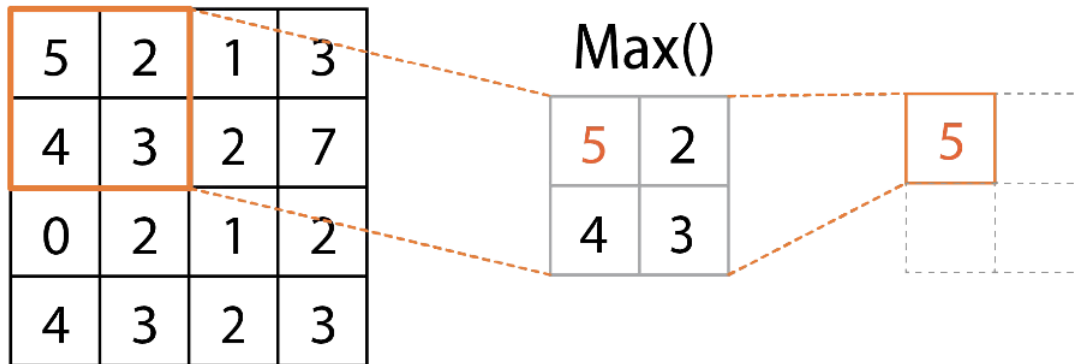


FIGURE 2.4. – Exemple de max pooling. Extrait de [FIDLE](#).

Cela permet au fur et à mesure des couches de diminuer le nombre de paramètres et donc la mémoire utilisée. L'opération de pooling apporte une non-linéarité supplémentaire au réseau et est invariante aux petites translations, accordant donc l'importance aux caractéristiques plutôt qu'à leur localisation précise. Cette opération ne comporte pas de paramètre optimisable.

2.1.4. Fonction d'activation

Mentionnées plus tôt, les couches d'activation correspondent à des fonctions non linéaires. Le terme de "fonction d'activation" vient de l'équivalent biologique "potentiel d'activation", seuil de stimulation qui, une fois atteint, entraîne une réponse du neurone. Elles effectuent une transformation non linéaire qui modifie la représentation spatiale des sorties d'une couche, permettant le passage simple ou la transformation d'une information en sortie. L'intérêt des fonctions d'activation est d'éviter la combinaison d'opérations linéaires entre neurones, dans le cas contraire les réseaux stackés en plusieurs couches seraient équivalents à un réseau à une couche unique linéaire, ne permettant pas la résolution de problèmes complexes et encore moins non linéaires. Les fonctions d'activation ont leurs caractéristiques propres :

- **Non-linéarité** : Cette non-linéarité crée une frontière entre chacune des couches pour qu'elles ne puissent pas se simplifier.
- **Différentiable par morceau** : Cette propriété permet de créer des optimisations basées sur les gradients.
- **Étendue** : Quand la plage d'activation est finie, les méthodes d'apprentissage basées sur les gradients sont plus stables (impact sur un nombre de poids limités). Quand la plage est infinie, l'apprentissage est généralement plus efficace (impact sur davantage de poids).
- **Monotone** : Lorsque la fonction est monotone, la surface d'erreur associée avec un modèle monocouche est certifiée convexe et donc minimisable ([\[Wu09\]](#)).
- **Douce (dérivée monotone)** : Les fonctions à dérivée monotone ont été montrées comme ayant une meilleure capacité à généraliser dans certains cas.

- **Identité en 0** ($f(x) \approx x$ quand $x \approx 0$) : Ces fonctions permettent de faire un apprentissage rapide en initialisant les poids de manière aléatoire.

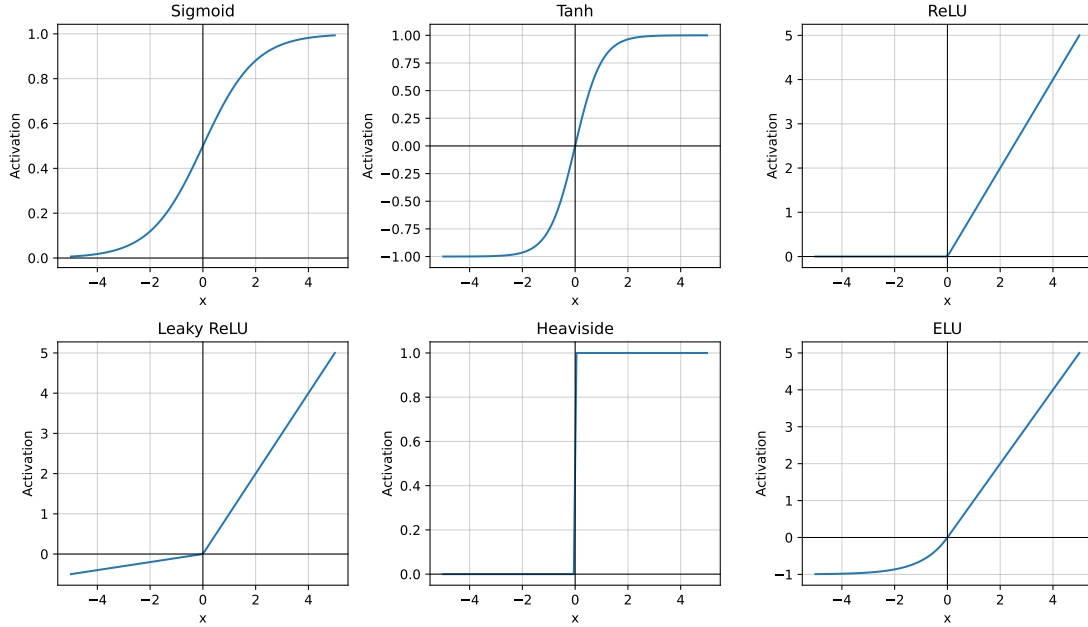


FIGURE 2.5. – Exemple des fonctions d’activation usuelles.

Dans ce manuscrit, nous allons particulièrement utiliser la fonction *Regularized Linear Unit* (ReLU) présentée en figure 2.5. C’est une fonction continue, non différentiable en 0 qui possède une dérivée nulle pour les valeurs négatives. En pratique, elle est très utilisée comme fonction d’activation de base. La fonction ReLU est préférable, car il en résulte la formation de réseau de neurones plus rapide ([KSH12]), sans faire une différence significative sur leur performance.

2.1.5. Loss

Le propagation en avant d’un réseau de neurones consiste en l’estimation d’une quantité $\hat{y}$ à partir de son entrée x . On peut modéliser la propagation du réseau paramétré par l’ensemble de ses paramètres θ tel que pour une donnée i , $f_{\theta}(\mathbf{x}_i) = \hat{\mathbf{y}}_i$. Dans le cas d’un apprentissage supervisé, dont on connaît la quantité réelle $\mathbf{y}$, on peut alors contrôler la qualité d’estimation par rapport à la vérité de terrain. C’est ce type d’apprentissage qui sera utilisé dans ce manuscrit. On peut donc définir une loss $\mathcal{L}$ sur l’ensemble des données qui permet de quantifier les erreurs d’estimation. On a alors, pour une paire $(\mathbf{x}_i, \mathbf{y}_i)$, une loss associée $l(\mathbf{y}_i, \hat{\mathbf{y}}_i)$, avec l la loss. On définit alors la loss totale $\mathcal{L}$ telle que la moyenne des losses sur l’ensemble $\mathcal{A} = (\mathbf{x}_i, \mathbf{y}_i)_{i \in 1, 2, \dots, N}$ de

données et de leurs labels correspondants :

$$\mathcal{L}_{\theta}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N l(\mathbf{y}_i, \hat{\mathbf{y}}_i) = \frac{1}{N} \sum_{i=1}^N l(\mathbf{y}_i, f_{\theta}(\mathbf{x}_i)) \quad (2.3)$$

On souhaite logiquement, minimiser cette quantité de façon à obtenir des quantités estimées aussi proche que possible des quantités réelles.

2.1.6. Optimiseurs et rétropropagation des gradients

Je propose une description purement qualitative de l'actualisation des paramètres du réseau. Pour une description plus complète et théorique de la démarche, je revoie le lecteur à [Goodfellow](#).

Afin d'entraîner un réseau de neurones, l'algorithme d'apprentissage supervisé vise à ajuster les paramètres du modèle (poids, filtres et biais) pour minimiser la loss, qui mesure l'écart entre les prédictions du modèle et les valeurs attendues. La méthode principale pour accomplir cette optimisation est appelée la rétropropagation des gradients (*backpropagation* en anglais), un processus essentiel dans le Deep Learning. La rétropropagation est une technique qui, à partir de l'erreur calculée sur les données de sortie, "remonte" dans le réseau en ajustant les poids de chaque couche. Elle utilise les dérivées partielles de la loss par rapport aux paramètres du modèle pour calculer les gradients et ajuste ces derniers de manière à réduire progressivement l'erreur. Concrètement, la rétropropagation se base sur la descente de gradient, qui ajuste les paramètres dans le sens opposé au gradient de la loss, pour atteindre un minimum local de cette fonction. Ce processus est itératif et se poursuit jusqu'à ce que la loss soit suffisamment petite ou que les changements soient négligeables. Typiquement, l'ajustement des paramètres θ d'une couche n d'un réseau, se fait tel que

$$\theta_k^{(n)} = \theta_k^{(n)} - \Gamma \frac{\mathcal{L}_{\theta}}{\partial \theta_k^{(n)}} \quad (2.4)$$

où Γ est un terme pondérateur. Pour une efficacité accrue, il est courant d'utiliser des optimiseurs plus avancés, qui adaptent dynamiquement la taille des ajustements des paramètres à chaque itération. Une méthode d'optimisation couramment utilisée dans les réseaux de neurones profonds est l'optimiseur Adam (pour *Adaptive Moment Estimation* ([KB14])). Adam adapte les taux d'apprentissage pour chaque paramètre en fonction des moments d'ordre un (moyenne des gradients) et d'ordre deux (moyenne des carrés des gradients), ce qui permet une convergence plus rapide et stable que la descente de gradient classique. Adam est devenu l'optimiseur de référence pour la plupart des modèles en Deep Learning, car il combine l'efficacité de plusieurs méthodes d'optimisation et améliore la stabilité de l'entraînement des modèles profonds.

En effet, l'algorithme Adam combine donc les avantages de deux autres optimiseurs populaires : l'optimiseur de momentum (moment d'ordre 1, [Sut+13]) et l'optimiseur de RMSProp (moment d'ordre 2, [Zou+19]) permettant d'ajuster les taux d'apprentis-

sage de manière adaptative pour chaque paramètre du modèle et étant efficace dans un large éventail de tâches d'apprentissage automatique.

Maintenant que tous les concepts nécessaires au réseau de neurones ont été développés ci-dessus, nous pouvons nous intéresser à la stratégie d'apprentissage.

2.2. Stratégie d'apprentissage

Répartition du jeu de données

Dans un premier temps, il faut constituer un jeu de données conséquent afin de représenter au mieux les multiples réalisations des données d'entrée possibles. Une fois le jeu de données créé, on peut alors le séparer en 3 parties. La première, la plus conséquente, est le jeu de données d'entraînement. Il représente généralement 70 à 80% de l'ensemble des données. Ensuite, 10 à 15% des données composent le jeu de validation. Ce jeu de données va permettre, à la fin de chaque époque (voir ci-dessous), de valider les performances du réseau sur des données qu'il n'a pas vues durant l'entraînement. On vérifie alors la généralisation du réseau au fur et à mesure des époques. Enfin, les 10 à 15% des données restantes sont utilisées pour le jeu de données de test. Une fois le réseau entraîné, ces données servent à vérifier les performances du réseau en opération.

Étape d'entraînement

1. **Batch** : Lors de l'entraînement, le jeu de données d'entraînement est séparé en plusieurs *batches*, principalement pour éviter des problèmes liés à un manque d'espace mémoire. Il assure notamment une stabilité lors de l'entraînement grâce à une variété d'éléments au sein de chaque lot : si les poids du réseau étaient mis à jour après chaque donnée (descente de gradient stochastique), les gradients peuvent alors être très fluctuants, rendant difficile la convergence vers un minimum global de la loss. En utilisant des *batches*, les gradients sont calculés comme la moyenne des gradients de plusieurs exemples, ce qui réduit la variance. Cela conduit à des mises à jour des poids plus stables et peut aider à éviter les minima locaux.
2. **Propagation** : À chaque étape, un batch est ainsi propagé dans le réseau afin de réaliser une estimation. La propagation en avant permet de transformer une entrée à travers une série de transformations linéaires et non linéaires pour produire une sortie qui peut être utilisée pour la régression ou la classification. C'est le processus fondamental qui permet au réseau de neurones de faire des inférences à partir des données. La sortie du réseau est ensuite comparée au résultat attendu dans le cas d'un apprentissage supervisé, et l'erreur est calculée.
3. **Backpropagation** ([[Ver05](#)]) : Pour chaque batch, c'est ensuite le principe mathématique de backpropagation qui permet de calculer les gradients des erreurs en fonction de chacun des paramètres du réseau, puis l'optimiseur vient mettre

à jour les différents paramètres du réseau dans chaque couche en fonction du paramètre de learning rate fixé par l'utilisateur.

4. **Époques, loss de training et loss de validation** : Lorsque l'ensemble des batchs du jeu de données est propagé à travers le réseau, l'ensemble des données d'entraînement ont donc été vues, les poids du réseaux sont actualisés ce qui termine une époque. Les données de validation sont alors propagées dans le réseau. On enregistre alors 2 losses : la loss d'apprentissage, quantifiant l'erreur sur les données d'apprentissage ; et la loss de validation, quantifiant l'erreur d'estimation sur le jeu de données de validation, non vu pendant l'apprentissage. Une fois ces losses calculées, une nouvelle époque démarre.

Dans l'ensemble de ce manuscrit, les entraînements des réseaux de neurones sont effectués à l'aide du centre de calcul du LAM, sur des GPU RTX et A50. Durant cette stratégie d'entraînement, plusieurs paramètres sont à choisir afin d'obtenir une valeur finale de la loss la plus petite possible, caractérisant une optimisation "optimale" des paramètres du réseau de neurones. Notamment, l'architecture du réseau doit être efficacement choisie pour correspondre aux dimensions du problème. La première étape dans le processus de conception d'un réseau neuronal est de choisir une architecture adaptée au problème à résoudre. Elle doit être sélectionnée en fonction des caractéristiques des données et des exigences du problème. Une fois l'architecture du réseau établie, on peut ensuite ajuster les hyperparamètres afin d'optimiser les performances du modèle. Dans cette optique, examinons les différentes architectures de réseaux utilisées durant la thèse et leur utilisation.

2.3. Architecture

Notre attention se porte particulièrement sur 4 architectures : l'AlexNet/VGG-Net, le ResNet, l'EfficientNet et enfin le *Variational Auto-Encoder* (VAE).

2.3.1. AlexNet & VGG-Net

Principe

L'architecture AlexNet ([KSH12]) est la première architecture de référence très performante pour les réseaux de neurones convolutifs. C'est notamment cette architecture qui a remporté le challenge ImageNet ILSVRC 2012 (challenge visant à évaluer et comparer les algorithmes de vision par ordinateur en classifiant des millions d'images dans mille catégories différentes) haut la main avec des performances nettement supérieures aux autres architectures. Elle est notamment la première architecture convolutive à empiler deux couches convolutives l'une au-dessus de l'autre.

C'est aussi le cas pour l'architecture VGG-Net ([SZ14]), qui remporta le concours ILSVRC 2 ans plus tard, avec son architecture tout aussi simple, combinant 2 à 3 couches convolutives suivies d'une couche de max pooling. Ce bloc successif répété 6 à 8 fois, est suivi d'un réseau dense de deux couches cachées et une couche de sortie, utilisant

uniquement des filtres 3×3 .

Application

Notre travail s'est donc inspiré de ces formes d'architecture afin de les appliquer à notre étude. En conséquence des contraintes énoncées plus tôt, ces architectures ont été modifiées et allégées dans le but d'étudier la taille idéale des réseaux permettant d'obtenir un phasage à la limite de la diffraction. On propose donc de concevoir notre propre réseau, que l'on nommera VGG-Net par simplicité. Notre VGG-Net comporte 5 couches convolutives, composées toutes de filtres 3×3 , et 3 couches entièrement connectées. La première couche convolutive est suivie d'un max-pooling (premier bloc convolutif) puis deux couches convolutives se suivent, suivies par un max-pooling, composant le second bloc convolutif, puis il en est de même pour les deux couches convolutives suivantes, formant le troisième bloc convolutif. La particularité de notre étude réside dans l'évolution du nombre de filtres de chaque couche. On nomme $N1$, $N2$ et $N3$ le nombre de filtres des blocs convolutifs 1, 2, 3, respectivement. On pourra retrouver dans le tableau Tab.2.1 un récapitulatif des filtres des couches convolutives et la taille des réseaux associés. Un schéma de l'architecture VGG-Net est également présenté ci-dessous.

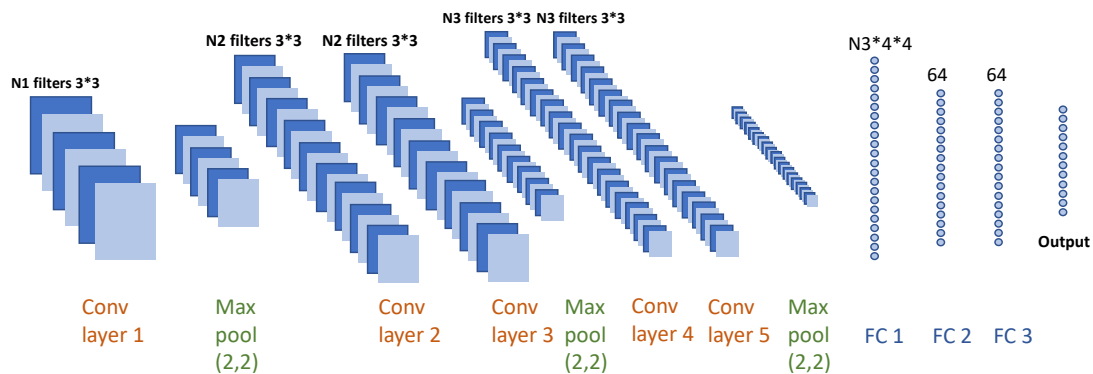


FIGURE 2.6. – Schéma de l'architecture VGG-Net utilisée.

On peut remarquer que le nombre de neurones de la première couche entièrement connectée (FC1) est dépendant du nombre $N3$ de filtres mais aussi d'un facteur 16. Ce facteur est déterminé par la taille de l'image d'entrée (ici, 64×64), cette image est alors transformée à travers les couches convolutives jusqu'à en extraire $N3$ imagerie de taille 4×4 .

2.3.2. ResNet

Principe

Un an plus tard, en 2015, c'est le ResNet ([He+16]) qui emporte le challenge ILSVRC. Cette architecture permet d'augmenter la profondeur des réseaux et d'assurer une bonne actualisation des poids du réseau : l'architecture du ResNet est très similaire à celle du précédent réseau. La différence réside dans la *skip connection* : le signal à l'entrée de la couche est ajouté en sortie de cette dernière en plus de la sortie originale de la couche. Formellement, si $\mathbf{x}$ est l'entrée d'une couche et $F(\mathbf{x})$ est la sortie de cette couche, la sortie finale de la couche sera $F(\mathbf{x}) + \mathbf{x}$.

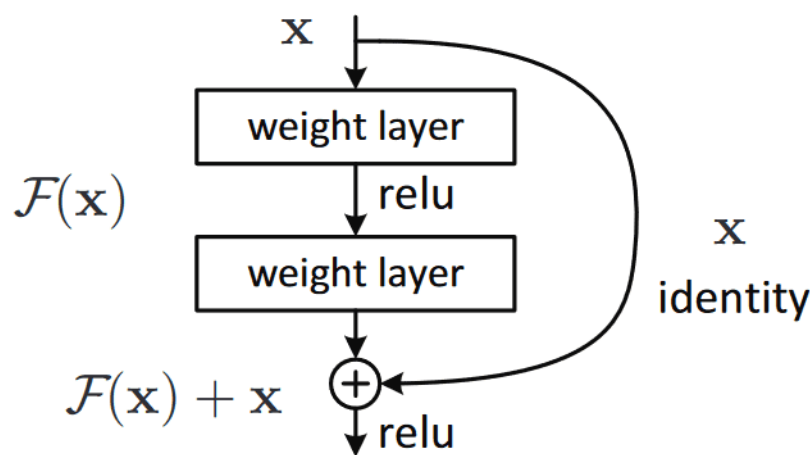


FIGURE 2.7. – Schéma d'une skip connection. Extrait de [AI summer](#).

Cela crée un chemin à travers lequel les informations peuvent être propagées sans être altérées. Lors de la backpropagation, cela permet une meilleure propagation du gradient jusqu'à la première couche du réseau, souvent difficilement modifiable à cause de la chute du gradient ([Hoc98]) et permet d'éviter des problèmes numériques de valeurs qui pourraient tendre vers 0.

Application

Dans notre étude, nous allons développer notre propre ResNet, du même nom. L'architecture est très semblable à celle de notre VGG-Net. On place uniquement une skip connection au départ de la couche Conv2 pour arriver en sortie de la couche Conv3, puis une autre au départ de la couche Conv4 jusqu'à la sortie de la couche Conv5. Un schéma de l'architecture utilisée est présenté en figure 2.8.

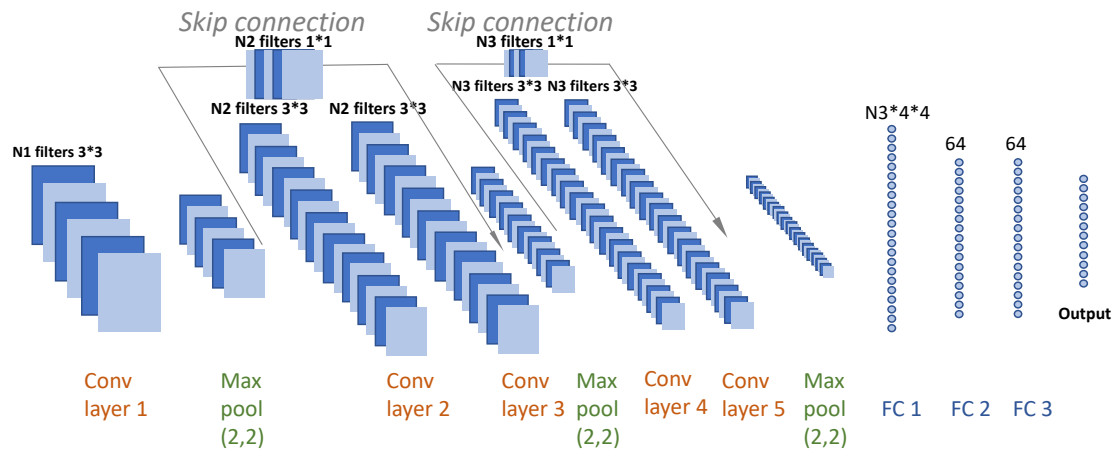


FIGURE 2.8. – Schéma de l'architecture ResNet utilisée.

Dans le cas du VGG-Net et du ResNet, c'est le nombre de filtres qui va venir déterminer la taille du réseau et donc le nombre de paramètres permettant d'estimer la phase à partir d'une PSF. On propose donc le tableau Tab.2.1 qui répertorie les différents réseaux et leurs tailles.

Label	N1	N2	N3	# parameters	Model size [MB]
NN1	4	8	16	13.10^3	0.134
NN2	8	16	32	31.10^3	0.278
NN3	16	32	64	94.10^3	0.94
NN4	32	64	128	326.10^3	2.6
NN5	64	128	256	$1.2.10^6$	11

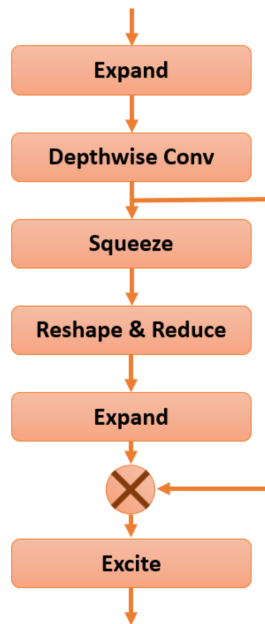
TABLEAU 2.1. – Taille des modèles Resnet en fonction du nombre de filtres convolutifs. Ce tableau donne également une approximation de la taille des réseaux VGG-Net, puisque l’architecture Resnet ne comporte que légèrement plus de paramètres internes en raison de la *skip connection* qui relie le nombre de filtres de la couche d’entrée au nombre de filtre de la couche de sortie.

2.3.3. EfficientNetB0

Principe

EfficientNet ([TL19]) est une architecture conçue pour atteindre un bon équilibre entre précision et efficacité calculatoire. Elle repose sur une méthode d'optimisation qui permet de dimensionner de manière automatique les différentes composantes du réseau, telles que la largeur, la profondeur et la résolution d'entrée. Cela permet

d'obtenir des modèles compacts et performants, adaptés à une utilisation sur des dispositifs avec des ressources limitées. L'EfficientNet a démontré des performances remarquables sur divers ensembles de données tout en étant plus efficace en termes de ressources par rapport à d'autres architectures plus complexes. L'EfficientNetB0 est la version de base de l'architecture EfficientNet. Conçue pour être efficace en termes de ressources tout en maintenant de bonnes performances, cette version initiale présente une taille relativement réduite par rapport aux versions plus avancées de l'EfficientNet. Malgré sa simplicité, l'EfficientNetB0 a démontré une capacité à obtenir des résultats compétitifs sur diverses tâches de classification d'images, en utilisant une combinaison optimisée de largeur, de profondeur et de résolution d'entrée. L'architecture est basée sur des blocs MBConv, qui combinent une convolution depthwise (chaque canal est convolué séparément), une expansion de dimension, une convolution 1x1 et éventuellement une opération d'attention squeeze-and-excitation. Ces blocs permettent d'extraire efficacement les caractéristiques des données tout en minimisant le nombre de paramètres et le coût calculatoire du modèle.



Opérateur	#Channels	#Layers
Conv3x3	32	1
MBConv1,k3x3	16	1
MBConv6,k3x3	24	2
MBConv6,k5x5	40	2
MBConv6,k3x3	80	3
MBConv6,k5x5	112	3
MBConv6,k5x5	192	4
MBConv6,k3x3	320	1
Conv1x1 & Pooling & FC	1280	1

FIGURE 2.9. – Exemple d'un bloc MBConv. Extrait de [Research-Gate](#).

TABEAU 2.2. – Tableau de récapitulatif de l'architecture EfficientNetB0.

Application

L'architecture EfficientNetB0 en tant que telle est initialement conçue pour la classification de 1000 classes à partir d'une image RGB. Afin de rendre le modèle adaptable

à notre utilisation, nous utilisons une couche d'entrée convolutive permettant de créer 3 canaux à partir de notre image d'entrée. Quant à la couche de sortie, elle est composée d'une couche entièrement connectée qui transforme les 1000 valeurs de sortie en 12 coefficients correspondant aux dimensions de notre étude.

2.3.4. Variational Auto-Encoder

Principe

Le Variational Auto-Encoder ([KW13]), introduit en 2013, est une extension des auto-encodeurs. Les auto-encodeurs sont des réseaux neuronaux utilisés pour apprendre une représentation compacte des données en entrée, appelée espace latent, en minimisant la différence entre les données d'entrée et les données reconstruites en sortie. Ils se composent de deux parties principales : un encodeur qui transforme les données d'entrée en une représentation compressée, et un décodeur qui reconstruit les données à partir de cette représentation. Les auto-encodeurs sont souvent utilisés pour la réduction de dimensionnalité, la détection d'anomalies ou la génération de données. Cependant, les autoencodeurs traditionnels, bien qu'efficaces dans la compression des données et la génération de répliques, présentent des lacunes importantes. Leur incapacité à fournir une représentation latente structurée et une mesure d'incertitude a conduit au développement des autoencodeurs variationnels. Les VAEs eux sont probabilistes, ils modélisent la distribution des données d'entrée dans un espace latent en utilisant des distributions probabilistes. En régularisant la distribution de la représentation latente, les VAEs offrent une meilleure généralisation et une gestion plus fine de l'incertitude. A partir de tirages de moyennes et d'écarts type d'une gaussienne sur l'espace latent, les VAEs peuvent générer des données similaires et non-identiques à celle d'entraînement. Les VAEs ne nécessitent pas d'entraînement supervisé et leur loss se compose de deux termes.

— Partie loss de reconstruction

Cette partie de la loss mesure la différence entre les données d'entrée originales x et les données reconstruites $\hat{x}$. Elle est souvent calculée à l'aide de la *Binary Cross-Entropy* (BCE) ou de l'erreur quadratique moyenne (*Mean Square Error* (MSE)) en fonction de la nature des données. La loss de reconstruction pousse le VAE à apprendre à reproduire les données d'entrée le plus fidèlement possible. Ici la BCE est choisie car elle comporte un avantage fort par rapport à la MSE : généralement, la fonction d'activation de la dernière couche est une sigmoïde, ce qui peut entraîner une saturation de la loss due au plateau de cette dernière. Cette saturation empêche les algorithmes d'apprentissage basés sur les gradients de progresser. Afin d'éviter cela, il est bon d'avoir un logarithme dans la fonction loss pour annuler l'exponentielle de la sigmoïde. C'est pourquoi la binary cross-entropy est préférée, utilisant un logarithme à la différence de la MSE.

$$\text{BCE}(\mathbf{x}, \hat{\mathbf{x}}) = -\frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (x_{i,j} \log(\hat{x}_{i,j}) + (1 - x_{i,j}) \log(1 - \hat{x}_{i,j})) \quad (2.5)$$

Dans cette équation, $x_{i,j}$ représente la valeur du pixel à la position (i, j) dans l'image originale, et $\hat{x}_{i,j}$ représente la valeur prédite par le modèle pour le pixel à la position (i, j) . H est la hauteur de l'image et W sa largeur, donc $H \times W$ est le nombre total de pixels dans l'image.

— **Partie loss de regularisation**

La seconde partie est appelée divergence KL et cherche à minimiser la distance entre les distributions latentes apprises notée q et une distribution, généralement normale $\mathcal{N}(0, 1)$, notée p . KL est la divergence de Kullback-Leibler, mesurant la distance entre les distributions $q(z|x)$ et $p(z)$:

$$\text{KL} = -\frac{1}{2} \sum_{i=1}^L (1 + \log \sigma_i^2 - \mu_i^2 - \sigma_i^2) \quad (2.6)$$

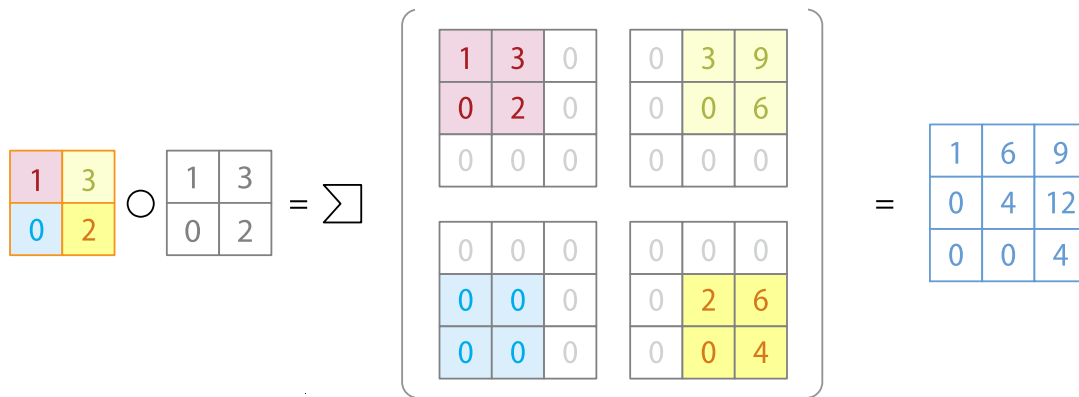
où μ_i et σ_i^2 sont respectivement la moyenne et la variance de la distribution de la représentation latente $q(z|x)$, et L est la dimension de la représentation latente. Ainsi, la loss KL mesure la divergence entre la distribution de la représentation latente apprise $q(z|x)$ et la distribution normale $p(z)$, contribuant ainsi à la régularisation du VAE en contraignant la représentation latente à suivre une distribution normale.

Enfin, la loss totale $\mathcal{L}_{\text{VAE}}$ se compose d'une somme pondérée de la loss de reconstruction et de la loss de regularisation :

$$\mathcal{L}_{\text{VAE}} = k_1 \times \text{BCE}(x, \hat{x}) + k_2 \times \text{KL} \quad (2.7)$$

où k_1 et k_2 sont des hyperparamètres de pondération, qui contrôlent l'importance relative de la perte de reconstruction et de la divergence KL dans l'entraînement du VAE. Cette formulation de la perte totale permet de guider l'apprentissage du VAE en encourageant à la fois une reconstruction précise des données d'entrée et une distribution de représentation latente conforme à la distribution a priori choisie.

Enfin, dans le but d'étudier le détail de l'architecture d'un VAE, nous allons rapidement évoquer les couches convolutives transposées. Ces couches, à l'inverse des couches de convolution permettent d'étendre l'image. Cette opération permet d'augmenter la dimension spatiale d'une image : supposons un tenseur d'entrée de dimensions $n_h \times n_w$ et un noyau de dimensions $k_h \times k_w$. En faisant glisser le noyau avec un pas de 1 pour n_h fois sur chaque ligne et n_w fois sur chaque colonne, cela produit une image intermédiaire. Chaque image intermédiaire est alors un tenseur de dimensions $(n_h \times k_h + 1) \times (n_w \times k_w + 1)$.

FIGURE 2.10. – Exemple convolution transposée. Extrait de [FIDLE](#).

où $\odot$ représente le symbole de convolution transposée, et $\sum$ représente la somme entre les matrices entre parenthèses. Cette convolution transposée permet à partir des valeurs tirée de l'espace latent d'augmenter la dimension des cartes de caractéristiques afin de générer une image en sortie.

Application

Le variational auto-encoder utilisé dans les travaux est un auto-encoder convolutif : constitué de couches convolutives il permet une meilleure appréciation des caractéristiques des images en entrée.

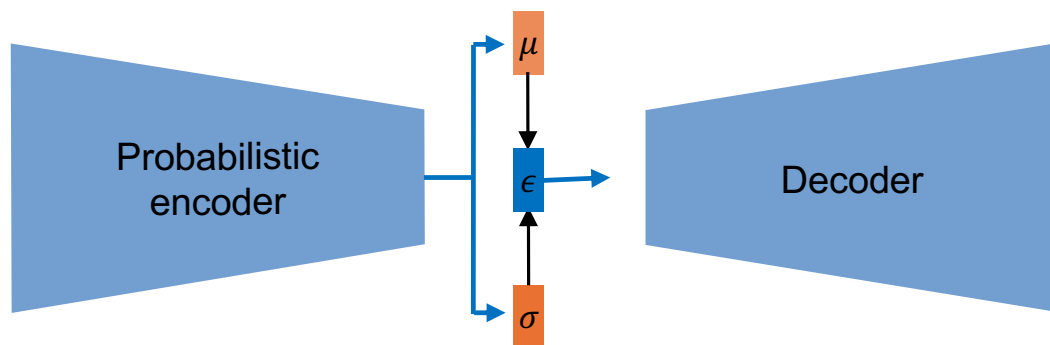


FIGURE 2.11. – Principe de l'architecture d'un VAE.

Layer	Type	Parameters
enc1	Conv2d	(1, 32, (3, 3), (1, 1), (1, 1))
enc2	Conv2d	(32, 64, (3, 3), (2, 2), (1, 1))
enc3	Conv2d	(64, 64, (3, 3), (2, 2), (1, 1))
enc4	Conv2d	(64, 64, (3, 3), (1, 1), (1, 1))
linearenc	Linear	(16384, 16)
mu_layer	Linear	(16, 8)
log_var_layer	Linear	(16, 8)
lineardec	Linear	(8, 16384)
dec1	ConvTranspose2d	(64, 64, (3, 3), (1, 1), (1, 1))cha
dec2	ConvTranspose2d	(64, 64, (3, 3), (2, 2), (1, 1), (1, 1))
dec3	ConvTranspose2d	(64, 32, (3, 3), (2, 2), (1, 1), (1, 1))
dec4	ConvTranspose2d	(32, 1, (3, 3), (1, 1), (1, 1))

TABLEAU 2.3. – Caractéristiques des couches du VAE

L'encodeur apprend donc la moyenne μ et l'écart type σ des caractéristiques qu'il extrait des images. La taille des vecteurs μ et σ correspond à la taille de l'espace latent. Lors de l'entraînement du réseau, un tirage aléatoire ϵ selon les loi normale $\mathcal{N}(\mu, \sigma)$ de l'espace latent est effectué, puis cette reparamétrisation ϵ est utilisée en entrée du décodeur afin de reconstruire une image à partir de ce tirage.

2.4. Méthodologie pour le phasage

2.4.1. Simulation de données

Dans notre application, les données réelles étant manquante, on se doit de simuler les données d'entrée des réseaux. La physique de la formation d'image nous donne une représentation très fiable des PSFs et image d'objet étendu attendus et permet donc de générer des jeu de données cohérents. Les caractéristiques physique des données simulées ont été présentée partie 1.4.2. Une partie concernant la génération d'objet étendu sera développée partie 5.2.1.

2.4.2. Choix de la loss

Dans le cadre de la régression de paramètres à partir de phénomènes physiques, il est idéal de sélectionner des fonctions de perte personnalisées appropriées pour garantir une adéquation optimale aux diverses exigences des problèmes rencontrés. Par exemple, afin de répondre au mieux à la minimisation de l'erreur de front d'onde sur la pupille, la loss utilisée pour le point source est directement le WFE. Pour le cas de l'objet étendu, une partie de la loss fait intervenir le modèle physique de formation d'image durant l'apprentissage.

2.4.3. Hyperparametres

Une fois les ensembles de données et les fonctions de perte déterminées, l'étape suivante consiste à choisir les hyperparamètres. Comme évoqué précédemment, les principaux paramètres modifiables pour améliorer le processus d'entraînement sont la taille des batchs et le taux d'apprentissage. Contrairement à la taille du batch, qui reste constante tout au long de l'entraînement, le taux d'apprentissage peut être adapté en fonction du nombre d'époques écoulées. Ainsi, un mécanisme de programmation dynamique du taux d'apprentissage, également appelé "scheduler", est mis en œuvre pour ajuster progressivement le taux d'apprentissage au fil des époques. Les détails de ces schedulers seront exposés dans les sections relatives aux résultats respectifs.

2.4.4. Le phasage par Deep Learning

L'apprentissage se base sur une stratégie supervisée à l'aide de données simulées par les modèles physiques de formation d'image. Dans les deux prochains chapitres, nous aborderons une méthode directe d'apprentissage qui consiste en l'inférence directe des paramètres de la phase à partir de l'image de la PSF. Une nouvelle méthodologie sera présentée au chapitre [5](#).

2. Methodologie – 2.4. Méthodologie pour le phasage

Dans ce chapitre, nous avons exploré en détail les concepts fondamentaux des réseaux de neurones, allant des multilayer perceptrons aux architectures plus avancées comme VGG-Net, ResNet et EfficientNetB0. Nous avons également examiné les caractéristiques des couches de convolution, des fonctions d'activation mais aussi les optimiseurs, ainsi que la manière dont ces éléments interagissent dans le processus d'apprentissage. Ces connaissances seront mises en œuvre dans le prochain chapitre, où nous appliquerons ces architectures à notre étude sur le phasage. Nous évaluerons les performances de ces modèles dans leur contexte respectif et concluerons quand à la pertinence de l'utilisation des réseaux de neurones.

3. Étude d'une source ponctuelle

Dans ce chapitre, j'étudie le comportement des méthodes d'apprentissages pour le phasage en plan focal dans le cas où l'objet observé est un point source. Dans un premier temps, je porte mon attention sur des cas simplifiés, c'est à dire où on estime un très petit nombre de coefficients dans un problème simple à peu de segment : j'analyse la réponse des réseaux de neurones à l'estimation de ces coefficients. Dans un second temps j'étudie l'impact de l'asymétrie de pupille, mentionnée partie 1.2.2. Les limitations de l'identification du piston (partie 1.4.2) seront ensuite présentées. Je détaille enfin les résultats obtenus sur une pupille pleine composée de 4 segments : les résultats seront exposés dans un cas à fort flux, bruité, de déploiement, et en boucle fermée. Les méthodes de Deep Learning seront comparées à l'état de l'art. Enfin une ouverture sur le phasage du JWST et sur la mesure du mode pétale de l'ELT sera présentée. Dans l'ensemble de ce chapitre, les données utilisées pour l'entraînement et le test des réseaux de neurones sont simulées à partir d'un simulateur de formation d'image que j'ai développé en Python.

Sommaire

3.1. Étude d'une pupille simplifiée	98
3.1.1. 1 Dimension : Tip	98
3.1.2. 2 Dimensions - 1 Pupille	101
3.1.3. 1 Dimension - 2 Pupilles	103
3.1.4. 2 Dimensions - 2 Pupilles	106
3.2. Étude sur pupille complète	110
3.2.1. Optimisation de la surface d'asymétrie	111
3.2.2. Résultats	114
3.2.3. Choix de l'architecture	115
3.3. Limitation de l'estimation du piston	116
3.4. Propagation de bruit	119
3.4.1. Performance à SNR constant	119
3.4.2. Robustesse au SNR	120
3.5. Influence des aberrations de hauts ordres	121
3.6. Comparaison aux méthodes d'AFOPF applicables	123
3.6.1. Diversité de Phase	123
3.6.2. Image Sharpening	124
3.6.3. Comparaison et temps d'estimation	125
3.7. Scenario de phasage complet	128

3. Étude d'une source ponctuelle

3.8. Application à de plus grands télescopes	131
3.8.1. Réseaux de neurones	132
3.8.2. Résultats	132
3.9. Application au petaling de l'ELT (sans turbulence)	135
3.9.1. Formation d'image	135
3.9.2. Réseau de neurones	137
3.9.3. Résultats	137

3.1. Étude d'une pupille simplifiée

Dans cette partie, on s'intéresse à un cas simplifié de notre problème (estimation d'un unique coefficient sur un unique segment). Pour étudier ce cas, nous allons utiliser une simple architecture de réseau de neurones afin de réaliser l'estimation des coefficients de la phase incidente à partir de PSFs à une ou deux pupilles, comprenant 1 ou deux coefficients chacune. Cette petite dimensionnalité des réseaux et des données d'entrée permet d'extraire et d'interpréter les paramètres internes appris par le réseau et de justifier le choix d'architectures plus complexes. J'ai développé, entraîné et testé l'ensemble des architectures présentées dans cette partie à l'aide de la librairie Pytorch.

3.1.1. 1 Dimension : Tip

On propose donc d'étudier la capacité d'un réseau de neurone à estimer un unique coefficient de tip sur une pupille carrée. L'effet de ce tilt ne déforme pas la PSF et la décale verticalement sur le détecteur, dans un sens ou dans l'autre. Le problème est donc linéaire. Dans le cas d'une résolution d'un problème linéaire, les couches de max-pooling, fonction d'activation et autres transformations non-linéaires ne sont pas nécessaires dans ce cas précis d'application. On propose donc une architecture composée uniquement d'une couche entièrement connectée, sans couche d'activation. Ce faisant, l'image en entrée du réseau est linéarisée et chaque poids et biais correspond à 1 pixel de l'image indépendamment.

3. Étude d'une source ponctuelle – 3.1. Étude d'une pupille simplifiée

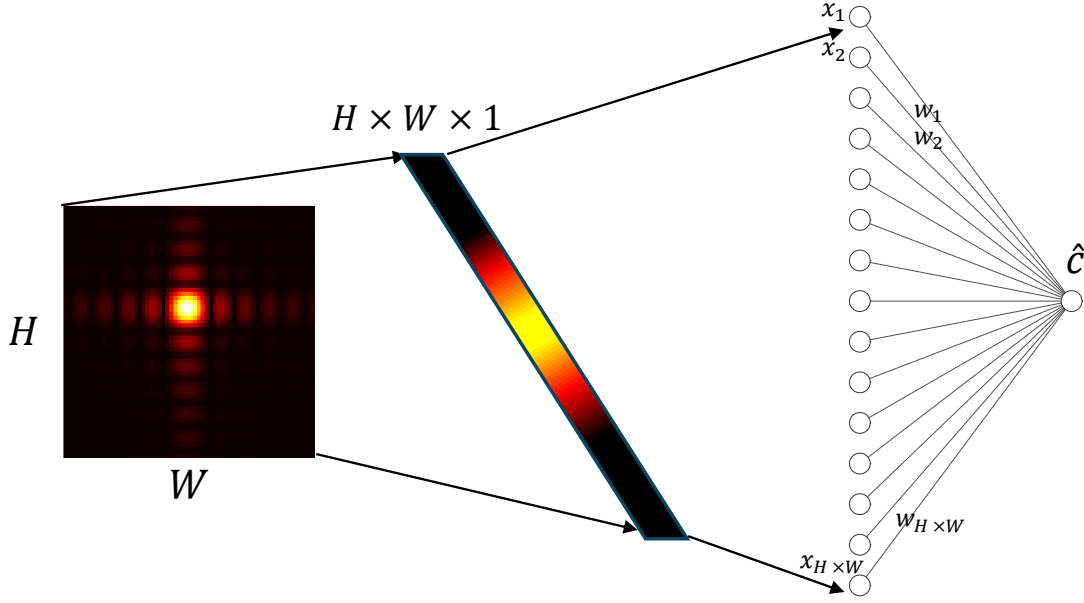


FIGURE 3.1. – Architecture d'un réseau de neurones entièrement connectée à une seule couche d'entrée entièrement connectée à un seul neurone de sortie sans couche cachée.

On note alors la sortie du réseau, correspondant au(x) coefficient(s) estimé(s) $\hat{c}$ telle que

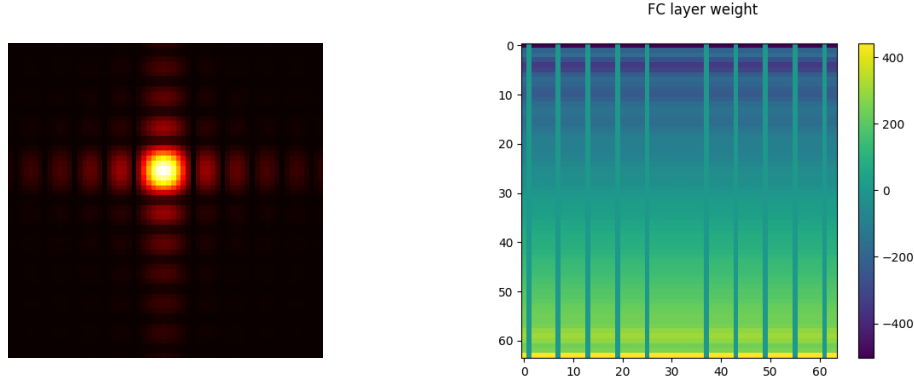
$$\hat{c} = \left(\sum_{n=0}^{H \times W} w_n \times x_n \right) + b \quad (3.1)$$

Où w_n représente le poids du neurones lié au pixel x_n , $H \times W$ correspond au nombre total de pixel de l'image, enfin, b le biais. Dans le cas présenté, $\hat{c}$ représente le tip estimé. En affiliant un nombre de neurones correspondant à la dimension de l'image d'entrée on peut alors reconstituer la carte des poids de la couche entièrement connectée comme une image ayant les mêmes dimensions que celle de l'image de la PSF en entrée. Le réseau agit alors comme la somme de la multiplication pixel à pixel de leur intensité et les poids correspondants de la couche d'entrée, à un biais près. On peut alors reformuler l'équation 3.1 telle que

$$\hat{c} = \left(\sum_{i=0}^H \sum_{j=0}^W w_{i,j} \times x_{i,j} \right) + b \quad (3.2)$$

On entraîne alors ce réseaux de neurones à partir d'un jeu de données de 1000 images et de leurs labels correspondants sur 30 époques utilisant une loss RMSE. On analyse ensuite les paramètres internes du réseaux appris.

3. Étude d'une source ponctuelle – 3.1. Étude d'une pupille simplifiée



(a) Exemple de PSF en entrée du réseau en échelle racine. Le décalage de la PSF par rapport au centre correspond au tilt introduit. (b) Carte des poids de la couche entièrement connectée. Les poids sont sans unité.

FIGURE 3.2. – Exemple de PSF en entrée et carte de poids

On peut alors observer les poids de la couche entièrement connectée en figure 3.2. Le biais étant négligeable face aux amplitudes des poids (ici, $b = 1.21$), on décide de l'ignorer ici. On remarque dans un premier temps une rampe formée par les poids des pixels selon Y variant dans un intervalle $[-400, 400]$ et constante selon X. C'est cette fonction rampe qui a été apprise par l'apprentissage du réseau en minimisant sa fonction de coût. On remarque aussi des franges verticales sur la carte des poids du réseau dont les poids sont strictement nuls. Ces franges, dues au formalisme matriciel sont localisées au même endroit où les pixels de la PSF s'annulent. Ces annulations, visibles uniquement en étalant l'échelle de la PSF par la fonction racine, n'ont donc aucune influence sur le calcul du tilt. L'opération en sortie du réseau, dictée par l'équation 3.2 peut être approximée :

$$\hat{c} = \sum_{i=0}^H \sum_{j=0}^W w_{i,j} \times x_{i,j} \quad (3.3)$$

or d'après la figure 3.2, en ignorant les franges nulles, on a $\forall (i, j) \in \mathbb{N}^2 w_{i,j} \approx z \times i$, avec z une constante multiplicative. On peut alors écrire

$$\hat{c} = z \sum_{i=0}^H i \sum_{j=0}^W x_{i,j} \quad (3.4)$$

On pose $L_i = \sum_{j=0}^W x_{i,j}$ la somme par ligne. La sortie du réseau devient alors

$$\hat{c} = z \sum_{i=0}^H i L_i \quad (3.5)$$

L'opération effectuée par le réseau est alors semblable à l'opération usuelle pour le

3. Étude d'une source ponctuelle – 3.1. Étude d'une pupille simplifiée

calcul d'un déplacement de position, le centre de gravité discret :

$$CoG = \sum_{i=0}^H i f(i) \quad (3.6)$$

Ou $f(i)$ est la valeur de la fonction dont on cherche le centre de gravité, évalué au pixel i . Évidemment, ce cas extrême n'est pas là pour apporter une alternative à la mesure du centre de gravité. Ce cas d'étude est étudié afin de montrer le fonctionnement interne d'un réseau, et que finalement l'apprentissage d'un réseau à une couche entièrement connectée mène à des méthodes connues.

3.1.2. 2 Dimensions - 1 Pupille

Afin d'étudier une complexité supérieure, on augmente le problème d'une inconnue. Cette fois le réseau doit identifier tip et tilt pour une PSF qui se déplace dans le champ. L'architecture est une nouvelle fois composée d'une seule couche d'entrée entièrement connectée à 2 neurones de sortie :

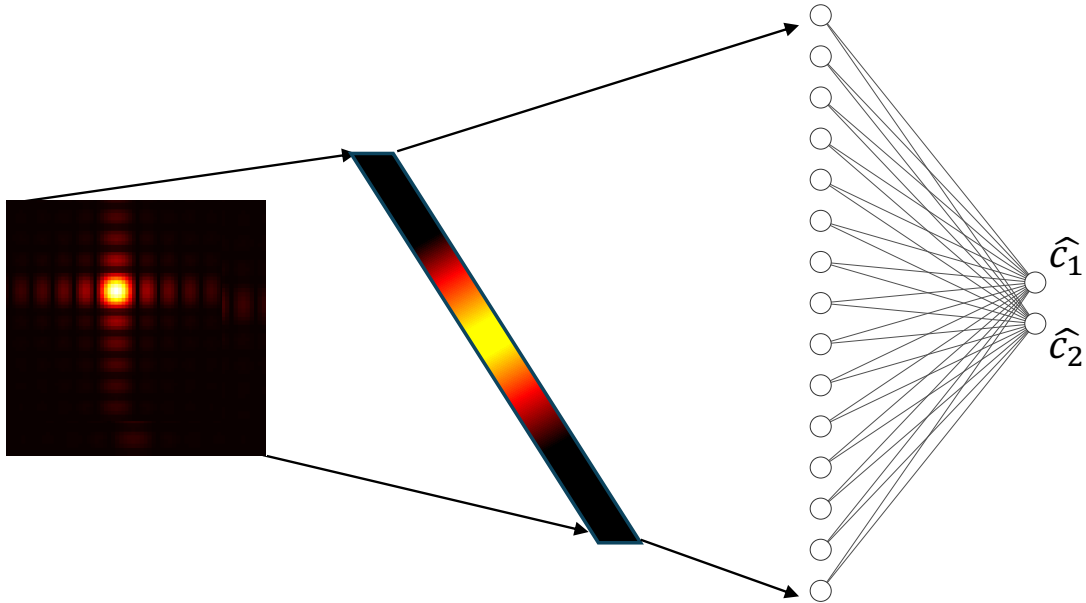


FIGURE 3.3. – Architecture d'un réseau à une couche entièrement connectée et deux neurones de sortie.

Les neurones sont connectés à la première et à la seconde sortie. Cela signifie qu'ils comportent maintenant 2 cartes de poids et 2 biais distincts. Les couches sorties

3. Étude d'une source ponctuelle – 3.1. Étude d'une pupille simplifiée

retournent alors :

$$\begin{aligned}\hat{c}_1 &= \left(\sum_{i=0}^H \sum_{j=0}^W w_{i,j}^1 \times x_{i,j} \right) + b^1 \\ \hat{c}_2 &= \left(\sum_{i=0}^H \sum_{j=0}^W w_{i,j}^2 \times x_{i,j} \right) + b^2\end{aligned}\tag{3.7}$$

Où $\hat{c}_1$ et $\hat{c}_2$ représentent les deux coefficients estimés (ici tip et tilt). On propose une nouvelle fois de s'intéresser aux poids de la couche entièrement connectée une fois le réseau entraîné avec l'aide d'une loss RMSE.

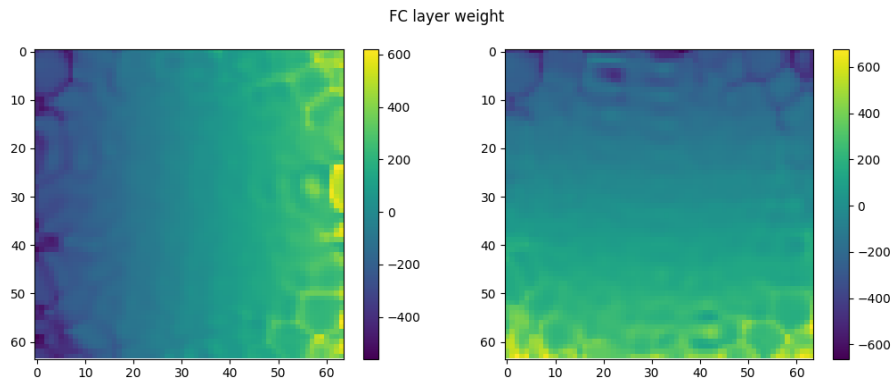


FIGURE 3.4. – Carte des poids de la couche entièrement connectée.

Les cartes de pondération inférées par l'apprentissage du réseau de neurone ressemblent aux cartes de calcul du centre de gravité, cette fois-ci dans les deux directions. Une carte pour l'horizontale, une carte pour la verticale. Cependant des effets apparaissent aux bords de ces cartes dû au jeu de données d'entraînement vraisemblablement non suffisant pour avoir des PSFs qui couvrent tout le plan focal.

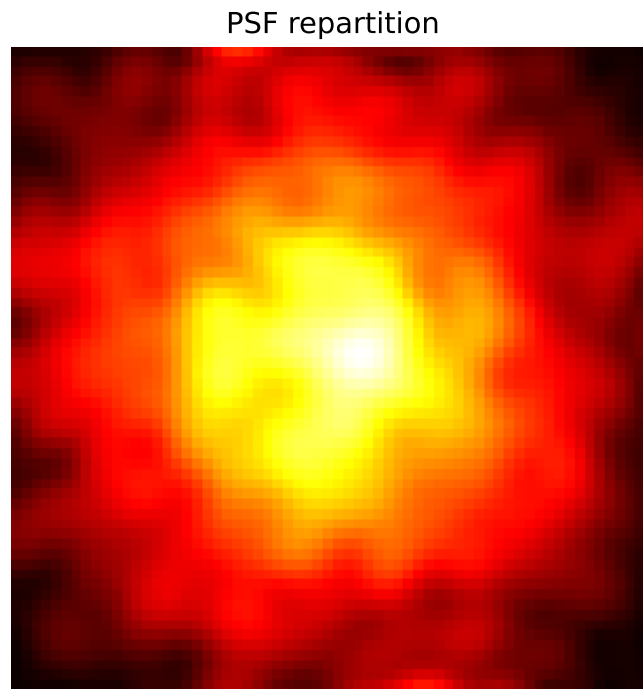


FIGURE 3.5. – Répartitions des 1000 PSFs d'entraînement.

En effet, la figure 3.5 montre que la plupart des PSFs d'entraînement sont réparties au centre de l'image, et se repartissent dans le plan focal en atteignant très peu les bords de l'image. C'est pourquoi les effets de bord peuvent apparaître. Néanmoins, avec un jeu de données dont les PSFs ne couvrent pas l'intégralité du plan focal, le réseau apprend correctement ces poids, même dans les zones très peu vues lors de l'entraînement, montrant donc une bonne généralisation de l'apprentissage légèrement plus loin que les données d'entraînement. Ces motifs changent lorsque l'on effectue un nouvel entraînement avec un autre tirages de PSFs.

Dans le cas présenté ici, la relation entre intensité des tip-tilts et intensité des pixels est presque linéaire étant donné que la déviation de la PSFs est proportionnelle à l'amplitude de l'inclinaison du miroir. Les deux cas de figures sont donc relativement simple à apprendre pour un réseau de neurones, même peu complexe.

3.1.3. 1 Dimension - 2 Pupilles

On propose de résoudre un nouveau problème : une pupille à deux segments, et de mesurer un coefficient de tilt par segment. On déplace alors chacune des deux sous-PSFs suivant une seule direction, verticale, correspondant à un tip différent

3. Étude d'une source ponctuelle – 3.1. Étude d'une pupille simplifiée

sur les deux sous-pupilles. Les sous-PSFs peuvent alors être totalement séparées, ou superposées. Un exemple est donné figure 3.6.

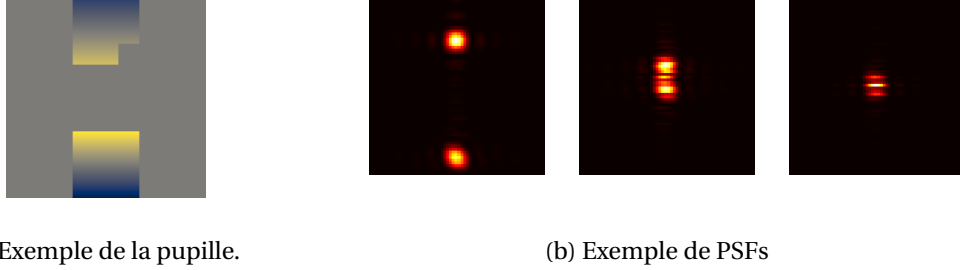


FIGURE 3.6. – Exemple de PSF en entrée et carte de poids

On peut remarquer que la sous-pupille du haut est tronquée sur le coin en bas à droite permettant d'identifier aisément quelle PSF appartient à quelle pupille (puis dans des cas plus complexe, à remonter à la partie paire de la phase, comme mentionné partie 1.2.2). Étant donné que la distribution spatiale de la sous pupille tronquée est donc moins étendue, alors le contenu spectral contenu dans la sous-PSF associée est plus étendu. Une nouvelle fois, une architecture entièrement connectée sans couche d'activation est utilisée. Cependant, la loss est cette fois la contribution linéaire des tilts, étant chacun manifesté sur une pupille différente. Une fois l'apprentissage effectuée, on observe que le réseau exploite de cette caractéristique spatiale pour identifier les tilts :

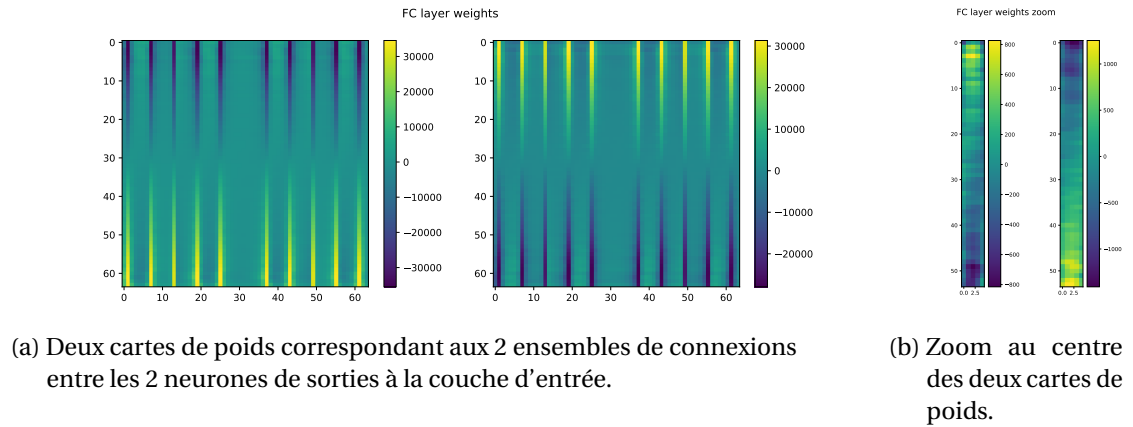


FIGURE 3.7. – Carte des poids de la couche entièrement connectée et zoom sur la bande centrale.

On remarque une nouvelle fois ces franges verticales. Cette fois ci, plutôt que d'annuler le produit par le pixel, elles sont de très hautes intensité pour mettre en évidence la différence entre les deux sous-PSFs. Étant donné que la pupille globale est plus

3. Étude d'une source ponctuelle – 3.1. Étude d'une pupille simplifiée

grande que les sous-pupilles, la PSF résultante de l'interférence des deux sous-PSFs est moins étendue, ces bandes de hautes intensité permettent alors de repérer ou non l'interférence des sous-PSFs. En analysant l'intensité entre deux franges de la carte de poids on retrouve une nouvelle fois la rampe rappelant le centre de gravité. Cette fois le centre de gravité est moins lisse que lors des cas précédents car il prend aussi en compte l'interférence entre les deux sous-PSFs. Les deux sous pupilles étant distribuée verticalement, les interférences ne se font que dans la direction verticale et laisse donc observer une largeur de la PSF inchangée mais une hauteur 3 fois plus petite. Ces lignes horizontales permettent d'estimer l'amplitude des aberrations et se retrouvent sur la carte de poids.

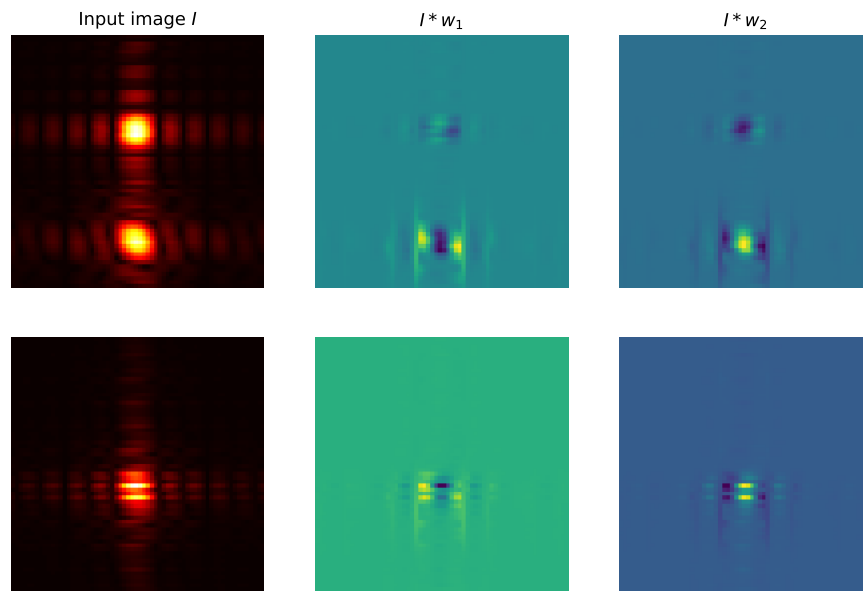


FIGURE 3.8. – Exemple de multiplication terme à terme de la carte de poids et de deux PSFs (échelle racine).

En figure 3.8, w_1 et w_2 sont les 2 vecteurs correspondant aux poids de chaque ensemble de connections de la couche d'entrée avec les 2 neurones de sortie. On remarque deux cas distincts :

- Les sous PSFs sont séparées : alors les franges verticales de hautes intensités permettent d'identifier quel PSFs appartient à quelle pupille et permet de déterminer l'amplitude des tips correspondant.
- Les sous PSFs interfèrent : alors des franges horizontales apparaissent au niveau de l'interférence des PSFs, ces interférences sont identifiées lorsque l'on zoom sur les cartes des poids w_1 et w_2 , et permettent d'extraire les tips des pupilles.

Dans ce cas, le problème inverse s'éloigne d'un régime linéaire, néanmoins, l'apprentissage du réseau montre une capacité à s'adapter aux variations des intensités du détecteur, et offre encore la possibilité d'interpréter l'apprentissage.

3.1.4. 2 Dimensions - 2 Pupilles

Maintenant intéressons nous à un cas plus complexe. On simule alors 2 pupilles, soumises à un tip et un tilt chacune. 4 coefficients doivent alors être identifiés à partir de la PSF seule et de l'architecture entièrement connectée. La loss quand à elle est composée de la somme linéaire entre les segments de la somme quadratique sur les modes. Cette fois-ci l'entraînement du réseau de neurones converge mais ne tend pas vers une valeur de front d'onde résiduelle suffisante. En effet le réseau apprend bien à minimiser sa fonction de coût, mais à un seuil trop élevé. Observons en détails comment se passe l'entraînement du réseau.

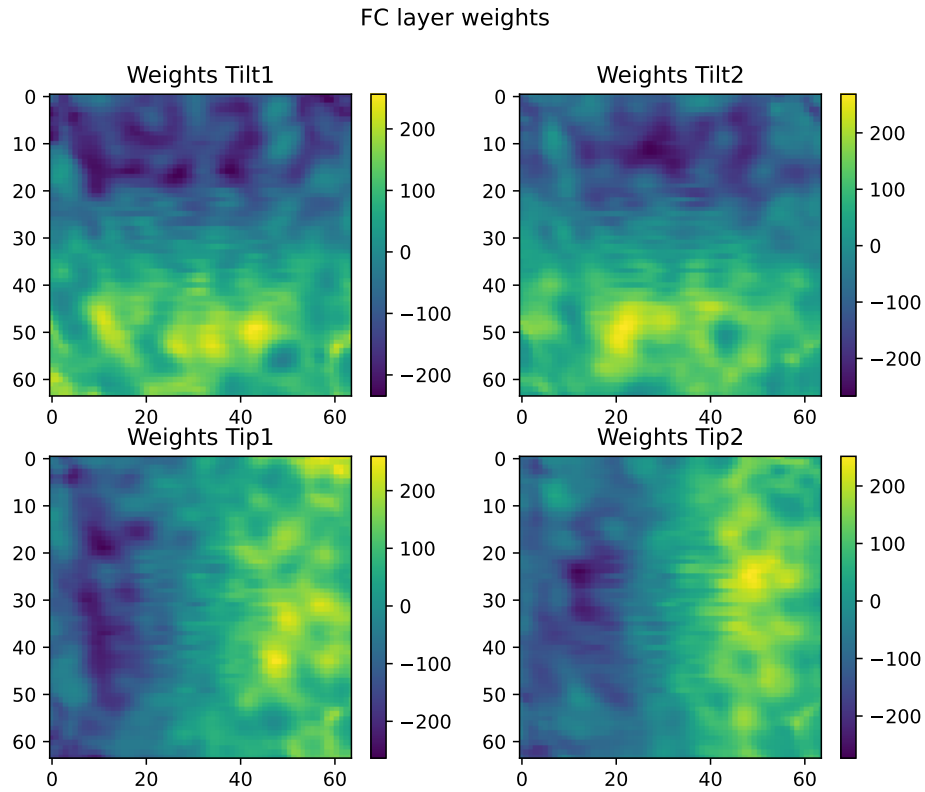


FIGURE 3.9. – Carte des poids de la couche entièrement connectée.

On remarque en figure 3.9, que les poids associés à l'identification des tips, et des tilts sont très similaires entre les pupilles. Le réseau apprend donc à identifier les paramètres mais sans les dissocier. Cependant on reconnaît la rampe caractéristique

3. Étude d'une source ponctuelle – 3.1. Étude d'une pupille simplifiée

de l'identification des tip-tilts. En reconstituant les PSFs issues des coefficients estimés par le réseau on retrouve une caractéristique significative :

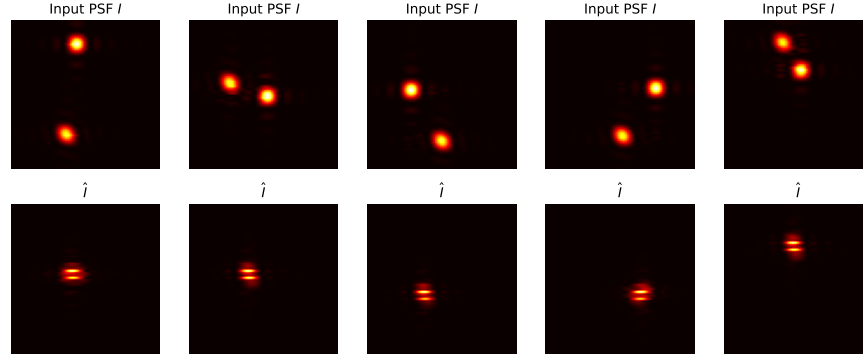


FIGURE 3.10. – Exemple de PSF en entrée du réseaux, I , et des PSFs estimées (reconstituées $\hat{I}$ à partir des coefficients estimés par le réseaux).

On voit ici que le réseau estime la PSF comme étant au centre des deux sous-PSFs d'entrée. Cela permet, dans une moindre mesure, de minimiser l'erreur d'estimation et de ne pas estimer des valeurs aberrantes. Comme énoncé en partie 2.1.5, l'entraînement des réseaux se base sur la minimisation d'une fonction de coût, ici le WFE. Cependant, cette fonction, dépendante des paramètres internes du réseau, peut trouver un minimum local important, voire même un minimum global qui ne correspond pas aux critères de performance attendus.

Étudions pourquoi le réseau a optimisé ses poids de façon à attribuer les mêmes valeurs entre tip et tilt. Estimer les mêmes valeurs entre tip et tilt correspond à un minimum local de la fonction de coût et bloque l'optimisation lors de l'apprentissage. Prenons un exemple de PSF où les sous-PSFs sont séparées. On vérifie alors l'erreur d'estimation pour un tirage aléatoire de coefficients et on la compare à l'erreur d'estimation pour ce même tirage aléatoire, avec une contrainte $tip_1 = tip_2$ et $tilt_1 = tilt_2$.

3. Étude d'une source ponctuelle – 3.1. Étude d'une pupille simplifiée

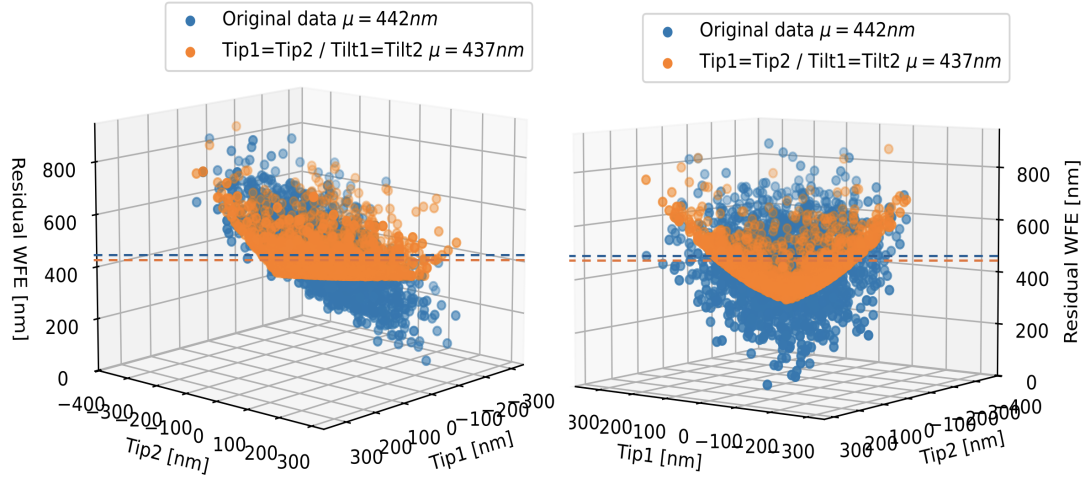


FIGURE 3.11. – Illustration de la minimisation de la moyenne du WFE par $tip_1 = tip_2$ et $tilt_1 = tilt_2$.

Égaliser les tips-tilts (points orange) figure 3.11 fait observer un plateau dans l'erreur de front d'onde car il limite l'estimation correcte dans le cas où les deux sous-PSFs ne sont pas superposées. Ce plateau observé comporte une moyenne d'erreur de front d'onde de $\mu_{condition} = 437 nm$ (courbe pointillée orange) contre $\mu_{original} = 442 nm$ (courbe pointillée bleue) pour les coefficients distribués sans contraintes. Le front d'onde est donc minimisé lorsque les coefficients sont contraints à $tip_1 = tip_2$ et $tilt_1 = tilt_2$. De plus, on peut aussi observer que l'erreur se minimise bien lorsque les deux PSFs superposées se trouvent sur la ligne entre les deux sous-PSFs du champ :

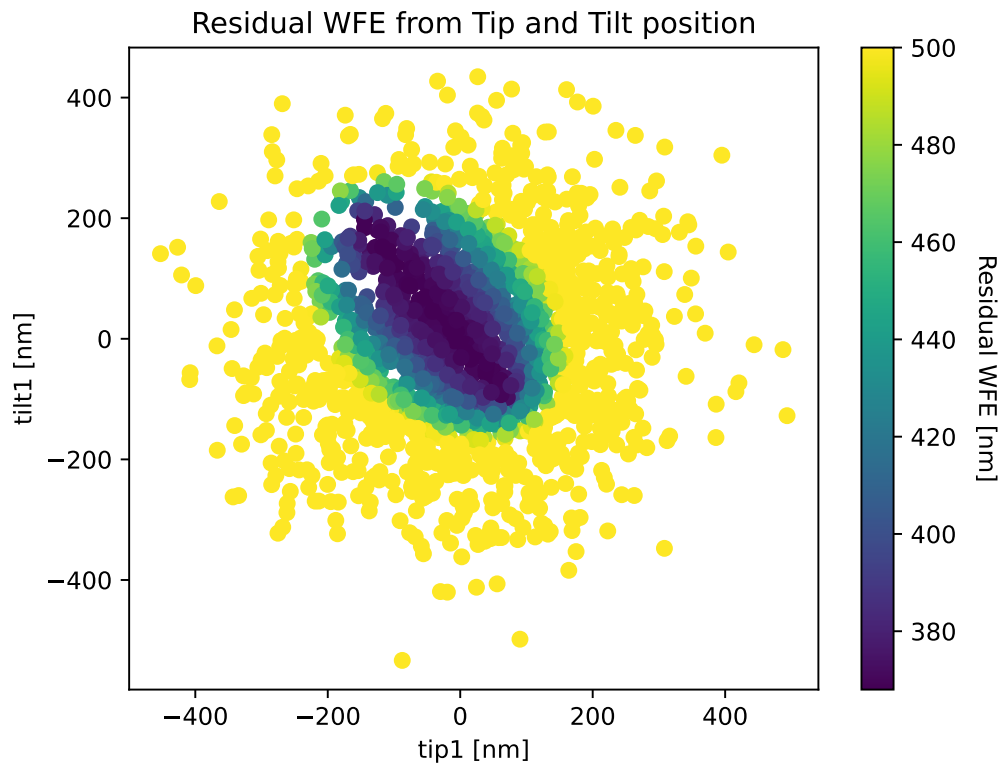


FIGURE 3.12. – Vue du dessus de l'erreur pour les tip-tilts égaux. Un seuillage est effectué à une erreur de 500 nm pour apprécier la dynamique des plus petites erreurs de front d'onde.

Cette ligne correspond aux valeurs comprises dans la zone où le front d'onde est le plus faible, figure 3.12. On minimise donc l'erreur initiale, donc la fonction de coût est optimisée, sans atteindre une valeur satisfaisante.

On remarque à travers cet exemple que la simple inférence par une seule couche entièrement connectée est trop légère pour construire une fonction de coût avec un minimum global plus précis dans un cas aussi complexe. Une information supplémentaire peut être donnée au réseau pour qu'il apprenne plus efficacement les 4 paramètres de la phase : l'information spatiale. C'est la tout l'intérêt des couches convolutives. On propose maintenant d'étudier un réseau composé cette fois de deux couches convolutives et d'une couche entièrement connectée.

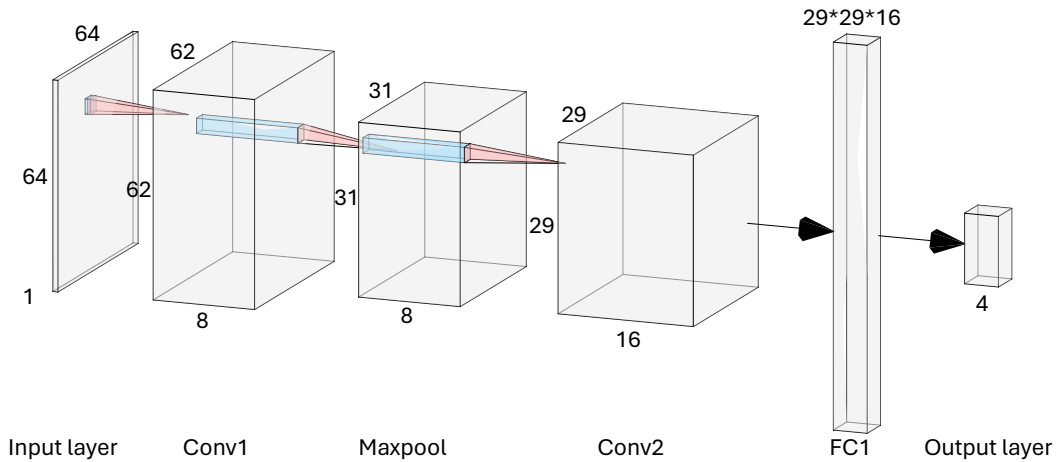


FIGURE 3.13. – Architecture pour le cas 4 coefficients composée de 2 couches convolutives (le 3eme bloc représente la couche de max-pooling uniquement) et une couche entièrement connectée.

Le lecteur pourra retrouver en annexe une partie sur l'analyse des filtres, poids et biais de cette architecture une fois entraînée (voir annexe A). Cette architecture permet d'estimer correctement les 4 paramètres de la phase à un front d'onde global autour de quelques nanometres.

On comprend ici que le saut de 2 coefficients à 4 implique une complexification conséquente de l'architecture. Le problème de phasage étudié dans ce manuscrit comporte 12 coefficients à identifier à partir de 4 segments indépendants, les architectures présentées seront de ce fait, plus complexes que celles abordées dans cette partie. Cette partie explore un cas d'école qui converge vers des résultats relativement interprétables. Les propriétés explorées ici auraient pu bénéficier aux travaux ultérieurs, mais l'aspect interprétable des réseaux de neurones se perd au profit de méthodes plus effiaces pour l'analyse de front d'onde.

Conclusion

La complexification des architectures nécessite une abstraction de l'interprétabilité au profit des performances souhaitées.

3.2. Étude sur pupille complète

On propose maintenant d'étudier le problème dans son intégralité : identifier piston-tip-tilt des 4 segments à partir de PSF directement. On travaille dans un premier temps avec des données brutes : des PSFs composées uniquement d'aberrations

de PTT. L'architecture de réseau de neurones proposée dans un premier temps est l'architecture semblable à un VGG-net telle que décrite partie 2.3, que j'ai développée sur mesure.

3.2.1. Optimisation de la surface d'asymétrie

En partie 1.2.2 nous avons vu que une des premières limitations que l'on retrouve est l'identification de la partie paire de la phase. Dans cette partie, nous allons montrer que les méthodes de Deep Learning requièrent aussi l'ajout de diversité aux images d'entrée pour l'AFOPE. Cette diversité peut être apportée, notamment, en brisant la centro-symétrie de la pupille. C'est cette surface d'asymétrie que nous proposons d'optimiser ici.

Décomposition de la formation d'image en partie paire et impaire

On propose d'explicitier la décomposition en partie paire et impaire dans notre étude. Toute décomposition de fonction en partie paire et impaire se fait selon la transformation

$$\begin{aligned} p(h(x)) &= \frac{1}{2}(h(x) + h(-x)) \\ i(h(x)) &= \frac{1}{2}(h(x) - h(-x)) \end{aligned} \quad (3.8)$$

Soit V la matrice de transformation de la fonction $h(x)$, et V' la matrice de transformation de la fonction $h(-x)$. On peut alors établir la matrice T de transformation entre V et V' tel que $V' = TV$. Dans ce cas, l'équation 3.8 devient

$$\begin{aligned} p(h(x)) &= \frac{1}{2}(1 + T)V \\ i(h(x)) &= \frac{1}{2}(1 - T)V \end{aligned} \quad (3.9)$$

Sachant que dans notre cas, le passage entre V et V' correspond à une rotation de 180 degrés de la carte de phase, les pistons réalisent alors une symétrie et les tips et tilts une anti-symétrie. On prend comme entrée un vecteur $\mathbf{x}$ telle que

$$\mathbf{x} = (P_1, P_2, P_3, P_4, Ti_1, Ti_2, Ti_3, Ti_4, Tl_1, Tl_2, Tl_3, Tl_4) \quad (3.10)$$

avec P_k , Ti_k et Tl_k les pistons, tips et tilts respectivement du segment k . Alors, en organisant les segments de 1 à 4 en commençant par le segment du haut puis dans le sens trigonométrique (voir figure 3.14).

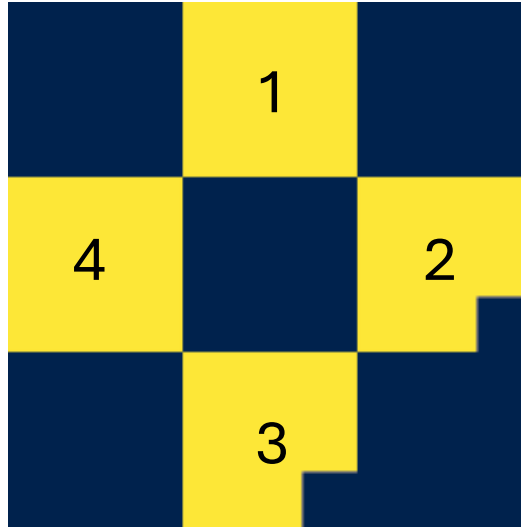


FIGURE 3.14. – Numérotation des segments de la pupille AZIMOV.

On obtient alors la matrice T telle que

$$T = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix} \quad (3.11)$$

Maintenant que l'on peut décomposer la phase en une partie paire et une partie impaire on peut alors analyser les résultats de l'estimation des réseaux de neurones. Une architecture VGG-net est alors développée, à partir des paramètres étudiés en partie 2.7. Afin d'entraîner le réseau de neurones à effectuer l'estimation de la phase sur pupille entière, la loss utilisée pour l'entraînement du réseau est le WFE vu en équation 1.22. On propose d'entraîner un premier réseau à partir de 10^5 PSFs et de leurs labels correspondants, sur 100 époques.

A la fin de l'entraînement, le réseau converge vers une valeur de la loss de validation autour de 43 nm , loin des 15 nm des spécifications. En décomposant la phase reconstruite à partir des estimations et calculant l'erreur de front d'onde sur la partie paire et la partie impaire on obtient pour 1000 PSFs de test :

3. Étude d'une source ponctuelle – 3.2. Étude sur pupille complète

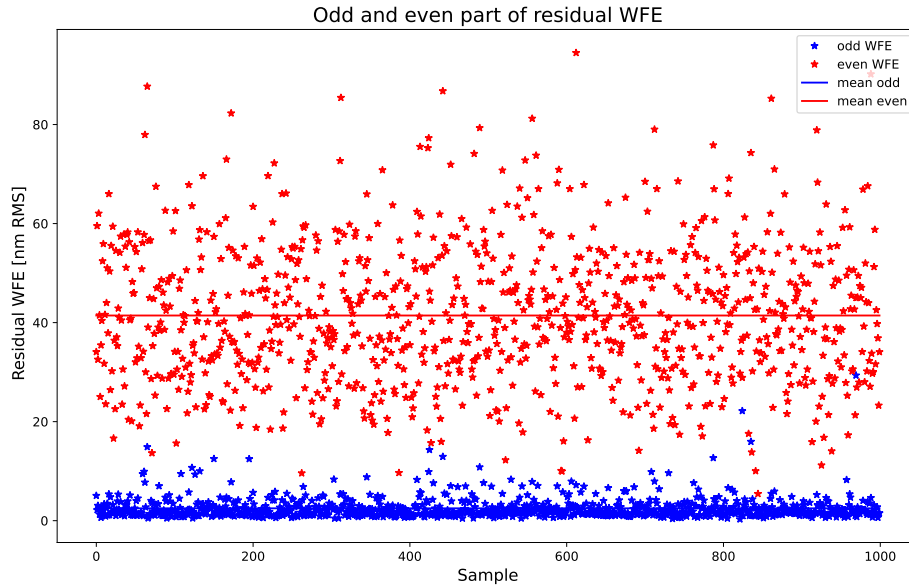


FIGURE 3.15. – Décomposition de l'erreur de la phase reconstruite en fonction de sa partie paire (rouge) et impaire (bleu).

Avec un front d'onde résiduel de 41.43 nm , la partie paire du front d'onde représente 94% de l'erreur totale. La partie impaire des résidus du front d'onde sont en moyenne de 2.49 nm pour un front d'onde résiduel total moyen de 43.92 nm . L'intérêt de briser la symétrie de la pupille est donc crucial pour estimer correctement le front d'onde.

Relation asymétrie - WFE

On propose donc de déterminer la surface à briser sur la pupille. Cette surface est choisie de façon à ce que deux carreaux soient coupés sur le bord bas-droite des segments 3 et 4 de la pupille. On étudie alors l'erreur d'estimation du front d'onde en fonction de la surface de symétrie brisée .

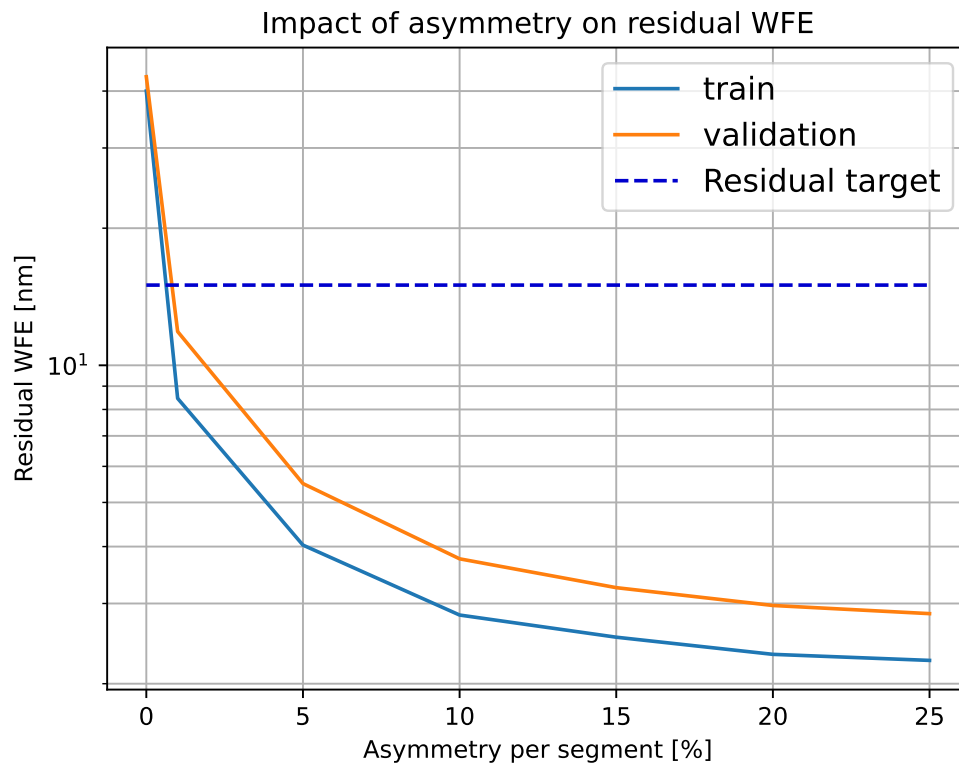


FIGURE 3.16. – Front d'onde résiduel en fonction de la surface d'asymétrie de la pupille pour la loss d'entraînement et la loss de validation.

La figure 3.16 présente les résultats de l'estimation des coefficients de la phase en fonction de la part d'asymétrie de la pupille. Le pourcentage représenté est le pourcentage par segment, la surface totale coupée relative par rapport à la pupille totale est donc cette valeur divisée par deux. On note une nette baisse de la tendance du WFE résiduel autour de 10% de surface coupée.

Conclusion

C'est cette valeur de 10% de la surface de 2 segments qui sera choisie dans la suite de notre étude afin d'observer une perte totale de 5% de la surface.

3.2.2. Résultats

Une fois le choix de la surface coupée sur la pupille, on propose d'étudier la qualité d'estimation des réseaux de neurones en fonction de leur architecture et de leur nombre de paramètres. Le choix de l'architecture du réseau pouvant être un paramètre déterminant due à la capacité calculatoire limitée à bord d'un nanosatellite,

un choix de l'architecture optimale est essentiel. On propose donc d'étudier les deux architectures présentées en partie 2.3, l'architecture Resnet (que j'ai développé sur mesure), et l'architecture VGG-net. Ces architectures sont entre autres dépendantes du nombre de filtres des couches convolutives N1,N2 et N3. C'est ce nombre qui va venir différencier les réseaux étudiés et donc leur taille. Pour rappel, le tableau § recense le nombre de paramètres et la taille des réseaux étudiés.

3.2.3. Choix de l'architecture

Maintenant que le choix des architectures est effectué, on propose donc d'entraîner ces différents réseaux sur 300 époques, avec 10^5 PSFs réparties en $9 \cdot 10^4$ pour l'entraînement, 10^4 pour la validation et 10^4 échantillons supplémentaires pour le jeu de donnée de test. Un batch de données de 32 PSFs est utilisé en plus d'un learning rate scheduler :

$$LR(e) = \begin{cases} 10^{-4} & \text{Si } e < 70 \\ 10^{-6} + (10^{-4} - 10^{-6}) \exp^{-(e-70) \cdot 0.0125} & \text{Sinon} \end{cases} \quad (3.12)$$

Où e représente les époques lors de l'entraînement. On répertorie alors les résultats en fonction de la taille du réseaux et de l'erreur d'estimation et on compare les performances des deux architectures énoncées. Les performances sont comparées sur les mêmes jeux de données.

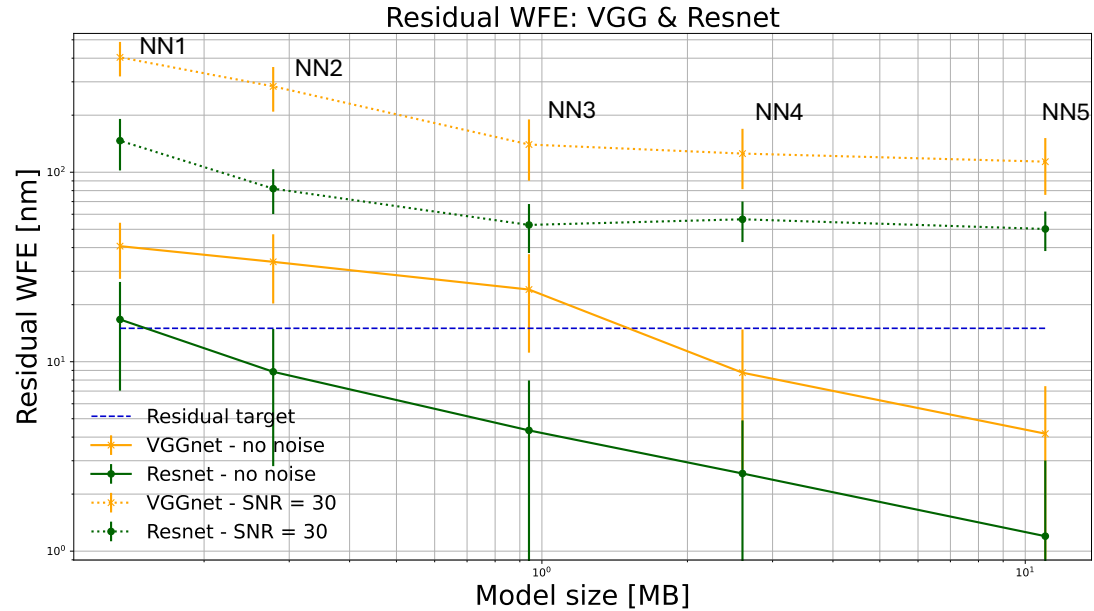


FIGURE 3.17. – Comparaison entre VGG-net (courbes orangées) et Resnet (courbes vertes) entraînés sur données brutes et testés sur données non bruitées (trait plein) et données bruitées (pointillés).

3. Étude d'une source ponctuelle – 3.3. Limitation de l'estimation du piston

On présente en figure 3.17 la qualité d'estimation en fonction de la taille du réseau en MB (§). Comme attendu, plus le réseau comporte de paramètres internes, meilleure est la qualité d'inférence. Cette tendance se retrouve pour les deux types d'architecture. Néanmoins, il faut veiller au sur-apprentissage pour les plus gros réseaux : l'estimation est certes meilleure, mais encore plus fidèle au modèle avec lequel les données ont été générées. En terme de qualité d'estimation, l'architecture Resnet est nettement plus performante que l'architecture VGG-net dans le cas de données bruitées ou non. La *skip connection* ajoutée au Resnet permet une meilleure optimisation des premières couches du réseau même dans le cas de réseau peu profond comme ici. De plus lors de l'entraînement des réseaux, l'architecture Resnet a montré une convergence bien plus rapide que pour le VGGnet qui présente un plateau après peu d'époques, pour converger ensuite vers la valeur finale.

Conclusion

Les performances présentées par les architectures VGG-Net et Resnet indiquent que le Resnet est bien plus performant que le VGG-Net. C'est donc le Resnet qui sera utilisé pour la suite de l'étude.

3.3. Limitation de l'estimation du piston

La première utilisation des Resnets sera pour l'identification et la mise en avant de l'erreur de piston mentionnée partie 1.4.2. Les manifestation du Piston et des tip-tilts sont très différentes en plan focal. L'un vient modifier les interférences entres sous-PSFs pendant que les deux autres les déplacent déplacent. La manifestation particulière du piston sur la PSF attire notre attention dans cette partie où nous proposons de vérifier par les méthodes de Deep Learning ces spécificités. Dans un premier temps on propose donc de générer plusieurs jeu de données, à tip et tilt nuls, mais en faisant varier les pistons suivant un tirage uniforme dans un intervalle donné. On entraîne ensuite un réseau de neurones par jeu de données et on compare alors les performances des différents réseaux.

3. Étude d'une source ponctuelle – 3.3. Limitation de l'estimation du piston

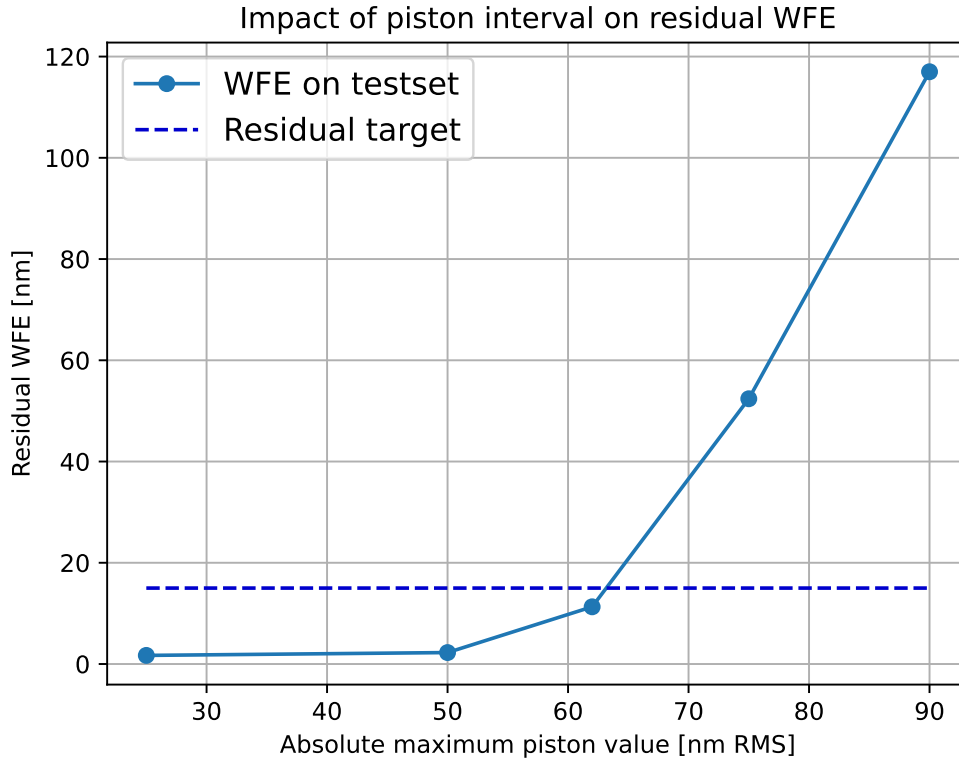


FIGURE 3.18. – Comparaison de la qualité d'estimation du piston pour des réseaux entraînés avec des jeux de données correspondant à différentes distributions uniformes de piston.

Pour rappel, l'équation 1.21 indique que la valeur du piston RMS est 4 fois plus petite que la valeur *Peak to Valley* (PV). Ce qui signifie que pour une valeur de piston de 50 nm RMS la valeur PV est de 200 nm PV soit $\frac{\lambda}{4}$. La figure 3.18 présente l'erreur d'estimation de différents réseaux en fonction de la distribution des pistons de son jeu de donnée d'entraînement. Dans le cas étudié ici uniquement, les pistons suivent une distribution uniforme dont les bornes sont représentées par les valeurs en abscisse sur le graphique. On retrouve une très bonne qualité d'estimation tant que l'on reste dans l'intervalle de piston inférieur à 50 nm RMS (soit un intervalle PV $[-\frac{\lambda}{4}, \frac{\lambda}{4}] \text{ nm}$). Pour de plus grandes valeurs de piston, l'estimation se dégrade rapidement, validant ainsi l'intervalle présenté équation 1.31. Étudions le réseau entraîné avec une distribution de piston suivant le tirage uniforme $\mathcal{U}(-75, 75) \text{ nm RMS}$, ayant un front d'onde résiduel de 52.4 nm .

3. Étude d'une source ponctuelle – 3.3. Limitation de l'estimation du piston

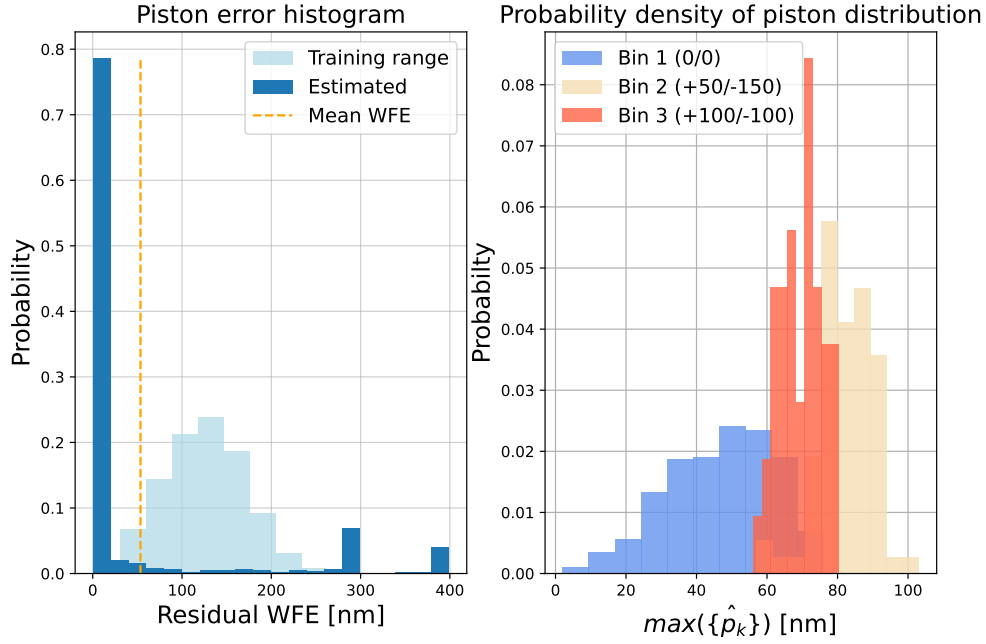


FIGURE 3.19. – Histogramme des erreurs de piston pour le réseau entraîné sur des PSFs dont les pistons varient entre $[-75, 75]$ nm.

Pour un jeu de données de 1000 PSFs, autour de 79% des échantillons de test sont correctement estimés. On peut remarquer 3 tendances dans les erreurs d'estimation. La première barre qui correspond un front d'onde correctement estimé, une seconde barre autour de 300 nm de WFE et une troisième barre autour de 400 nm de WFE. Ces trois barres sont représentées respectivement en bleu, rouge et beige sur le graphique de droite. Ce graphique représente la valeur absolue de la plus haute valeur des 4 pistons identifiés lors de l'inférence. Il permet donc d'identifier les sauts de lambda qui ont été estimés. On remarquera alors que la seconde barre amenant un WFE d'environ 300 nm correspond aux sauts de λ indiqués par l'équation 1.32, alors que la troisième barre, identifiée en orange sur le graphique de droite, correspond aux erreurs décrites équation 1.30. Ces erreurs surviennent en raison de la largeur de l'intervalle de 150 nm RMS, autorisant aux identifications des pistons un écart δ additionnel de 150 nm RMS. C'est ce qui explique la tendance de la courbe figure 3.18, qui augmente progressivement à partir de 50 nm RMS, puis augmente alors plus vite arrivé à 75 nm RMS.

Conclusion

Les erreurs dues au saut de λ sont une limite à la formation d'image. L'estimation de la phase par des réseaux de neurones et donc limitée, elle aussi, par la longueur d'onde.

3.4. Propagation de bruit

3.4.1. Performance à SNR constant

On propose maintenant d'étudier la propagation de bruit à travers les réseaux de neurones. 2 scénarios distincts sont étudiés :

- **Scenario 1**, l'erreur de modèle : Ce scénario consiste à entraîner un réseau de neurone sur des données non bruitées, puis à le tester sur des données bruitées à différents SNRs en mesurant la qualité de l'estimation.
- **Scenario 2**, limite des réseaux : les réseaux sont entraînés avec des données bruitées à SNR fixe et ensuite testé sur des données au SNR correspondant.

On peut alors tester les 5 différentes tailles de réseau.

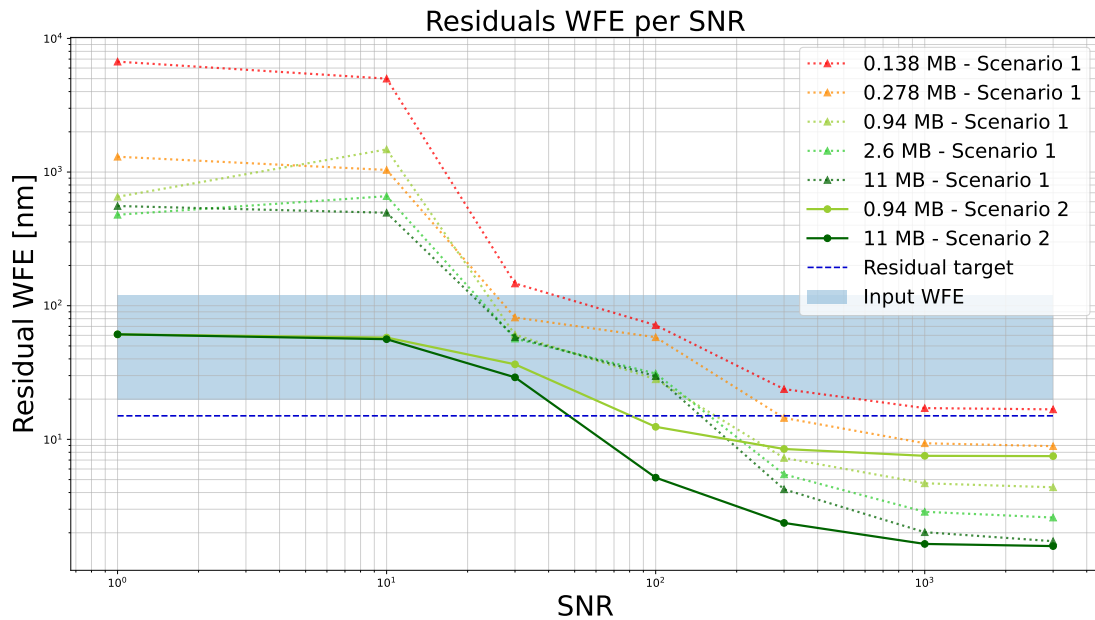


FIGURE 3.20. – Comparaison de la qualité d'estimation des réseaux entraînés sans bruit (pointillé) ou entraînés avec des données bruitées (trait plein) en fonction du SNR.

Nous avons réalisé une comparaison de performance entre les deux scénarios basée sur l'erreur résiduelle de front d'onde après correction. La figure 3.20 illustre cette erreur résiduelle de front d'onde en fonction du SNR. Les lignes en pointillés représentent les modèles entraînés sans bruit, tandis que les lignes pleines représentent les modèles entraînés avec des PSFs bruitées (chaque point de la courbe correspond donc à un modèle entraîné à un SNR spécifique). Comme attendu, la performance des modèles s'améliore à mesure que leur taille augmente. La limite de diffraction, définie par $WFE < Residual\ target$, n'est pas atteinte par le plus petit modèle (ligne rouge en pointillés) même à des SNR élevés, à l'image du réseau NN1 entraîné sans bruit figure

3.17. Pour un faible SNR, où le bruit du détecteur prédomine sur le bruit de photon, le *scénario 1* génère une erreur d'estimation plus grande que le WFE initial : cela entraînerait une augmentation du WFE s'il était utilisé en boucle fermée. En revanche, l'estimation des réseaux du *scénario 2* reste dans la moyenne de la distribution de WFE d'entrée : pendant l'entraînement, les modèles minimisent leur fonction de perte en définissant tous les coefficients de phase à 0 afin d'éviter toute divergence car la géométrie de la PSF ne peut pas être correctement identifiée dans le bruit. À des SNR élevés, dans les deux scénarios, le plus grand modèle (NN5) converge vers la même valeur de WFE. Cependant, le modèle NN3 du *scénario 2* présente une légère limitation, atteignant un WFE légèrement inférieur à 10 nm , alors que le même modèle entraîné sans bruit performe mieux.

Conclusion

Les réseaux estiment correctement les paramètres de la PSF pour des $SNR \leq 100$. La stratégie d'entraîner les réseaux sur des données bruitées donne les meilleurs résultats à partir du moment où l'on peut estimer en amont le SNR des données.

Le SNR ne peut être qu'estimé et ne correspond donc pas forcément aux valeurs attendues avec précision, ce dernier pouvant varier au cours du temps. Une approche alternative consiste à entraîner des modèles en utilisant une gamme de SNR tirée uniformément dans un intervalle afin d'évaluer leur performance.

3.4.2. Robustesse au SNR

On propose donc maintenant d'entraîner 2 réseaux (NN3 et NN5) sur une gamme variable de SNR entre 30 et 100, et de comparer leur performance par rapport aux réseaux entraînés à SNR fixe.

3. Étude d'une source ponctuelle – 3.5. Influence des aberrations de hauts ordres

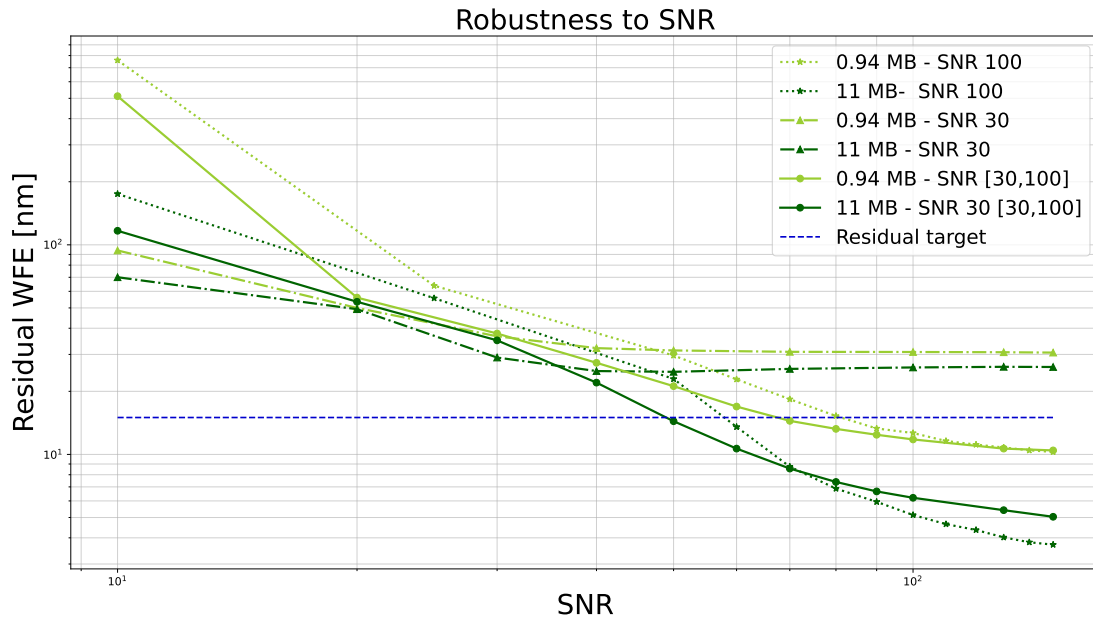


FIGURE 3.21. – Robustesse des réseaux en fonction du SNR. La courbe verte foncé illustre la qualité de l'estimation du NN5; la courbe verte clair celle du NN3.

La figure 3.21 illustre la comparaison de la qualité de l'estimation du front d'onde entre les modèles entraînés sur un SNR fixe et ceux entraînés sur des SNRs variables. Les modèles entraînés à SNR variables montrent une bonne performance, atteignant plus rapidement la limite de diffraction à SNR=50 pour le NN5 et SNR=70 pour le NN3. De manière surprenante, même à des SNRs élevés, les modèles entraînés sur des SNRs variables démontrent des performances comparables à celles des modèles entraînés sur un SNR fixe. Cependant, dans un régime de flux très bas, les performances se détériorent et divergent.

Conclusion

En général, un SNR minimum de 50 est nécessaire pour atteindre la limite de diffraction avec un modèle relativement petit.

3.5. Influence des aberrations de hauts ordres

On souhaite étudier maintenant la robustesse de notre méthode face aux aberrations de plus hauts ordres. Ces aberrations consistent en des perturbations de front d'onde à plus haute fréquence spatiale que les coefficients de piston, tip et tilt estimés sur les segments. Elles ont pour origine les défauts de réalisation des optiques (erreurs

3. Étude d'une source ponctuelle – 3.5. Influence des aberrations de hauts ordres

de polissage, de forme). Elles participent à la dégradation de la qualité d'image, de la même façon que les erreurs de phasage ou d'orientation des segments. Cependant dans le cas de la plate-forme miniaturisée qui nous intéresse, nous ne cherchons pas à les corriger (d'ailleurs nous ne pouvons pas le faire). Par contre elles peuvent venir perturber la mesure des coefficients corrigeables (PTT). Elles ne sont donc pas estimées par le réseau, elles interviennent uniquement comme une source de additionnelle de bruit et n'interviennent pas dans le calcul du WFE résiduel. On présente deux scénarios comme décrit ci-dessus.

- **Scénario 3** : Chaque modèle est entraîné avec un ensemble de données où les aberrations de haut ordre ne sont pas présentes dans la formation des données d'entraînement. Mais elles sont présentes dans les données de test. Le réseau est ainsi testé sur des PSF générées avec PTT et des aberrations de haut ordres.
- **Scénario 4** : Chaque modèle est entraîné avec des échantillons générés à partir de PTT et d'une carte de phase d'aberration de haut ordre. Le modèle est ensuite testé sur des données souffrant de la même amplitude d'aberration de haut ordres.

Pour ces deux scénarios, les 12 coefficients de PTT suivent la même distribution. Ce sont les seuls coefficients inférés par le modèle puisque les aberrations de hauts ordres ne peuvent pas être compensées, et qu'ils complexifieraient l'architecture et l'entraînement du réseau. Le *Residual WFE* (figure 3.22) se réfère ici à la partie PPT de la phase (c'est-à-dire en excluant les aberrations de haut ordres).

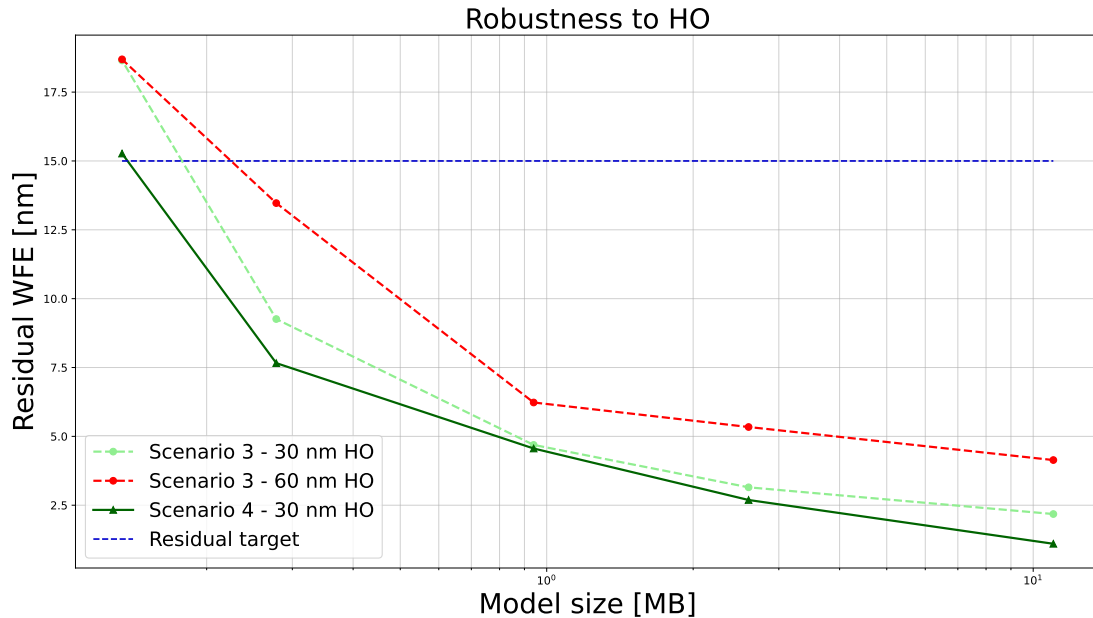


FIGURE 3.22. – Robustesse des réseaux a des aberrations de plus haut ordre. En vert les courbes des deux scénarios identifiés. En rouge les réseaux entraînés sans HO testé avec 60 nm de haut ordres.

Les performances des réseaux dans les deux scénarios sont très satisfaisantes, avec presque tous les réseaux atteignant la limite de diffraction attendue. Un avantage majeur des réseaux du scénario 3 réside dans sa robustesse : l'introduction d'aberrations inconnues de 30 à 60 nm ne dégrade l'estimation des coefficients que de quelques nanomètres, permettant de continuer à produire une estimation correcte même en cas de plus fortes perturbations. On peut également remarquer que le réseau le plus robuste aux variations des hauts ordres est le NN3 de taille environ 1 MB. En effet, la différence entre les scénarios 3 et 4 est relativement faible (environ 0.3 nm), tandis que, pour comparaison, le NN5 (11 MB) effectue une meilleure inférence des coefficients de phase mais présente une moins bonne robustesse, avec une différence environ trois fois plus élevée. Néanmoins, la performance reste satisfaisante et atteint les performances de la limite de diffraction. Cette disparité peut s'expliquer par l'augmentation beaucoup plus importante du nombre de paramètres pour le NN5 par rapport au NN3 (1 MB), le rendant ainsi plus enclin à surapprendre directement les données d'entraînement pour optimiser sa performance.

Conclusion

La robustesse des réseaux aux bruits et aux aberrations de hauts ordres permet d'atteindre une nouvelle fois la limite de diffraction imposée par les spécifications d'AZIMOV.

Cependant ces événements ne sont pas indépendants et arrivent (presque) simultanément lors de la formation d'image. On propose donc maintenant d'étudier les performances des réseaux NN3 et NN5 (§) face aux perturbations et de comparer ces derniers aux performances d'estimations des méthodes de l'état de l'art.

3.6. Comparaison aux méthodes d'AFOPF applicables

Dans un premier temps, rappelons le formalisme des méthodes de l'état de l'art, dans le cadre de notre application. Comme mentionné dans la partie 0.2.3, la Diversité de Phase et l'Image Sharpening sont les deux méthodes vues dans ce manuscrit. Cependant, dans le contexte d'AZIMOV, leur utilisation varie de l'application classique.

3.6.1. Diversité de Phase

En effet, les contraintes d'un télescope spatial à petit volume en orbite basse imposent deux inconvénients : l'objet d'observation défile très rapidement, il est donc impossible d'utiliser une image de diversité "temporelle" c'est à dire en alternant une image focalisée et une image défocalisée. Les objets observés dans les deux images ne seront pas identiques, or c'est une contrainte de la Diversité de Phase. De plus, comme aucune optique additionnelle au système imageur n'est envisagée, une image

détériorée par une aberration connue ne peut être acquise simultanément à une image au focus. La minimisation de critère doit donc s'effectuer sur la seule image non dégradée du plan focal. Finalement, on cherche à minimiser le critère entre le modèle de PSF et la PSF réelle tel que :

$$\widehat{\{c_i^k\}} = \operatorname{argmin}_{\{c_i^k\}} \|I - h(\{c_i^k\})\|^2 \quad (3.13)$$

cependant dans les cas de faible SNR, l'estimation des coefficients peut être faussée et mener à une correction très mauvaise. Une régularisation de Tikhonov permet de limiter cet effet. La Diversité de Phase étudiée ici est finalement définie par la minimisation du critère :

$$\widehat{\{c_i^k\}} = \operatorname{argmin}_{\{c_i^k\}} \|I - h(\{c_i^k\})\|^2 + \beta \sum_{i,k} c_i^{k2} \quad (3.14)$$

avec β un coefficient pondérateur de la régularisation de Tikhonov. La Diversité de Phase réalise alors une optimisation numérique de son critère à partir du modèle de formation d'image défini, en dérivant partiellement le critère par rapport aux coefficients à estimer afin de trouver les valeurs optimales des coefficients qui minimisent le critère de l'équation 3.14. Afin de minimiser cette quantité, l'algorithme de minimisation des gradients conjugués ([HS+52]) a montré les meilleures performances pour la Diversité de Phase.

3.6.2. Image Sharpening

L'IS consiste à maximiser une métrique directement mesurable dans l'image, l'avantage de cette méthode est qu'elle ne dépend d'aucun modèle. Cette méthode utilise directement une optimisation optique des composants du miroir primaire pour ajuster la métrique choisie. Dans notre cas, la métrique à maximiser est la somme des carrés de l'intensité des 4 pixels centraux de l'image telle que :

$$\widehat{\{c_i^k\}} = \operatorname{argmax}_{\{c_i^k\}} \sum_{x=x_c-W}^{x_c+W} \sum_{y=y_c-W}^{y_c+W} I(x, y) \quad (3.15)$$

avec x_c et y_c les coordonnées du centre de l'image I et W la taille de la fenêtre choisie (ici, 2×2). L'Image Sharpening réalise alors une optimisation optique de son critère à partir de l'analyse de l'image ainsi formée à chaque itération, en dérivant partiellement le critère par rapport aux coefficients à estimer afin de trouver les valeurs optimales des coefficients qui minimisent le critère de l'équation 3.15. L'algorithme de minimisation de Powell ([Pow64]) a montré les meilleures performances pour cette méthode. On cherche donc à comparer ces deux méthodes itératives à la méthode de Deep Learning étudiée en terme de performance et de temps d'exécution. On propose une comparaison des principales caractéristiques de ces méthodes tableau 3.1.

3. Étude d'une source ponctuelle – 3.6. Comparaison aux méthodes d'AFOPF applicables

	Deep Learning method	Image Sharpening	Phase Diversity
Data Driven	Yes	No	No
Model-free	Yes	Yes	No
Iterative	No	Very	Yes
On-board computing	Deterministic	Optical optimization	Numerical optimization
Off-board computing	Stochastic	None	None
Initial guess	No	Yes	Yes

TABLEAU 3.1. – Comparaison des méthodes d'estimation du front d'onde plan focal mono-image

Le tableau compare trois méthodes d'AFOPF mono-image : Deep Learning, l'Image Sharpening, et la Diversité de Phase. Le Deep Learning est basée sur les données et offre une estimation presque déterministe une fois entraîné, mais elle nécessite cet entraînement préalable et dépend uniquement des données fournies. En revanche, l'Image Sharpening utilise une optimisation optique et est itérative, tout comme la Diversité de Phase, qui utilise une optimisation numérique. Tandis que Deep Learning est n'est pas dépendant du moindre modèle et non itératif, les deux autres méthodes requièrent une condition initiale.

3.6.3. Comparaison et temps d'estimation

Afin d'évaluer les performances des différentes méthodes, elles ont été testées sur le même jeu de données : des PSFs aberrées avec 30 *nm* d'aberrations de haut ordres dont on étudie dans un premier temps la propagation de bruit pour les 3 méthodes testées sur 1000 PSFs. Dans l'étude de propagation de bruit, la méthode d'optimisation de la Diversité de Phase propose un critère d'arrêt très petit afin de mettre en avant la performance maximale d'un tel algorithme indépendamment du temps de calcul. Le critère d'arrêt est alors défini par $\epsilon^{PD} = 10^{-15}$.

3. Étude d'une source ponctuelle – 3.6. Comparaison aux méthodes d'AFOPF applicables

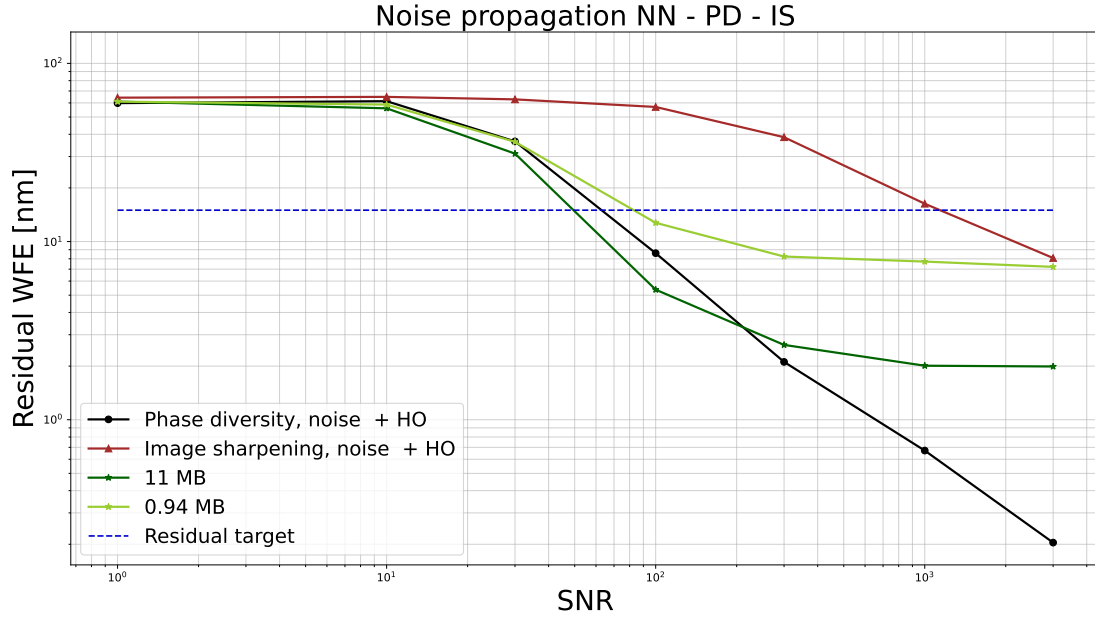


FIGURE 3.23. – Comparaison de la propagation de bruit de la Diversité de Phase à $\epsilon^{PD} = 10^{-15}$ (courbe noire), de l'Image Sharpening (courbe marron) et des réseaux NN3 (courbe verte claire) et NN5 (courbe verte foncée).

On remarque alors que lors de la propagation de bruit, les tendances des réseaux de neurones et de la Diversité de Phase se suivent. En cas de fort rapport signal à bruit, la Diversité de Phase performe bien mieux que les réseaux de neurones, estimant une erreur de front d'onde inférieure au nanomètre. L'Image Sharpening est plus sensible au bruit et à bas rapport signal à bruit performe nettement moins bien que les autres méthodes, étant basé uniquement sur l'intensité des pixels de l'image. On s'intéresse maintenant au temps de calcul de ces méthodes, Afin de prendre en compte le temps de calcul de l'Image Sharpening, un temps d'exposition de 0.1 ms est considéré et doit être pris en compte pour le temps de calcul total car une image doit être acquise à chaque itération de l'algorithme. Cependant, dans le cas de la simulation numérique des PSFs, à chaque itération le temps de calcul est soustrait du temps total, à la différence de la méthode de Diversité de Phase. Dans cette dernière, l'optimisation numérique nécessite bien un calcul numérique de la PSF à chaque étape, on ne soustrait donc pas le temps de simulation de PSF. Pour une comparaison équitable, on propose deux nouveaux critères d'arrêt de l'optimisation de la Diversité de Phase : moins restrictif que le premier, ils permettent une qualité d'estimation relative à une vitesse de calcul plus raisonnable. On propose maintenant $\epsilon_1^{PD} = 10^{-7}$ et $\epsilon_2^{PD} = 10^{-4}$. L'un permet une estimation fiable à un certain temps de calcul, l'autre permet de relaxer le critère pour converger plus vite, au pris d'une estimation moins précise. On propose de tester les dites méthodes pour des données bruitées à SNR = 100, toujours soumises aux 30 nm d'aberrations de haut ordres.

3. Étude d'une source ponctuelle – 3.6. Comparaison aux méthodes d'AFOPF applicables

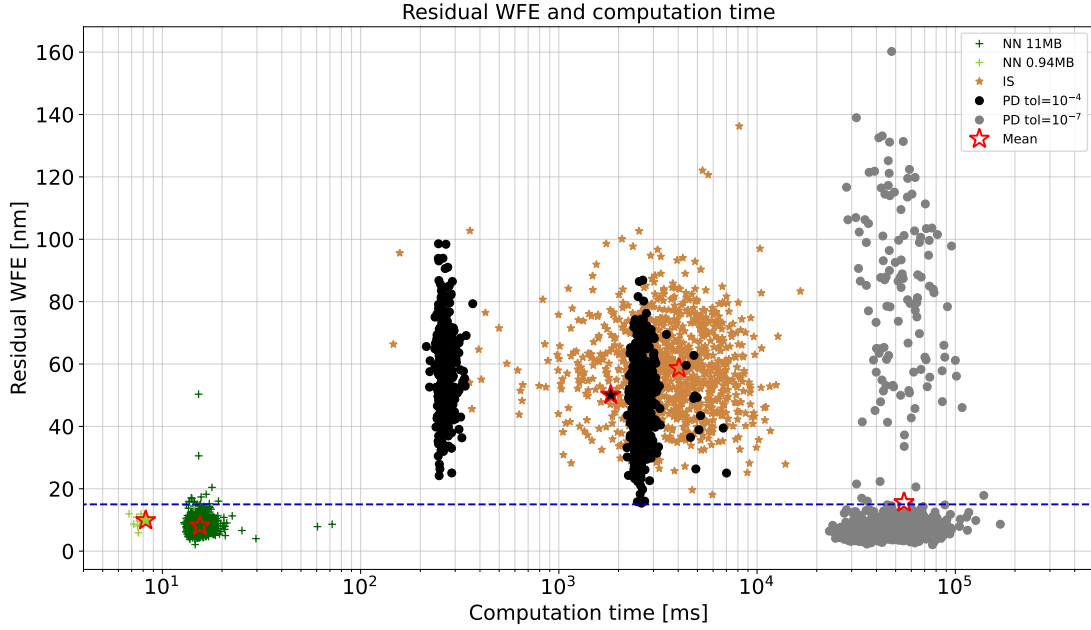


FIGURE 3.24. – Comparaison du temps de calcul et de la performance des 3 méthodes.

L'estimation des coefficients par réseaux de neurones présente une amélioration significative en termes de temps de calcul par rapport à la PD et l'IS. Les modèles, avec leur architecture relativement légère, sont beaucoup plus rapides, la majorité du temps de calcul étant consacré à l'entraînement (ce qui n'est évidemment pas pris en compte pour l'inférence de coefficients). La figure 3.24 illustre les performances des deux méthodes en fonction de leur temps d'exécution. Le critère $\epsilon^{PD} = 10^{-15}$ et $\epsilon_1^{PD} = 10^{-7}$ atteignent sensiblement les mêmes performances à SN = 100, mais avec un temps de calcul différent. Comparé à l'inférence des réseaux, qui est stable en termes de temps de calcul, la Diversité de Phase présente des limitations en termes de temps d'exécution. Tout d'abord, elle prend plus de temps que les inférences des réseaux, et deuxièmement, l'écart-type du temps de calcul est plus important. De plus, 2 clusters distincts sont identifiés lors du calcul de la Diversité de Phase avec un critère de tolérance de $\epsilon_2^{PD} = 10^{-4}$, ce qui indique que le nombre d'itérations pour atteindre la convergence dépend de la forme de la fonction de coût à optimiser. De meilleures estimations peuvent être obtenues avec un critère de convergence $\epsilon_1^{PD} = 10^{-7}$. Cependant, cela nécessite un temps de calcul plus long. De plus, bien que la Diversité de Phase atteigne en moyenne les performances de la limite de diffraction (étoile rouge), elle présente environ 100 valeurs aberrantes sur 1000 échantillons. Alors que l'Image Sharpening est largement répandu en termes de temps de calcul et n'est pas fiable, seul la Diversité de Phase suit les performances du Deep Learning en termes de WFE résiduel. Cependant, le temps d'estimation est au moins 3 ordres de grandeur plus lent par rapport à l'estimation des réseaux de neurones.

Conclusion

La méthode des réseaux de neurones semble être plus fiable pour estimer avec précision le front d'onde par rapport à une Diversité de Phase mono-image et à l'Image Sharpening.

3.7. Scenario de phasage complet

Maintenant que l'efficacité de notre méthode est démontrée et que ses niveaux de performance ont été comparées à l'état de l'art, on peut s'intéresser à la stratégie de phasage complète sur point source. Une fois en orbite dans l'espace, le télescope déploie son miroir primaire. La précision de déploiement constatée sur le modèle d'AZIMOV est de l'ordre de 2 à 5 μm de positionnement des segments. Cette précision représente une erreur de front d'onde très grande devant la longueur d'onde. Le déploiement des segments, est alors opéré par une méthode telle que ELASTIC ([VCM17]), permet de repérer toutes les sous-PSFs dans le champ de vue du détecteur. La stratégie de phasage en plan focal commence donc par une étape de phasage grossier (Coarse phasing), suivie d'une étape de phasage fin (Fine phasing). Comme expliqué partie 1.4.2, la PSF n'est pas sensible au piston lorsque les sous-PSFs sont éparpillées dans le plan focal. Par conséquent, le phasage doit alors s'effectuer en 2 étapes au minimum. La première consiste alors à recentrer les sous-PSFs au centre du champ, sans chercher à corriger l'erreur de piston, afin de les faire interférer au centre du champ. Le but est de faire se recouvrir géométriquement les images de chaque segment afin de faire apparaître un signal d'interférence entre les segments. Ce signal sera la signature du piston différentiel entre segment. Donc, dans un second temps, la PSF alors générée peut être corrigée pour du phasage fin avec l'identification du piston et des tip-tilts résiduels. On propose donc 2 réseaux de neurones en cascade, l'un pour l'identification des tips-tilts importants à partir d'un certain seuil de front d'onde aberrant; l'autre pour la correction des erreurs résiduelles et l'identification du piston. Un schéma de la stratégie de phasage par Deep Learning est présenté en figure 3.25.

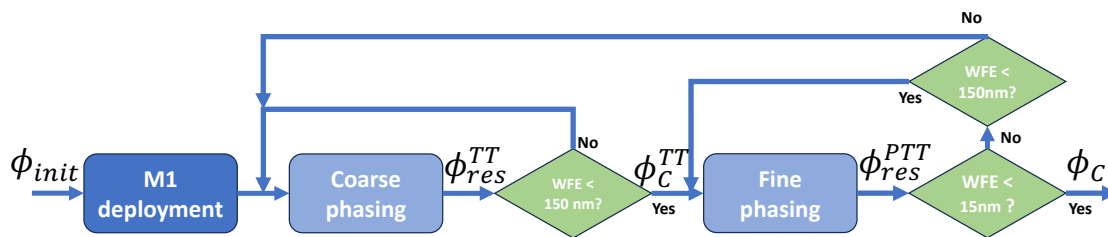
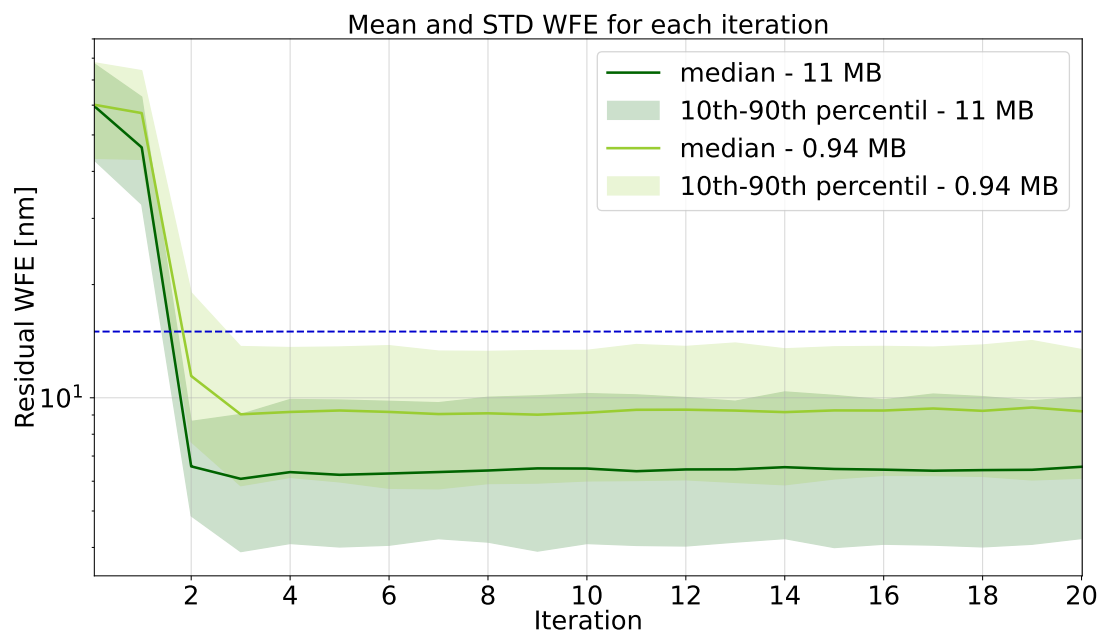


FIGURE 3.25. – Schéma de la stratégie de phasage complet du télescope.

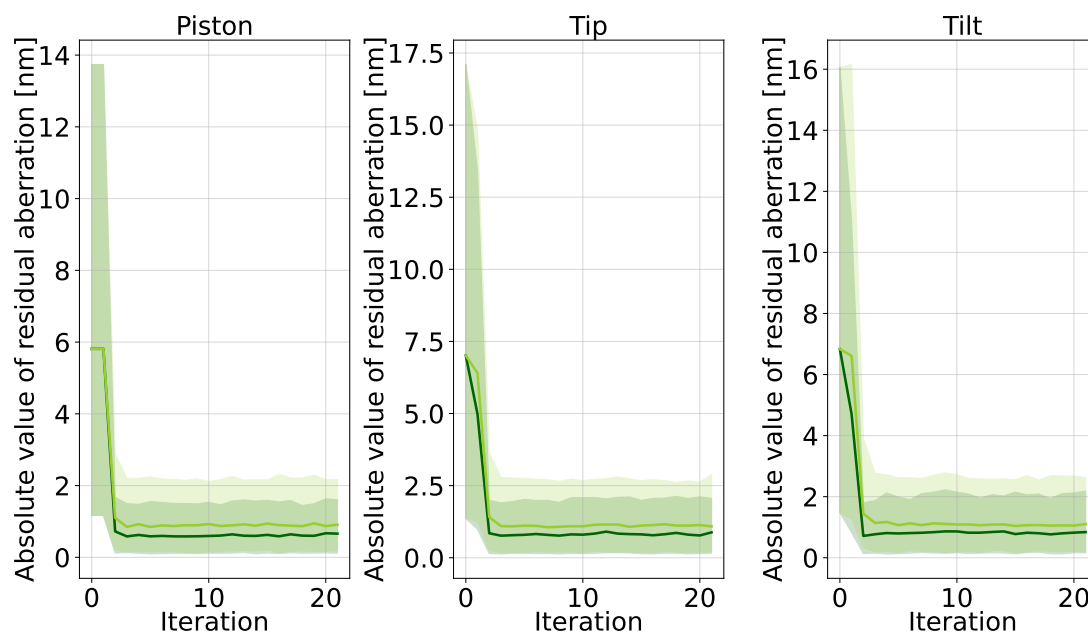
Cette stratégie de phasage peut être implémentée en boucle fermée pour tester le

3. Étude d'une source ponctuelle – 3.7. Scenario de phasage complet

régime d'opération complet du phasage du télescope. Un gain de boucle de 0.9 est appliqué et un tirage de bruit aléatoire tel que étudié en partie 3.4 est appliqué à chaque itération à un SNR de 100.



(a) WFE résiduel au cours d'une boucle fermée.



(b) Erreur moyenne sur les segments des 3 modes identifiés.

FIGURE 3.26. – Résultats des réseaux NN3 et NN5 en boucle fermée.

La boucle fermée proposée ici pour les réseaux NN3 et NN5 (§) converge vers des

3. Étude d'une source ponctuelle – 3.7. Scenario de phasage complet

valeurs résiduelles de front d'onde en dessous des spécifications de 15 nm . Le NN5 converge directement au bout de deux itérations jusqu'à sa performance nominale, le réseau NN3 au bout de 3 ou 4 itérations. Néanmoins, la vitesse de cette boucle est bien plus rapide que la vitesse d'évolution des perturbations, liée à l'orbite autour de la Terre (et donc à un temps caractéristique de 90 minutes). Les deux réseaux montrent une excellente stabilité en boucle fermée et permettent d'assurer un phasage constant et fiable en temps réel. On observe la correction des tip-tilt uniquement sur la figure 3.26 en cas de front d'onde trop important, le plateau observé à la première itération du piston résiduel montre bien la non correction (imposée) de ce dernier. Ensuite, après la deuxième itération, le réseau responsable de l'identification fine des tip-tilts corrige les aberrations restantes, convergeant finalement vers environ un nanomètre de piston et de tip-tilts résiduel, résultant en un WFE de moins de 10 nm sur toute la pupille. En présence de bruit à $SNR = 100$, l'écart-type de l'erreur d'estimation augmente sensiblement.

Conclusion

L'estimation reste largement en dessous de l'exigence requise et montre une grande stabilité pour 100 itérations de boucle fermée pour les deux modèles.

De plus, l'amplitude des tip-tilts de notre étude est limitée par la taille des images (ici, 64×64 pixel, ce qui revient, échantillonné à Nyquist avec une diffraction à 0.3 arcsec , à un FoV de 20×20 arcseconde), un front d'onde plus important peut aussi être corrigé à partir du moment où les sous-PSFs sont sur le détecteur.

Dans cette section, nous avons exploré les performances des réseaux neuronaux pour le phasage d'un télescope spatial déployable à 4 segments avec des aberrations de piston, de tip et de tilt dans la plage de $(20 - 120\text{ nm})$ dans le cas où l'objet est un point source. Les algorithmes existants les plus avancés sont considérés comme complexes à mettre en oeuvre en raison de leur caractère itératif et de leur charge calculatoire importante et longue pour les applications spatiales. Nous avons proposé une méthode alternative fondée sur les réseaux de neurones. Des architectures de réseaux neuronaux simplifiées ont été explorées, présentant des résultats prometteurs pour estimer avec précision les erreurs de phasage de piston-tip-tilt dans le cas réaliste du télescope AZIMOV. Notre stratégie de phasage peut aussi s'étendre à d'autres télescopes et d'autres problématiques. On propose maintenant d'étudier le cas du phasage de plus gros télescope allant jusqu'à 18 segments et le cas nominal de la mesure du mode pétale de l'ELT sans turbulence.

3.8. Application à de plus grands télescopes

La transition du phasage d'AZIMOV, un télescope déployable comprenant quatre segments de miroir, à celui du JWST, qui en compte dix-huit, représente un saut significatif dans la complexité. Avec un nombre de segments plus élevé, le défi de l'alignement devient considérablement plus complexe. Cependant, la méthode développée pour AZIMOV, utilisant le Deep Learning pour l'AFOPE, offre néanmoins un processus solide pour aborder le phasage de télescope de grande taille.

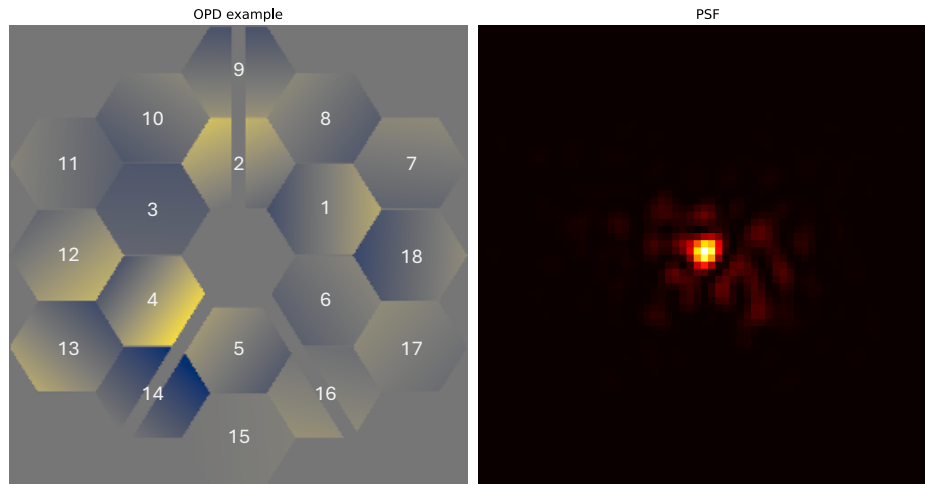


FIGURE 3.27. – Exemple de la pupille et d'une carte d'OPD, avec la PSF correspondante pour la pupille du JWST.

La figure 3.27 présente la pupille utilisée et le numéro des segments associé ainsi que la PSF correspondante à la carte d'OPD. On remarque que la pupille est par design non centro-symétrique due à la position des spiders, cela va directement servir de diversité d'amplitude. On propose d'étudier deux régimes haute performance, où les résidus de front d'onde sont faibles et les sous-PSFs géométriquement superposées, afin d'identifier les piston-tip-tilts résiduels. On propose alors deux distributions des coefficients en effectuant un tirage aléatoire sur l'ensemble des 54 paramètres à estimer, suivant une loi de distribution normale à $\sigma_{JWST}^1 = 30 \text{ nm PV}$ et $\sigma_{JWST}^2 = 70 \text{ nm PV}$. Comme énoncé dans la partie formation d'image, le piston est une nouvelle fois tiré selon une gaussienne d'écart-type σ et tronqué à l'intervalle $]-\frac{\lambda}{4}, \frac{\lambda}{4}[$. Les valeurs ici sont choisies en valeur *Peak to Valley (PV) PV* et non en valeurs RMS. On propose de quantifier l'erreur de front d'onde par la somme de la valeur absolue des erreurs PV. La distribution de l'erreur de front d'onde pour le JWST en valeur PV est présentée figure 3.28.

3. Étude d'une source ponctuelle – 3.8. Application à de plus grands télescopes

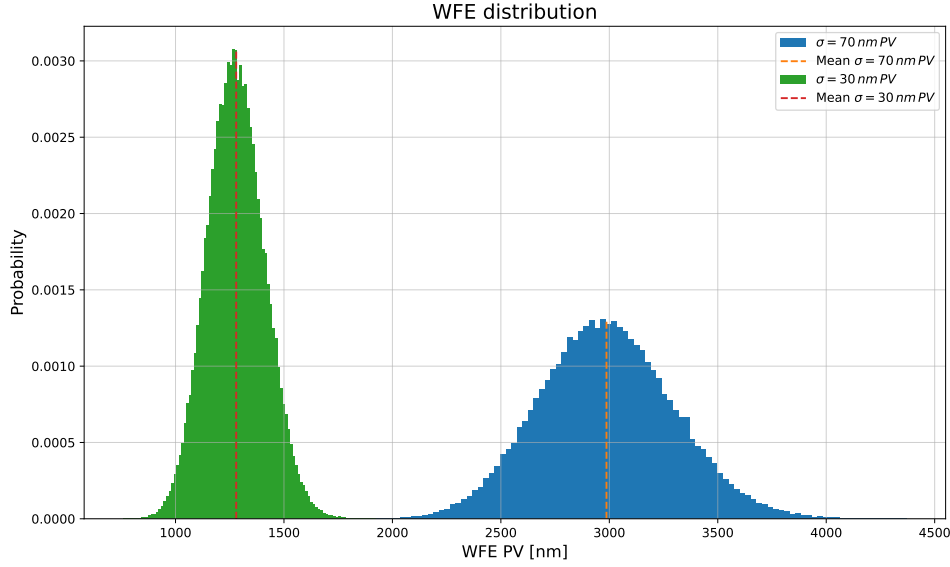


FIGURE 3.28. – Distribution du front d'onde PV total sur la pupille JWST.

On cherche une nouvelle fois, à partir d'une unique PSF, à déterminer les coefficients Piston-Tip-Tilt de chacun des segments.

3.8.1. Réseaux de neurones

Dans cette étude, nous envisageons d'évaluer les performances des deux réseaux Resnet NN3 et NN5. L'architecture reste la même pour les réseaux, néanmoins ils infèrent plus de coefficients, augmentant le nombre de paramètres internes en conséquence. L'entraînement est réalisé en utilisant 10^5 PSFs sur une période de 1000 époques. Le learning rate scheduler ajuste la valeur du taux d'apprentissage à partir de la 750ème époque et suit la même loi de décroissance décrite équation 3.12.

3.8.2. Résultats

La figure 3.29 expose l'amplitude des erreurs de l'estimation du PTT par segment pour 1000 PSFs du jeu de données de test. Pour chacun des segments, l'erreur est moyennée sur l'ensemble des 1000 PSFs testées et est présentée en fonction des modes piston, tip et tilt.

3. Étude d'une source ponctuelle – 3.8. Application à de plus grands télescopes

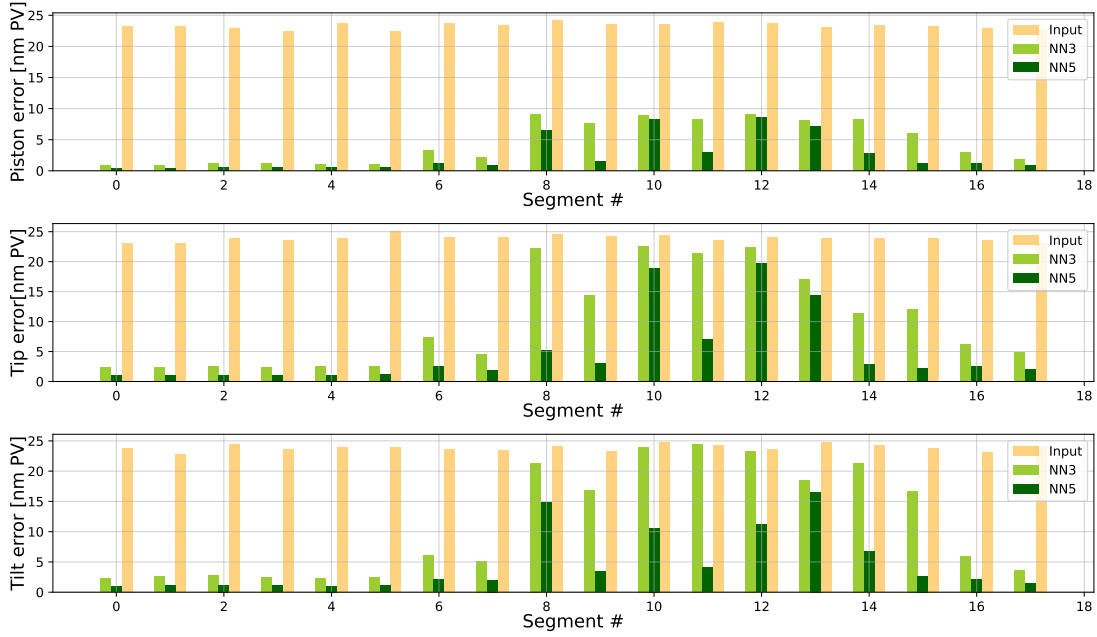


FIGURE 3.29. – Amplitude des erreurs d'estimation pour un front d'onde $\sigma = 30 \text{ nm PV}$.

Les aberrations de PTT sont correctement estimées dans les cas de faibles aberrations pour les premiers segments de 1 à 7, malgré l'augmentation du nombre de coefficients estimés et, par conséquent, un WFE plus élevé. Cependant on peut observer que pour les deux réseaux étudiés, l'erreur d'estimation est bien plus élevée pour les segments 8, 9, 10, 11, 12, 13 et 14. Ce sont ces segments qui a eux seuls représentent plus de 90% du front d'onde résiduel. On peut alors supposer que le réseau optimise les segments un à un et que la complexité des réseaux proposés n'est pas suffisante pour représenter l'ensemble des segments du JWST. La performance du réseau NN5 est nettement supérieure à celle du NN3. En effet, l'amplitude des erreurs (en vert foncé) est jusqu'à cinq fois plus petite pour le segment 11. Ici se dessine une première limite aux réseaux minimalistes pour le phasage de plus grands télescopes, une nette limitation au nombre de coefficients à estimer est montrée. On observe la même tendance lors du phasage grossier du JWST, à partir d'une erreur du front d'onde initial plus importante, certains segments sont mal estimés par le réseau alors que d'autre proposent une meilleure estimation.

3. Étude d'une source ponctuelle – 3.8. Application à de plus grands télescopes

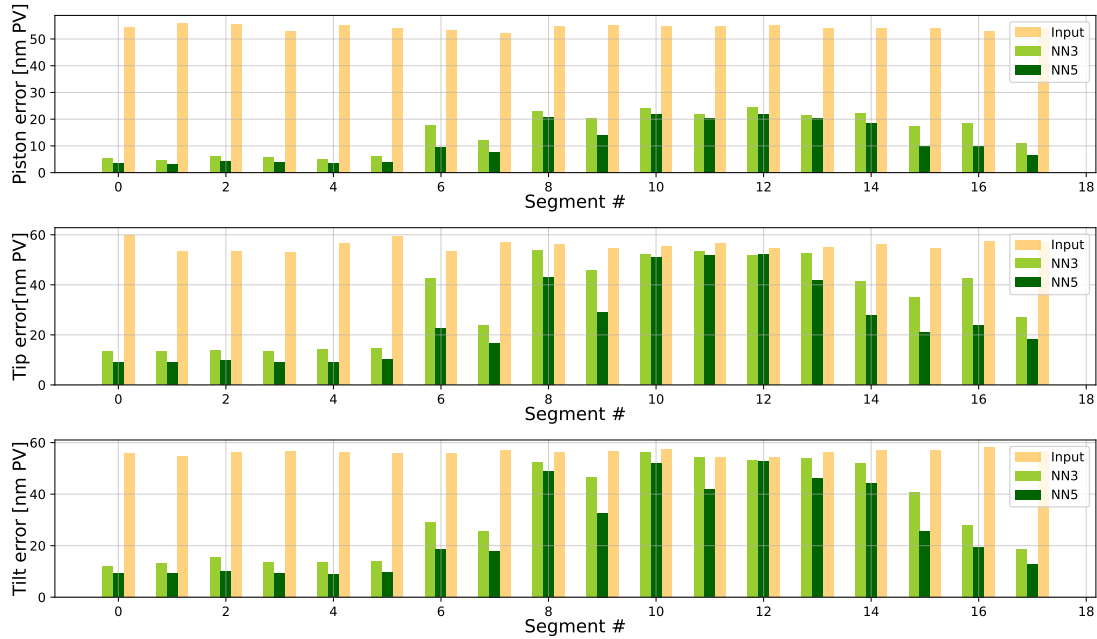


FIGURE 3.30. – Amplitude des erreurs d'estimation pour un front d'onde $\sigma = 70 \text{ nm PV}$.

La qualité d'estimation des modes sur les autres segments est impactée par l'amplitude du front d'onde incident aberré. En effet, ici, l'amplitude des erreurs a nettement augmenté lorsque le front d'onde incident augmente : les segments extérieurs de la pupille présentent en moyenne près de 100% d'erreur d'estimation sur le tip-tilt pour le NN3. Une nouvelle fois, le NN5 performe mieux qu'un réseau plus léger mais peine à identifier correctement le front d'onde pour environ 2/3 des segments. On observe une nette limite des réseaux légers pour l'estimation du front d'onde. Néanmoins, il est tout à fait possible, à l'image de la boucle fermée étudiée sur AZIMOV, de mettre en cascade les deux réseaux afin de réduire au mieux le front d'onde. En effet, le front d'onde résiduel des réseaux à $\sigma = 70 \text{ nm PV}$ est en moyenne de 765 nm PV pour le NN5 et de 994 nm PV pour le NN3. Ces valeurs correspondent exactement à la distribution à laquelle le réseau $\sigma = 30 \text{ nm PV}$ a été entraîné.

Cette observation met en évidence la limite de notre méthode pour les télescopes dotés d'un plus grand nombre de segments de miroir mais aussi ses limites. Elle démontre néanmoins son potentiel à contribuer à la correction des aberrations complexes dans les systèmes optiques avancés pour le suivi de phasage résiduel fin. Néanmoins, la qualité de l'estimation peut être nettement améliorée car des erreurs d'estimation persistent et peuvent être conséquentes sur un grand nombre de segments. Plusieurs perspectives peuvent être envisagées. Nous avons vu dans un premier temps que le phasage sur quatre segments par Deep Learning est une méthode efficace pour phaser un télescope segmenté. Une solution envisageable pour un télescope comportant plus de segments serait d'identifier des sous-parties de la pupille composées de

moins de segments tout en gardant les mêmes architectures légères afin de phaser séquentiellement les différents segments du télescope. Dans un autre contexte, le phasage pourrait toujours s'effectuer en une seule itération en utilisant cette fois des architectures de réseaux de neurones plus complexes. Nous avons vu qu'un réseau léger plus complexe (NN5) performe mieux que son jumeau plus léger (NN3). Des architectures plus complexes telles que l'EfficientNet, l'InceptionV3 ou encore d'autres architectures peuvent être envisagées dans le cas de télescopes de plus grande taille, étant moins limités en volume et en puissance de calcul que AZIMOV.

Conclusion

Ces résultats garantissent un phasage quasi-temps réel à faible coût calculatoire sans nécessité d'optique additionnelle, permettant ainsi de maintenir un phasage fin correcte à la limite de la diffraction dans le visible.

3.9. Application au petaling de l'ELT (sans turbulence)

Une application aux très grands télescopes peut être étudiée dans le cas ambitieux de l'ELT. Cette section présente une étude préliminaire et simplifiée de l'ELT et de son mode pétale. Le problème est complexe et dépasse le cadre de cette thèse. Je propose ici une ouverture, sans prétendre fournir une description approfondie et novatrice du problème, mais en explorant la possibilité d'appliquer ma méthode au cas du petaling de l'ELT. Je laisse à la communauté le soin de poursuivre ce travail de manière plus approfondie. Compte tenu de l'étude nominale et des simplifications apportées, les résultats préliminaires obtenus sont encourageants pour des applications futures en présence de turbulence. L'objectif de cette section est de poser les bases d'une étude sans turbulence pour la mesure du petaling de l'ELT. Dans cette étude, j'ai participé à la supervision des travaux de Charleston Chauvet avec Benoit Neichel et Morgan Gray. Les figures sont extraites de ([Cha+23]).

3.9.1. Formation d'image

Comme développé dans la partie 0.4.2, le télescope de l'ELT souffre du petaling ([Lev+24]), une segmentation de la pupille due aux supports du miroir secondaire (spiders), ce qui entraîne une discontinuité de phase du champ électromagnétique au niveau du plan pupille. Cette discontinuité rend la pupille de l'ELT semblable à une pupille découpée en 6 parties. Cette segmentation engendre de nombreuses contraintes. Comme toute erreur de piston, elle est limitée dans son identification pour les méthodes du plan focal. Cependant, l'ELT étant un télescope au sol avec un volume bien moins limité que les télescopes spatiaux, de nombreux instruments optiques sont prévus afin de faciliter la mesure du front d'onde. En particulier, pour

le mode Laser Tomography de HARMONI, un senseur de type Shack-Hartmann à 2×2 sous-pupilles est prévu pour faire la mesure des trois modes tip-tilt et focus sur l'étoile laser. Nous souhaitons utiliser ce senseur simple pour mesurer également les modes pétales de l'ELT. Cette idée simple se confronte cependant à une problématique : mesurer une discontinuité de phase à l'intérieur d'une sous-pupille, en analysant l'image plan-focal. Les sous-pupilles échantillonnées par le SH génèrent alors des sous-PSFs (ou spots, figure 3.31b) témoignant de l'information sur le piston différentiel entre les pétales compris au sein de la sous-pupille. Une particularité de ce dispositif est que lorsque les spiders sont alignés avec l'axe du SH (figure 3.31a), il est alors strictement impossible d'identifier les pistons différentiels entre l'hémicercle inférieur et l'hémicercle supérieur de la pupille, les deux parties étant complètement indépendantes. De plus, la pupille de l'ELT tourne constamment, modifiant ainsi la proportion de la taille des pétales vue par les sous-pupilles. En fonction de l'angle de rotation, 2 ou 3 segments peuvent se présenter dans la sous-pupille. C'est pourquoi la rotation est limitée à 30° maximum afin d'étudier un seul piston différentiel dans la sous-pupille en haut à droite et 2 pistons différentiels pour la sous-pupille en haut à gauche. L'avantage de l'étude par le SH permet d'apporter une centro-symétrie des cas entre les différentes sous-pupilles et que par rotation, les entraînements de 2 réseaux sur une paire de sous pupilles supérieures peuvent s'appliquer aux 2 autres sous pupilles inférieures.

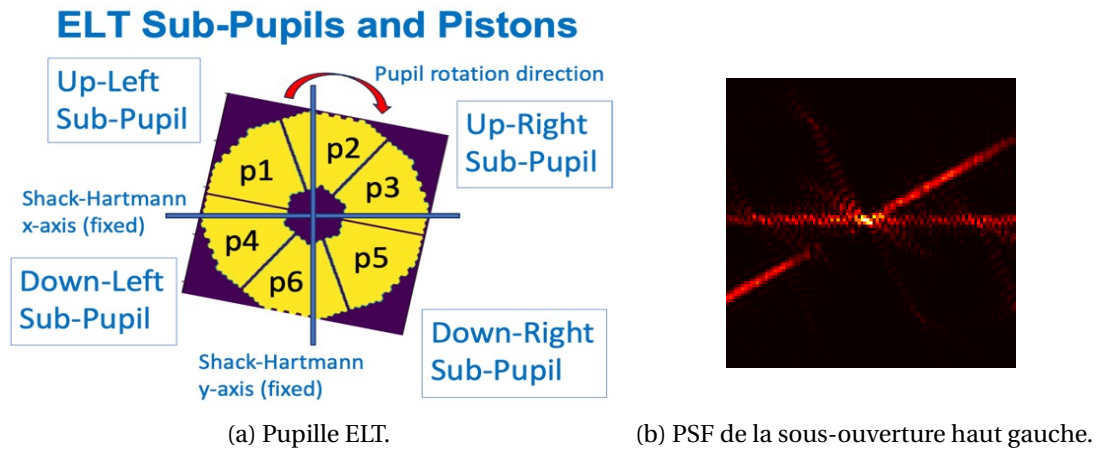


FIGURE 3.31. – Illustration de la pupille ELT segmentée par les spiders et PSF associée à la sous-ouverture haut-gauche du Shack-Hartmann. Extrait de [Cha+23]

Les PSFs ainsi générées correspondent à un certain angle de rotation de la pupille, qui est distribué durant l'entraînement des réseaux. Ces PSFs, de taille 97×97 , sont générées uniquement à partir de la sous-pupille délimitée par l'ouverture du SH, c'est pourquoi la forme des PSFs varie tant par rapport aux PSFs précédemment étudiées. Les valeurs des pistons sont tirées aléatoirement dans l'intervalle $[-\frac{\pi}{2}, \frac{\pi}{2}]$ à $\lambda = 2.2 \mu m$ qui équivaut à $[-550, 550] nm$. Les sous-PSFs sont échantillonnées à Nyquist et sont

générées avec un tirage aléatoire d'angle de pupille variant de 1° à 30° avec un pas de 1° . Quelle que soit la valeur du piston différentiel à estimer ou la configuration des effets optiques, 29 ensembles de données ont été générés. Chaque ensemble contient 3000 images de sous-PSF de test, toutes obtenues à partir d'un seul angle de rotation compris entre 1° et 29° , soit un total de 87 000 échantillons.

3.9.2. Réseau de neurones

Pour résoudre le problème de l'identification du piston, on propose une nouvelle architecture, basée une nouvelle fois sur l'architecture Resnet vue plus tôt, avec de légères modifications : le Resnet se compose de deux couches convolutives initiales. La sortie de ces couches est non seulement transmise à un premier bloc de deux couches convolutives supplémentaires, mais aussi via une *skip connection* dont la sortie est ajoutée avant la fonction d'activation de la deuxième couche convolutive de ce premier bloc. La sortie de ce premier bloc est ensuite transmise de la même manière à un deuxième bloc de deux couches convolutives supplémentaires, avec une autre *skip connection* ajoutant la sortie avant la fonction d'activation de la deuxième couche convolutives de ce deuxième bloc. Cette architecture continue avec deux blocs supplémentaires suivant le même schéma, suivis par une couche de max-pooling. Enfin, deux couches entièrement connectées de 16 neurones chacune et une couche dense finale sont ajoutées, produisant une seule valeur de piston différentiel. Pour que la valeur de l'angle de rotation soit prise en compte, cette information est concaténée aux 16 neurones de la première couche dense. Cela permet au réseau de ne pas avoir à estimer l'angle de rotation et de le prendre en compte comme une donnée d'entrée. La fonction de perte et métrique d'évaluation utilisée est la RMSE.

3.9.3. Résultats

Afin de compenser un maximum les effets de petaling de l'ELT, une valeur seuil est fixée à $RMSE_{min} = 35 \text{ nm}$ pour chaque angle à partir de laquelle on estime une correcte estimation des pistons différentiels. Les résultats sont présentés pour les sous-pupilles haut droite, et haut gauche.

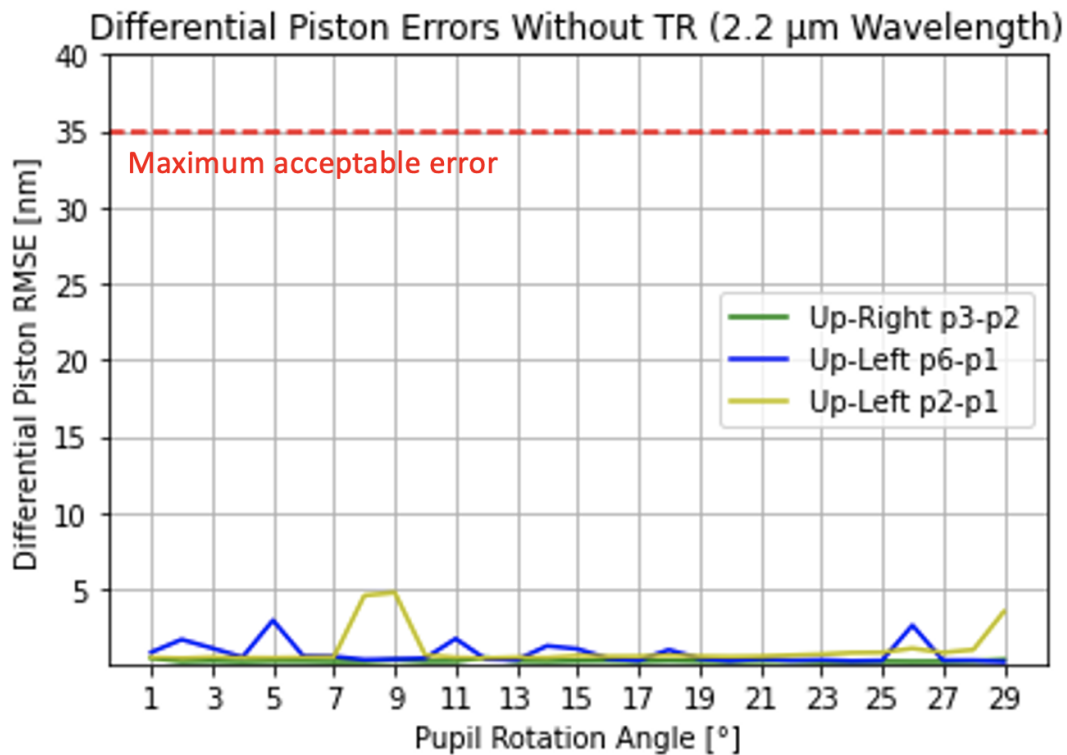


FIGURE 3.32. – Résultats d'estimation des pistons différentiels par réseaux de neurones présenté pour les 2 sous pupilles du haut.

Dans le cas présenté, en l'absence de Turbulence atmosphérique Résiduelle (TR) ou de résidus d'OA, l'estimation des réseaux est très précise et efficace. En particulier, lorsqu'un seul piston différentiel doit être identifié, l'erreur ne varie pas quel que soit l'angle. Dans le cas plus complexe de deux pistons par sous-pupilles, l'estimation est moins stable mais maintient une excellente performance. Évidemment, un cas plus qu'idéal est étudié. Cependant, ces travaux préliminaires ouvrent des perspectives prometteuses pour l'identification complète du mode pétale de l'ELT.

Conclusion

En s'appuyant sur les performances précédemment étudiées lors de l'ajout de sources de perturbations inconnues (partie 3.4 et 3.5), l'estimation précise des erreurs de pistons par les réseaux de neurones apparaît exploitable.

Des résultats plus élaborés en présence de turbulence atmosphérique et correspondant au système de LTAO d'HARMONIE sont disponibles à l'aide des travaux de Pierre Janin (voir proceeding SPIE 2024).

3. Étude d'une source ponctuelle – 3.9. Application au petaling de l'ELT (sans turbulence)

En résumé, les réseaux de neurones se révèlent être une méthode particulièrement efficace pour l'analyse du front d'onde en plan focal. Leur robustesse et leur efficacité ont été démontrées dans ce chapitre à travers différents scénarios et applications. Les réseaux de neurones offrent une nouvelle solution innovante pour assurer un phasage rapide, précis et robuste des télescopes. Cependant, étant une méthode d'apprentissage basée sur les données, leur performance dépend fortement du jeu de données utilisé pour l'entraînement. La possibilité d'une divergence entre les données d'apprentissage et les données expérimentales peut remettre en cause les performances observées précédemment. C'est pourquoi le quatrième chapitre de cette thèse se penchera sur l'importance du choix et de la proportion entre données expérimentales et données simulées.

4. Validation expérimentale

Afin de valider l'approche effectuée en simulation, il est nécessaire de tester cette dernière sur des données expérimentales. Nous avons mené une telle validation avec UK-ATC, en utilisant leur maquette optique du système AZIMOV. Cette validation a été menée dans le cadre d'un stage du fin d'étude de Yohann Daligaud que j'ai entièrement encadré début 2024. L'objectif de ce chapitre est de valider les réseaux développés au chapitre précédent et vérifier que ces derniers peuvent fonctionner en mode hybride, c'est à dire entraîné sur données simulées et en opération sur données expérimentales. Dans un premier temps, je propose d'examiner le banc optique permettant de générer des données expérimentales. Puis, dans le but de déterminer la proportion adéquate entre données simulées et données expérimentales, nous avons développé un jumeau numérique des données afin de simuler numériquement des données les plus proches possibles des données issues du banc. Je vérifierais notamment l'échantillonnage des données acquises en développant la méthode employée et j'ajusterai le simulateur de PSFs aux différents paramètres du banc. Enfin, j'appliquerai la méthode de Deep Learning aux données expérimentales et étudierai la balance entre données simulées et données expérimentales afin de déterminer un seuil optimal de données réelles requis pour la mise en œuvre de ma méthode sur un cas opératoire.

Sommaire

4.1. Acquisition de données	141
4.1.1. Banc optique	141
4.1.2. Génération des données expérimentales	142
4.2. Validation des données	144
4.2.1. Mesure de la fréquence de coupure	144
4.2.2. Ajustement des paramètres expérimentaux	147
4.3. Application aux réseaux de neurones	150
4.3.1. Validation simulée et expérimentale	150
4.3.2. Performance des R-Sim appliqué aux données expérimentale	153
4.3.3. Fine-tuning	154
4.4. Jeu de données optimal	157

4.1. Acquisition de données

4.1.1. Banc optique

Le banc optique utilisé pour la génération des PSFs expérimentales a été développé à UK-ATC, laboratoire en charge du développement d'AZIMOV. Le schéma du banc peut être trouvé ci-dessous.

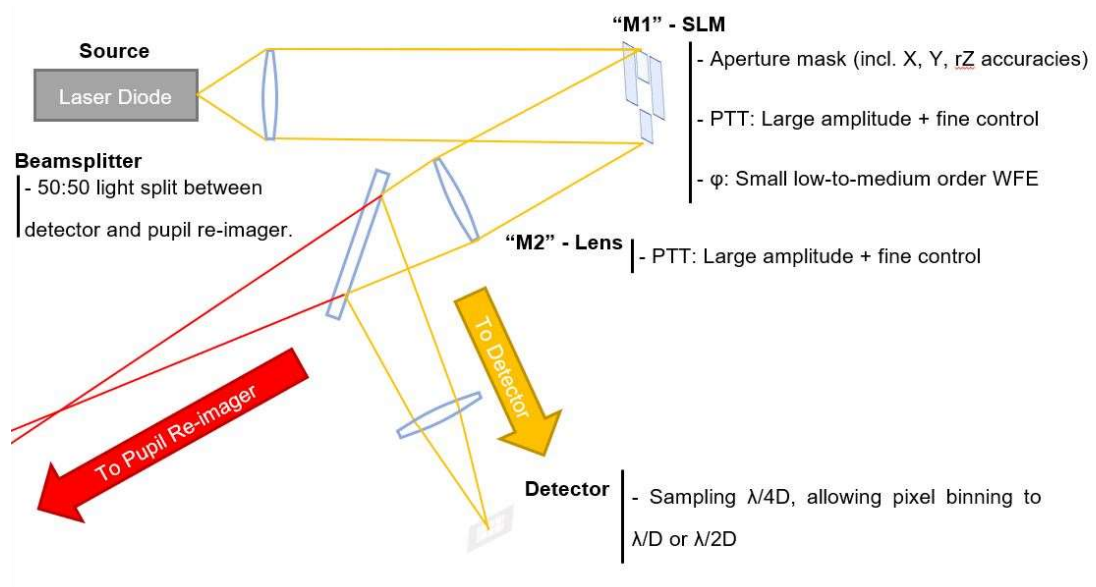


FIGURE 4.1. – Schéma de la maquette optique d'AZIMOV. Extrait du reporting du banc de UK-ATC.

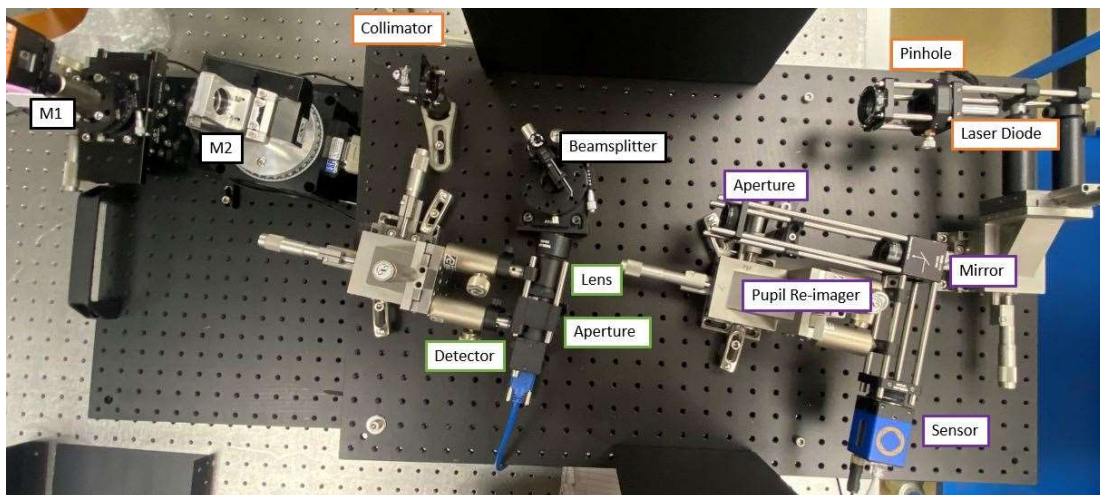


FIGURE 4.2. – Image du banc. Extrait du reporting du banc de UK-ATC.

La lumière monochromatique de la diode laser est collimatée et incidente à un *Spatial Light Modulator* (SLM), agissant comme un substitut de M1. Un SLM placé en faisceau collimaté a deux fonctions : la fonction pupille (génération de la forme particulière AZIMOV) et la fonction d'aberrations de phase (piston, tip et tilt sur chaque segment). Cela permet de contrôler l'erreur en piston, tip et tilt pour chaque segment individuel afin de tester les algorithmes de phasage, ainsi que d'introduire de petites erreurs de bas ordre. Une lentille double achromatique montée sur un ensemble de platines motorisées émule le miroir secondaire d'AZIMOV. Cette lentille est montée sur une platine PTT. La lumière est ensuite divisée par un séparateur de faisceaux 50 : 50. 50% de la lumière est donc focalisée sur le détecteur avec un échantillonnage de $\lambda/4D$, permettant un binning des pixels à λ/D ou $\lambda/2D$, selon la conception étudiée. Les 50% restants de la lumière passent par un point de focalisation avant d'être recollimatés sur un ré-imageur de pupille.

- **Source laser** La diode laser émet un faisceau monochromatique collimaté à $\lambda = 635 \text{ nm}$, grâce à un pinhole de $10 \mu\text{m}$ placé juste après l'ouverture du laser. La lumière diffractée crée ensuite un faisceau divergent, qui passe à travers un polariseur linéaire. La lumière est ensuite recollimatée par une lentille achromatique et arrive sur le SLM.
- **SLM** Le banc utilise un SLM LCOS (Liquid Crystal on Silicon) (Hamamatsu X15213-01). Les SLMs LCOS peuvent moduler la phase de la lumière incidente sur leur surface. La phase de la lumière est contrôlée spatialement en deux dimensions en appliquant des tensions à travers un cristal liquide via un réseau bidimensionnel d'électrodes (pixels). La tension appliquée par chaque électrode incline localement les molécules de cristal liquide, provoquant un changement d'indice de réfraction. Cela permet la création de front d'onde presque arbitraires pour la lumière incidente sur le SLM, limitée uniquement par la plage de décalages de phase que les cristaux liquides peuvent provoquer et la taille des pixels du SLM. Le SLM a un tableau de pixels de 1024×1272 avec chaque pixel séparé de $12,5 \mu\text{m}$, définissant le diamètre maximal du masque de $12,8 \text{ mm}$. Pour tenir compte des tolérances d'alignement et de fabrication des optiques, la conception optique a été optimisée pour créer un échantillonnage correct pour un masque AZIMOV de 12 mm . Le SLM doit être capable de définir simultanément un masque de pupille d'ouverture M1 AZIMOV et de simuler le piston-tip-tilt des pétales M1 individuels lors du déploiement.
- **Détecteur** La lumière issue du M2 est séparée par le beamsplitter. La moitié du flux incident arrive sur le détecteur optique. Le faisceau, focalisé sur le détecteur à l'aide d'une lentille possède une résolution de 1280×1024 pixels de taille $5.3 \mu\text{m}$. Le read-out noise du détecteur est de $11 e^-$ et les images sont encodées sur 10 bits, permettant une plage de 1024 valeurs de pixels.

4.1.2. Génération des données expérimentales

Les données sont générées en grande quantité en faisant varier la tension des pixels du SLM, ce qui permet de contrôler le déphasage en piston et tip-tilt. Les

4. Validation expérimentale – 4.1. Acquisition de données

variations du SLM suivent sensiblement la même distribution que celle étudiée ci-dessus théoriquement, à la différence que le tip et le tilt sont générés à partir de l'angle Ω du segment plutôt que de leur valeur RMS directement. On peut retrouver la valeur PV à partir de l'angle tel que :

$$c_i^k = \tan(\Omega_i^k) \times l_{SLM} \times n_{pix} \quad (4.1)$$

où l représente la taille d'un pixel du SLM et n_{pix} le nombre de pixel nécessaire au SLM pour simuler 1 segment. On retrouve ensuite la valeurs RMS en remontant à la base présentés partie 1.21.

On sépare ensuite 2 type de jeu de données :

- **Sous-PSFs comprennent entre 0 et 1 FWHM**, correspondant au régime de phasage fin.
- **Sous-PSFs comprennent entre 1 et 6 FWHM**, correspondant au régime de phasage grossier où les sous-PSFs sont éclatées dans le champ.

Les distributions des différents coefficient peut être retrouvée en figure 4.3 pour les 100000 données expérimentales générées par jeu de données.

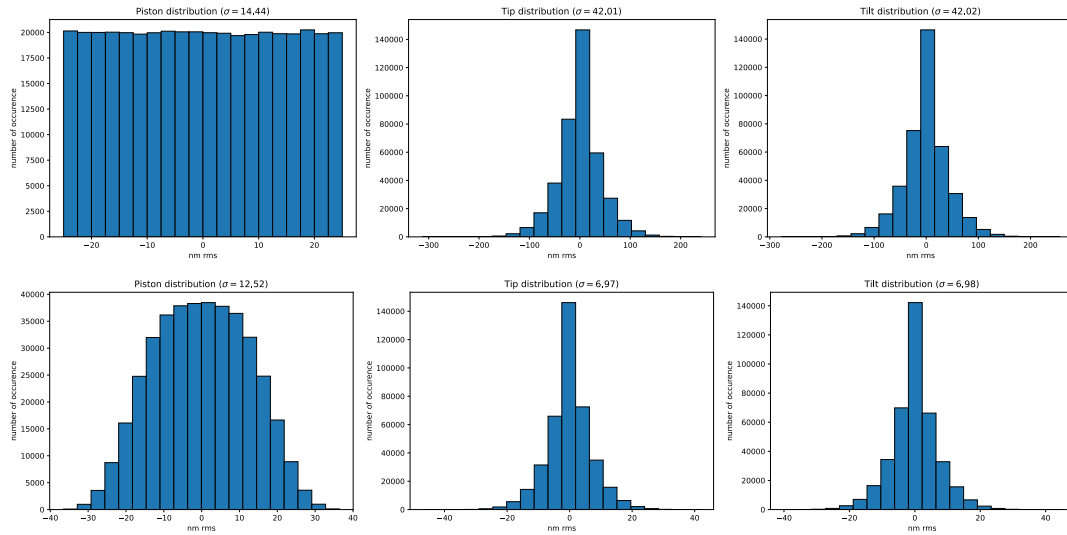
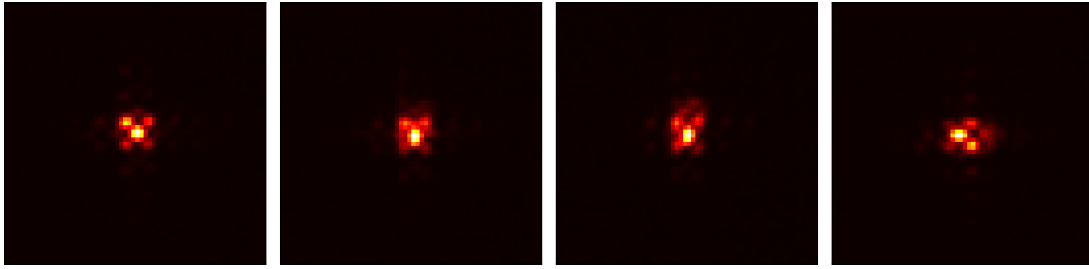
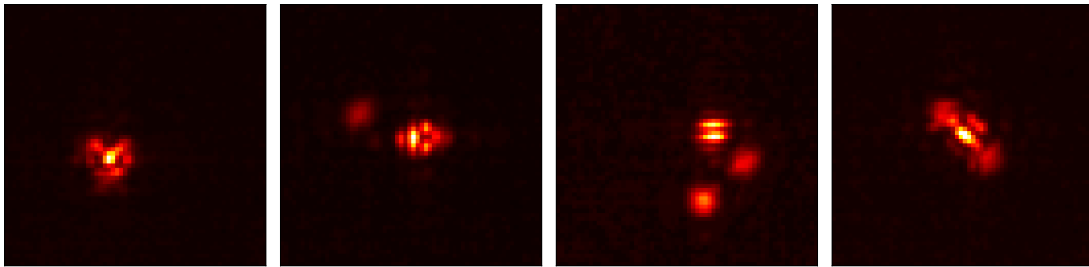


FIGURE 4.3. – Distribution des coefficients des PSFs expérimentales.

La première ligne représente la distribution des coefficients PTT pour le régime de déploiement. La seconde ligne correspond à la distribution des coefficients pour le régime de phasage fin. Afin de correspondre aux sous-PSFs entre 0 et 1 FWHM, l'écart type des distributions est sensiblement inférieur à celui utilisé pour les données simulées. Un exemple des PSFs expérimentales correspondantes peut être trouvé en figure 4.4.



(a) Exemple de PSFs en régime de phasage.



(b) Exemple de PSFs en régime de déploiement

FIGURE 4.4. – Exemple de PSFs expérimentales.

4.2. Validation des données

4.2.1. Mesure de la fréquence de coupure

On souhaite dans un premier temps mesurer l'échantillonnage des PSFs. Le facteur d'échantillonnage est un paramètre optique crucial et un élément clé du modèle de formation d'image. Même si, dans certains cas, comme lors de la formation sur des données expérimentales et leur utilisation pour notre méthode, connaître précisément le facteur d'échantillonnage peut sembler moins essentiel, il reste fondamental de s'assurer que les données sont au moins échantillonnées selon le critère de Shannon. Ce critère garantit que les détails de l'image sont correctement capturés sans perte d'information due au sous-échantillonnage. Il est donc judicieux de mesurer le facteur d'échantillonnage de manière complète dès le départ, car cela sera précieux par la suite, notamment lorsque nous simulerons des images avec un échantillonnage approprié. En anticipant cette étape, nous assurons que toutes les simulations futures respectent les normes d'échantillonnage, permettant une modélisation fidèle et précise des images. Cette mesure n'est pas complexe, mais la présence d'une pupille non circulaire demande de revisiter l'exercice.

L'échantillonnage théorique du banc est à $\lambda/4D$ avec des données ensuite concaténées afin d'obtenir un échantillonnage à λ/D . On cherche à retrouver cet échantillonnage à partir des données fournies par UK-ATC. On propose de développer une approche fréquentielle pour mesurer l'échantillonnage à partir de la fréquence de coupure de la FTO (première annulation de la FTO). Dans le cas d'une pupille monolithique circu-

4. Validation expérimentale – 4.2. Validation des données

laire, la FTO est circulaire et la fréquence de coupure ne varie donc pas selon l'angle θ à partir du centre. Cependant, la pupille du design d'AZIMOV est diluée et composée de segments carrés, la FTO est donc plus complexe. Pour expliquer la méthode et effectuer un premier test, on utilise des données simulées. Cela permet de valider le processus sans contraintes expérimentales réelles et de s'assurer que la méthodologie fonctionne correctement avant de l'appliquer aux données expérimentales.

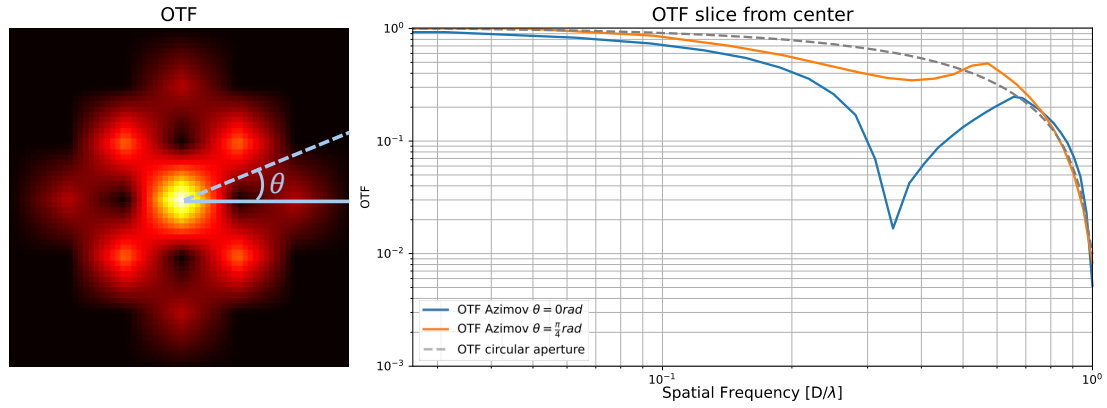


FIGURE 4.5. – Illustration de la FTO AZIMOV comparée à une FTO sur pupille circulaire sans obstruction centrale.

La figure 4.5 montre l'OTF parfaite pour une pupille de la forme de celle d'AZIMOV (image simulée) ainsi qu'une coupe de cette dernière dans deux direction. On constate que les fréquences spatiales transmises par le système optique ne sont pas constantes en fonction de la direction dans le domaine fréquentiel. La forme particulière de la pupille permet même d'imager des fréquences spatiales plus élevées dans certaines directions ($\theta = \frac{\pi}{4}$ rad) comparé à une pupille circulaire de même diamètre. Pour déterminer l'échantillonnage α des données, on s'intéresse à la valeur de la fréquence de coupure par rapport à la taille de l'image. Ce rapport β est dépendant de l'angle θ à partir de l'origine des fréquences. Cependant, lors d'un échantillonnage inférieur à Nyquist, la fréquence de coupure n'est pas identifiable pour les angles proches de $\theta = 0$ et $\theta = \frac{\pi}{2}$.

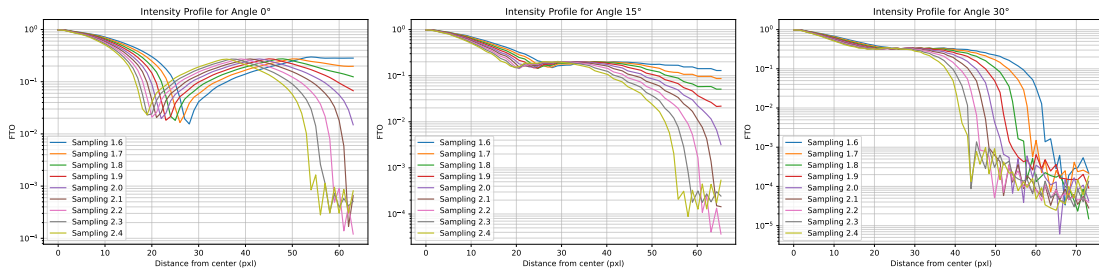


FIGURE 4.6. – Ratio de la position relative β en fonction de l'échantillonnage α pour 3 angles θ .

La figure 4.6 montre une coupe de la FTO en fonction de l'échantillonnage (différentes courbes variant de 1.5 à 2.4) et de l'angle θ (de gauche à droite, $\theta = 0^\circ$, $\theta = 15^\circ$ et $\theta = 45^\circ$). Dans un premier temps on repère un support fréquentiel plus étendu allant de 0 à 64. Un échantillonnage est aussi effectué lors du passage de la PSF à la FTO afin de faciliter l'identification de la fréquence de coupure. On remarque aussi que la fréquence de coupure de la FTO n'est pas positionnée à la même distance par rapport au centre en fonction des angles θ , alors que l'échantillonnage lui ne varie pas. Typiquement, pour un échantillonnage donné, en partant du centre, on sait que le rapport β entre la position du pixel de la fréquence de coupure x_{f_c} et la taille de la demi-image $N/2$ est

$$\beta = \frac{x_{f_c}}{N/2} \quad (4.2)$$

En fonction de la valeur de β , on peut déterminer si le critère de Shannon est respecté ou non. Si $\beta \leq 1$, alors la fréquence de coupure est inférieure ou égale à la fréquence de Nyquist. Dans ce cas, on est correctement échantillonné, ce qui garantit que toutes les informations pertinentes de l'image sont bien capturées sans perte due au sous-échantillonnage. Si la fréquence de coupure est à la fréquence de Nyquist, ce rapport est $\beta = 1$ car la fréquence de coupure se trouve au bord du support de la FTO, à $\theta = 0$ rad. Cependant, le rapport β n'est pas aussi évident pour d'autres angles θ . On étudie alors en simulation l'évolution du rapport β par rapport à l'échantillonnage de la PSF α et on cherche à approximer les courbes pour obtenir une estimation de β à partir de l'échantillonnage α .

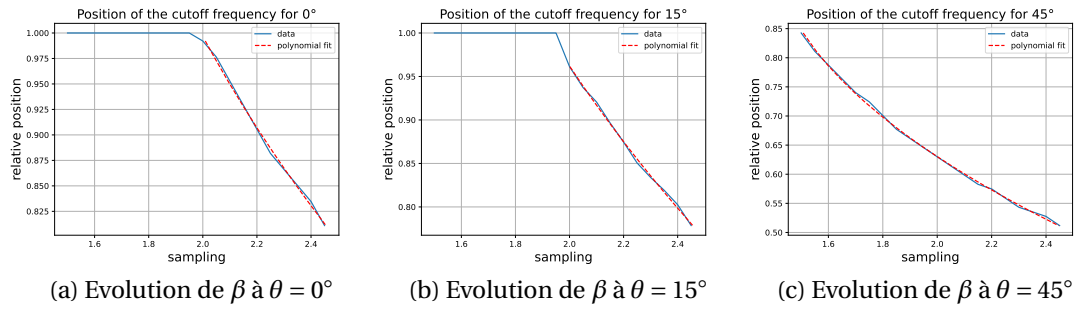


FIGURE 4.7. – Ratio de la position relative β en fonction de l'échantillonnage α pour 3 angles θ

On peut observer en figure 4.7 que pour de petits angles, un échantillonnage $\alpha < 1$ représente une fréquence de coupure qui n'est pas représentée sur le support fréquentiel : étant sous-échantillonnée, de l'information est perdue et on pose alors $\beta = 1$. Le cas échéant, l'angle en question ne sera pas utilisé pour la mesure de l'échantillonnage. La courbe d'évolution de β est ensuite ajustée sous forme d'une fonction polynomiale (courbe orange pointillée). On décrit alors un ensemble de fonctions $g_\theta(\alpha) = \beta$. En inversant chacune des fonctions g_θ , on peut alors retrouver l'échantillonnage estimé $\hat{\alpha}$ à partir du ratio β . On moyenne ensuite les différentes valeurs d'échantillonnage

pour $\theta \in [0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2}]$.

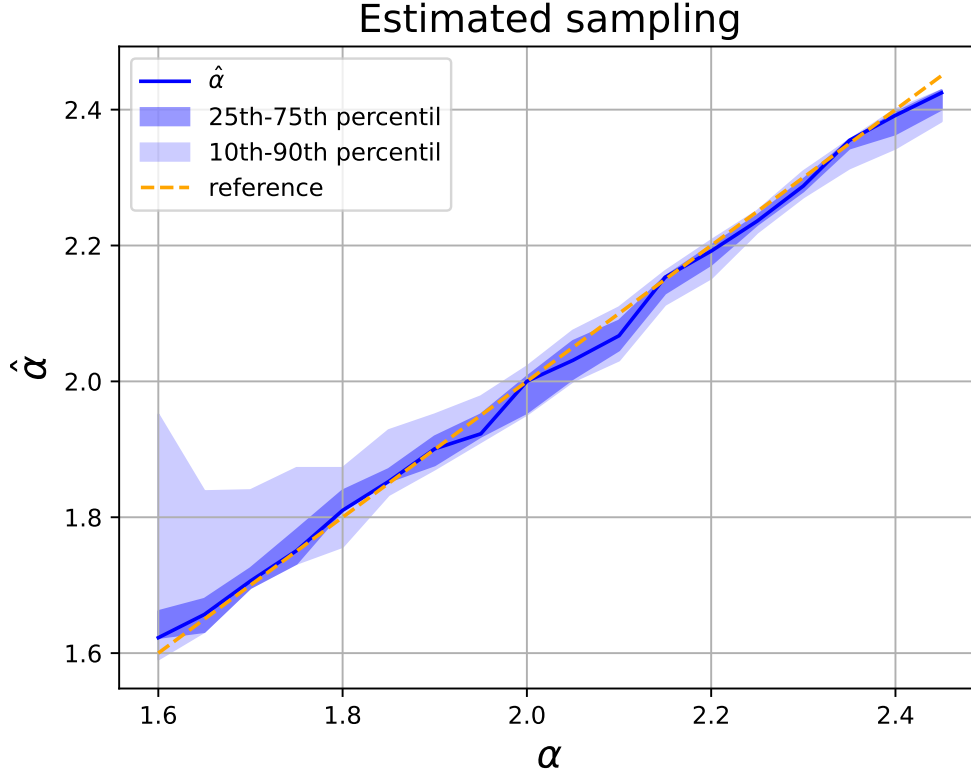


FIGURE 4.8. – Estimation de l'échantillonnage $\hat{\alpha}$.

On retrouve en figure 4.8, les résultats de l'estimation de l'échantillonnage $\hat{\alpha}$ des données simulées par rapport à l'échantillonnage réel α . L'estimation de l'échantillonnage est très précise et stable en fonction des échantillonnages significatifs pour notre étude. On repère une légère augmentation de l'erreur pour les échantillonnages relativement petits : comme étudié précédemment, le support fréquentiel de la FTO est alors trop petit pour contenir la FTO du système optique, on s'affranchit de la mesure de plusieurs angles, faussant ainsi la mesure. Néanmoins, la méthode est précise et robuste, permettant de vérifier maintenant l'échantillonnage des données expérimentales. Sur l'ensemble du jeu de données expérimentales, la valeur moyenne de l'échantillonnage des données est de 2.05 avec un écart-type de 0.041 permettant de valider un échantillonnage $\alpha_{exp} = 2$ des données expérimentales.

4.2.2. Ajustement des paramètres expérimentaux

Dans la section précédente nous avons vu le cœur de l'estimation de l'échantillonnage et son principe. Dans cette section nous allons étudier les paramètres spécifiquement expérimentaux que sont la quantification des données, le rapport signal à bruit et la réponse du pixel.

Quantification

La quantification des données est le processus de conversion de valeurs continues en valeurs discrètes pour simplifier la représentation numérique. Cela permet de réduire la précision en échange d'une représentation numérique plus compacte, mais introduit une erreur de quantification due à cette approximation elle doit donc être prise en compte. Les données sont simulées sur 32 bits alors que le détecteur de la caméra du banc optique propose une discrétisation des valeurs sur 10 bits, allant alors de 0 à 255 ADU avec un pas de 0.25 ADU soit 1024 valeurs. Une conversion doit être faite de 32 bits à 10 bits afin d'approcher au mieux en simulation les données expérimentale.

Flux et SNR

On cherche maintenant à modéliser le flux directement capté sur le détecteur : nous ne nous intéressons pas ici au flux émis par la source lumineuse, mais directement au flux reçu par le détecteur. Les données expérimentales, codées sur 10 bits, sont données en ADU et sont normalisées entre 0 et 255 avec un pas de 0.25. Afin de retrouver le flux f en photons on calcule alors

$$f = 4 \times \sum ADU \times \frac{G}{Q_E} \quad (4.3)$$

Avec G le gain en électrons par ADU et Q_E , l'efficacité quantique du détecteur. Les caractéristiques du détecteur indiquent un gain $G = 0.22 \text{ e}^-/\text{ADU}$ et une efficacité quantique $Q_E = 70\% \text{ e}^-/\text{photon}$. Le facteur 4 du calcul est dû au fait que la normalisation de la PSF soit faite entre $[0,255]$, ne représentant pas des ADU entiers. De plus, le bruit de lecture mentionné par la datasheet du détecteur est estimé à $\sigma_{\text{ron}} = 11.73 \text{ e}^-$. On peut alors remonter directement au SNR défini par l'équation 1.25. L'application numérique donne alors un $SNR_{\text{exp}} \approx 90$. C'est un régime de SNR qui a montré de bonnes performances en section 3.4 avec des fronts d'onde résiduels de quelques nanomètres. C'est ce SNR qui est pris en compte lors de la simulation des données.

Réponse et taille du pixel

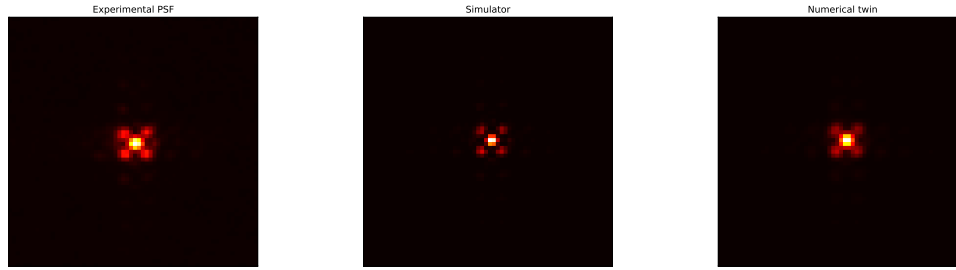
Lors de la simulation numérique des PSFs, ces dernières sont représentées dans un tableau correspondant aux pixels du détecteur, l'image du détecteur correspond exactement au module carré du champ électrique en plan focal. Cependant lors de l'acquisition de la lumière par un détecteur, une étape supplémentaire est effectuée. En effet le signal renvoyé par chaque pixel de taille p du détecteur correspond à l'intégrale de l'intensité du champ électrique sur la surface sensible du pixel. Cette opération se modélise par une opération de convolution par une fonction porte.

$$I_{\text{détecteur}} = h(P, \Phi) \otimes \Pi_p \quad (4.4)$$

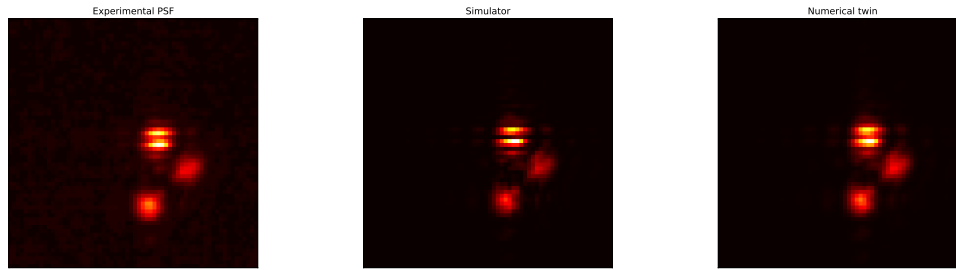
ou $I_{\text{détecteur}}$ correspond à l'image sur le détecteur, $h(P, \Phi)$ l'intensité au niveau du détecteur en plan focal et Π_p la fonction porte de largeur p , la taille du pixel. Cette

4. Validation expérimentale – 4.2. Validation des données

opération revient à flouter l'image des PSFs numériquement générées. Une fois ces paramètres pris en comptes, on peut alors simuler des données plus semblables aux données issus du banc, voir .



(a) Exemple de PSFs de régime de phasage.



(b) Exemple de PSFs de régime de déploiement

FIGURE 4.9. – Comparaison des PSFs expérimentales (gauche), issues du simulateur (milieu), et jumeaux numérique (droite) pour les régimes de phasage fin (haut) de déploiement (bas).

La figure 4.9 présente les PSFs expérimentales et simulées, on peut remarquer une nette similarité entre les deux PSFs, néanmoins des erreurs de modèles subsistent pouvant atteindre jusqu'à 28% d'erreur. Cependant, nous verrons que ces erreurs de modèle ne sont pas contraignantes par la suite. L'objectif de cette section était de générer numériquement des données qui se rapprochent des données expérimentales, dans le but de les modéliser de manière précise à l'aide d'un modèle physique. En examinant attentivement les aspects d'échantillonnage, du bruit de lecture et de photon, de la quantification des données et de la réponse des pixels, nous avons pu valider que notre modèle de formation d'image s'approche des caractéristiques d'acquisition des données expérimentales. Cette validation permet ainsi d'obtenir une simulation numérique des données plus semblable aux données expérimentales. Dans un second temps, nous allons donc exploiter ces données pour notre méthode d'identification des paramètres de la phase.

4.3. Application aux réseaux de neurones

Dans cette section, nous allons utiliser les réseaux neuronaux développés dans le chapitre 3, et les entraîner avec des données expérimentales et des données simulées numériquement. Dans un premier temps, nous évaluerons les réseaux uniquement sur les données simulées issues du jumeau numérique, puis exclusivement sur les données expérimentales pour valider notre méthode sur des données réelles. Enfin, nous étudierons un scénario plus réaliste, où le nombre de données réelles est limité et insuffisant pour constituer un jeu de données complet pour l'entraînement des réseaux. Dans ce cas, il est envisageable d'utiliser quelques données expérimentales pour parfaire les paramètres de nos réseaux et les conformer aux cas réels : c'est ce que l'on appelle le *fine-tuning*.

4.3.1. Validation simulée et expérimentale

L'une des différences majeures entre le simulateur de PSFs utilisé pour la validation sur point source et les données simulées conformes au banc est la quantification effectuée. En effet, les valeurs des PSFs simulées à partir du jumeau numérique ne peuvent prendre que 1024 valeurs contre 4×10^9 valeurs sur 32 bits pour les données simulées. La précision de quantification des PSFs est donc considérablement réduite. En plus du flou appliqué représentant la réponse des pixels sur le détecteur, les PSFs sont donc différentes. Cependant, l'entraînement des réseaux de neurones suit la même méthodologie, comprenant la même fonction de perte (le WFE défini équation 1.22) et les mêmes 100 000 données d'entraînement. Le taux d'apprentissage reste également identique. Deux réseaux sont donc entraînés de manière distincte : un Réseau entraîné sur données Expérimentales (R-Exp), et un second Réseau entraîné sur données Simulées (R-Sim), suivant la même distribution de paramètres que le jeu de données expérimental. On représente alors les résultats de l'inférence des réseaux sur le jeu de données de test correspondant aux données d'entraînement de façon à ce qu'un réseau R-Sim soit testé sur des données simulées et le réseau R-Exp soit testé sur des données expérimentales.

4. Validation expérimentale – 4.3. Application aux réseaux de neurones

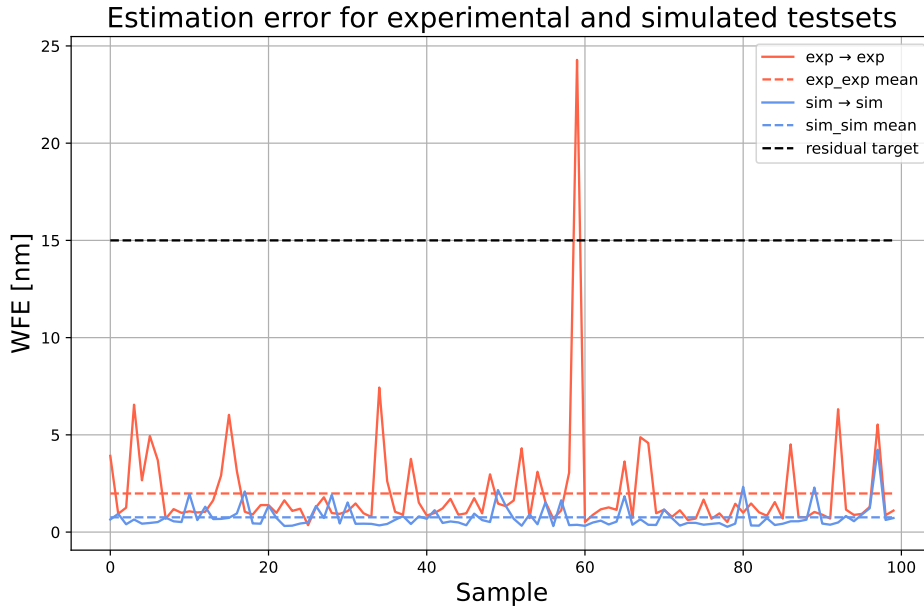


FIGURE 4.10. – Résultats de l'estimation des réseaux R-Exp et R-Sim sur les jeux de données de test de phasage fin.

La figure 4.10 montre les résultats de l'estimation des réseaux R-Exp (rouge) et R-Sim (bleu) testé sur des données expérimentales et simulées respectivement sur un jeu de 100 PSFs de test. Dans un premier temps les réseaux fonctionnent en moyenne très bien sur l'estimation des paramètres. On remarque une légère différence de qualité d'estimation entre les réseaux. En effet, le réseau R-Sim propose une erreur moindre (environ 0.8 nm en moyenne) comparée au réseau R-Exp (environ 2 nm de moyenne). Les données simulées étant fidèles au modèle de formation d'image utilisé, le réseau en apprend les caractéristiques lors de l'apprentissage et infère mieux les paramètres par la suite. Le modèle expérimental est, lui, soumis aux conditions d'acquisition des données qui peuvent être plus variables que notre modèle analytique, d'où la différence entre les deux. La PSF numéro 59 présente un front d'onde résiduel conséquent, l'erreur étant répartie sur l'ensemble des coefficients aucune conclusion n'a expliqué ce résultat.

4. Validation expérimentale – 4.3. Application aux réseaux de neurones

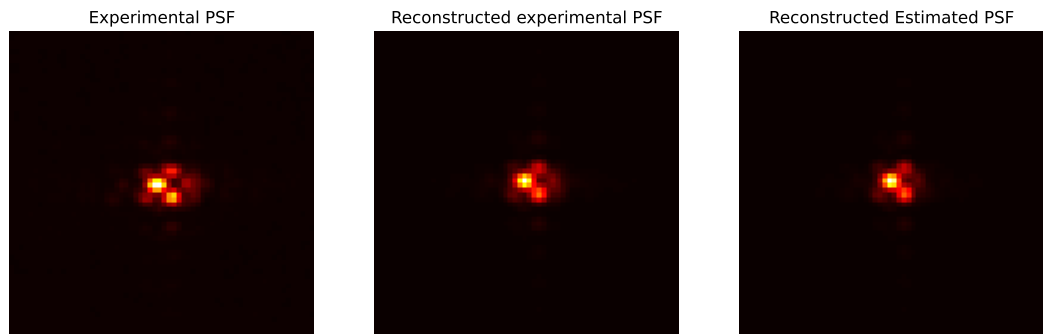


FIGURE 4.11. – PSF expérimentale (gauche), reconstruite avec les bons coefficients (milieu), reconstruite à partir des coefficients estimés (gauche).

On peut retrouver en figure 4.11, la reconstruction des PSFs réelles et estimées. On remarque que en dépit du biais entre simulateur et données expérimentales, l'estimation des réseaux est très cohérente et permet une identification fiable des erreurs de phasage.

Conclusion

Notre approche est validée sur des données expérimentales, permettant le phasage fin du télescope à partir de données plus réalistes.

Dans un second temps, on peut s'intéresser au phasage grossier du télescope, correspondant à des erreurs de segmentation plus conséquentes. Pour rappel, dans ce régime, on cherche uniquement à mesurer le tip et le tilt de chacun des segments, la PSF n'étant pas sensible aux erreurs de piston lorsque les sous-PSFs sont séparées dans le plan focal (Partie 3.3).

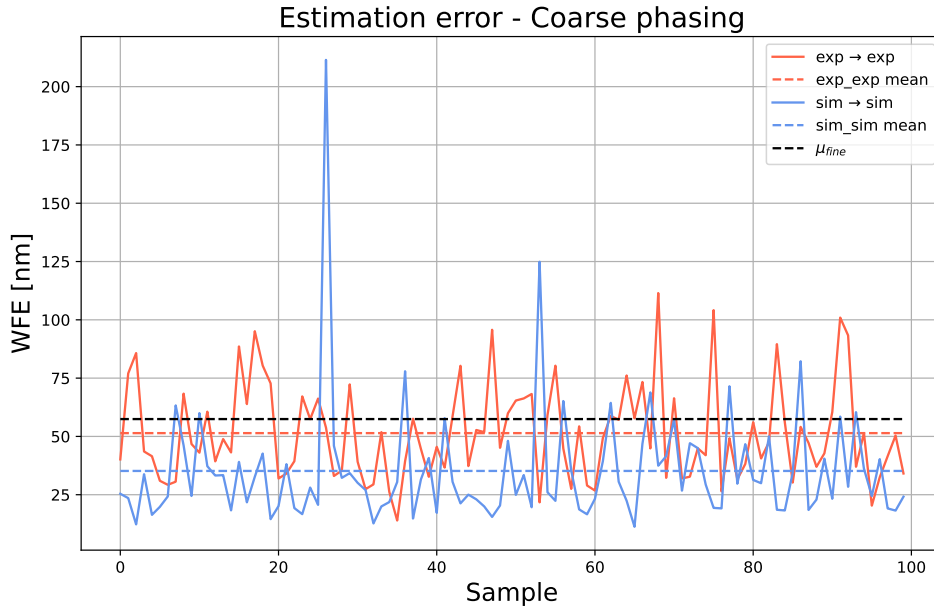


FIGURE 4.12. – Résultats de l'estimation des réseaux R-Exp et R-Sim sur les jeux de données de test du régime de déploiement.

Une nouvelle fois, les performances pour les deux jeux de données sont encourageant. Ils atteignent en effet un front d'onde résiduel moyen correspondant à la moyenne de la distribution des WFE du phasage fin, permettant de faire le lien entre les réseaux de phasage grossier et les réseaux de phasage fin sur des données expérimentales. Néanmoins une nouvelle valeur aberrante apparaît pour le réseau R-sim, repartie sur les tips et tilts de la PSF.

Conclusion

Néanmoins, notre approche est validée pour un phasage complet du télescope pour des données expérimentales à partir du moment où un jeu de données conséquent est disponible.

4.3.2. Performance des R-Sim appliqué aux données expérimentale

Le nombre de données réelles disponibles dans les cas d'application peut être très limité, et il est probable qu'elles ne soient pas suffisantes pour effectuer un entraînement complet de réseau de neurones. Nous proposons donc de tester le réseau R-Sim avec des données expérimentales afin d'évaluer la généralisation d'un type de données à l'autre.

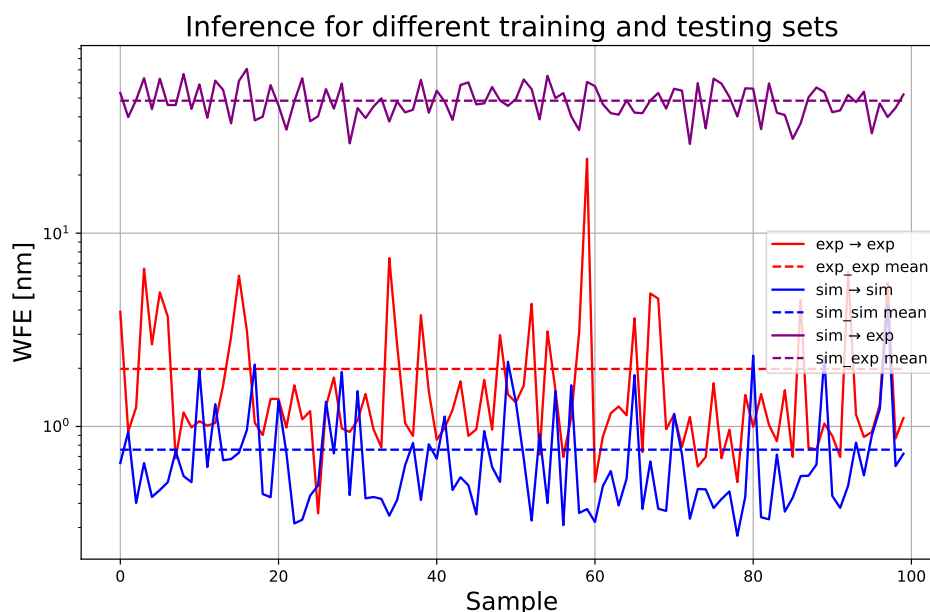


FIGURE 4.13. – Comparaison des estimations pour données simulées et expérimentales pour 100 PSFs.

La figure 4.13 compare cette fois ci les estimations effectuées sur le même type de donnée que lors de l’entraînement avec les performances du réseau entraîné sur données simulées et testé sur les données expérimentales (courbe violette). La performance affichée est clairement insuffisante.

Conclusion

L’écart entre le modèle de formation d’image et les données réelles est encore trop conséquent pour que le réseau puisse généraliser à partir des données simulées sur les données réelles.

Nous proposons donc d’améliorer nos méthodes afin qu’elles puissent convenir au mieux au phasage précis sur des données expérimentales.

4.3.3. Fine-tuning

Le modèle de formation d’image étant actualisé pour correspondre aux conditions du banc optique, il n’est néanmoins pas encore conforme aux données expérimentales. L’application directe du réseau R-Sim aux données expérimentales ne produit pas de résultats satisfaisants (proche de 100% d’erreur de front d’onde). Cependant, des techniques de *fine-tuning* permettent d’améliorer les réseaux déjà entraînés pour les adapter aux données réelles ou expérimentales. Cette approche, dite d’apprentissage par transfert, consiste à ajuster les paramètres d’un modèle pré-entraîné avec de nouvelles données lors d’un second entraînement, moins conséquent que le premier. Le

4. Validation expérimentale – 4.3. Application aux réseaux de neurones

fine-tuning peut être effectué de manière simple en modifiant légèrement l'ensemble des filtres, poids et biais du réseau. Cependant, une approche plus sophistiquée permet de profiter pleinement de l'entraînement effectué sur les données simulées. Les couches convolutives des réseaux de neurones servent principalement d'extracteurs de caractéristiques sur une image, et les filtres associés peuvent être fixés pendant le *fine-tuning*; les couches entièrement connectées effectuent la régression à partir des caractéristiques extraites par les couche convolutives. En effet, la différence entre le jumeau numérique du banc et les données expérimentales se situe majoritairement au niveau de l'intensité des pixels plutôt que sur l'influence et la localisation des erreurs de PTT au niveau de la PSF. Les résultats sur données simulées montrent que le réseau a appris à extraire ces informations à l'aide des couches convolutives, mais, une fois appliqué aux données expérimentales, ne les interprète pas correctement dans les couches entièrement connectées. Ce sont donc ces trois dernières couches entièrement connectées que nous voulons ré-optimiser pour améliorer la performance sur des données expérimentales, pendant que les filtres des couches de convolution, eux, seront gelés.

Lors de ce second apprentissage, 100 époques seront nécessaires à un learning rate initial de 5.10^{-5} puis qui décroît à la 50 ème époques jusqu'à 1.10^{-7} . La validation du réseau est effectuée sur des données expérimentales non vues pendant le second apprentissage du *fine-tuning*.

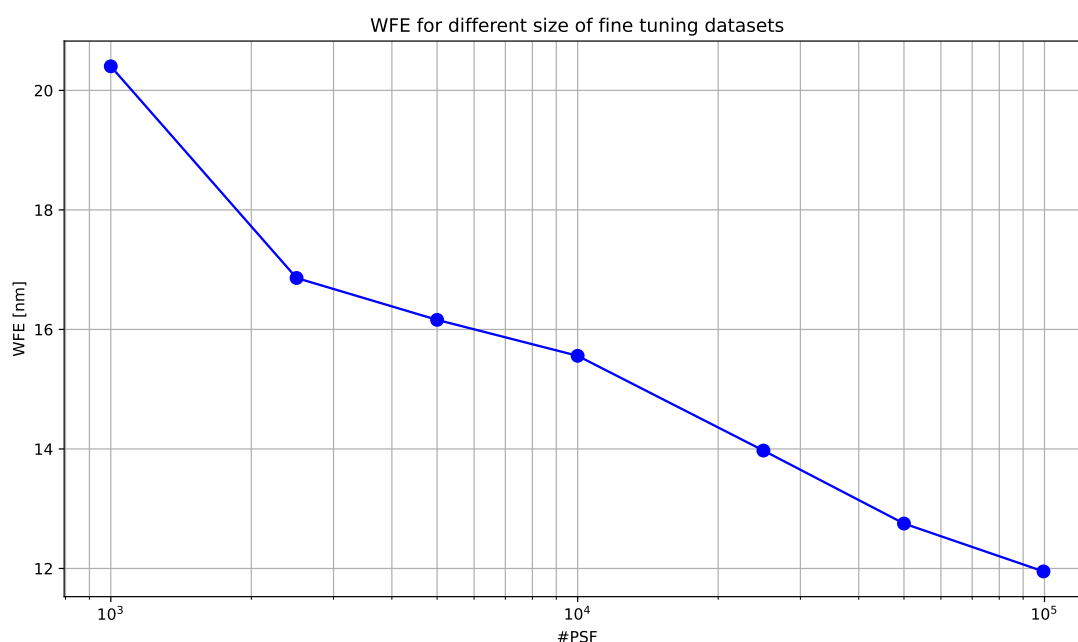


FIGURE 4.14. – Comparaison de l'erreur d'estimation des réseaux en fonction du nombre de donnée utilisée pour le *fine-tuning*.

La figure 4.14 présenter la valeur moyenne de la validation à la dernière époque d'entraînement en fonction du nombre de PSFs utilisée lors du *fine-tuning*. Finalement, le

4. Validation expérimentale – 4.3. Application aux réseaux de neurones

réseau entraîné sur des données expérimentales a optimisé le poids de ses couches entièrement connectées afin de s'adapter au biais entre données expérimentales et données réelles. Évidemment, la performance reste moindre par rapport au réseau testé sur leur propre jeu de données, loin de 2 nm des réseaux R-Exp. Néanmoins, le *fine-tuning* offre une perspective d'application des réseaux de neurones allant de la simulation à la réalité. On souhaite maintenant améliorer notre méthode. Le résultat ayant amélioré la performance des réseaux n'est néanmoins pas optimisé, on propose donc de d'utiliser 1000, 2000 et 5000 données expérimentales pour le *fine-tuning* et d'étudier l'impact des couches convolutives sur la qualité d'estimation. Pour ce faire, nous proposons donc cette fois-ci d'ajuster les poids des couches convolutives en plus des couches entièrement connectées pendant le second entraînement. Pour cela, les poids de certaines couches convolutives vont être "libérés" afin de pouvoir être ajustés à chaque époque. C'est ce que l'on appellera par la suite la libération des couches convolutives.

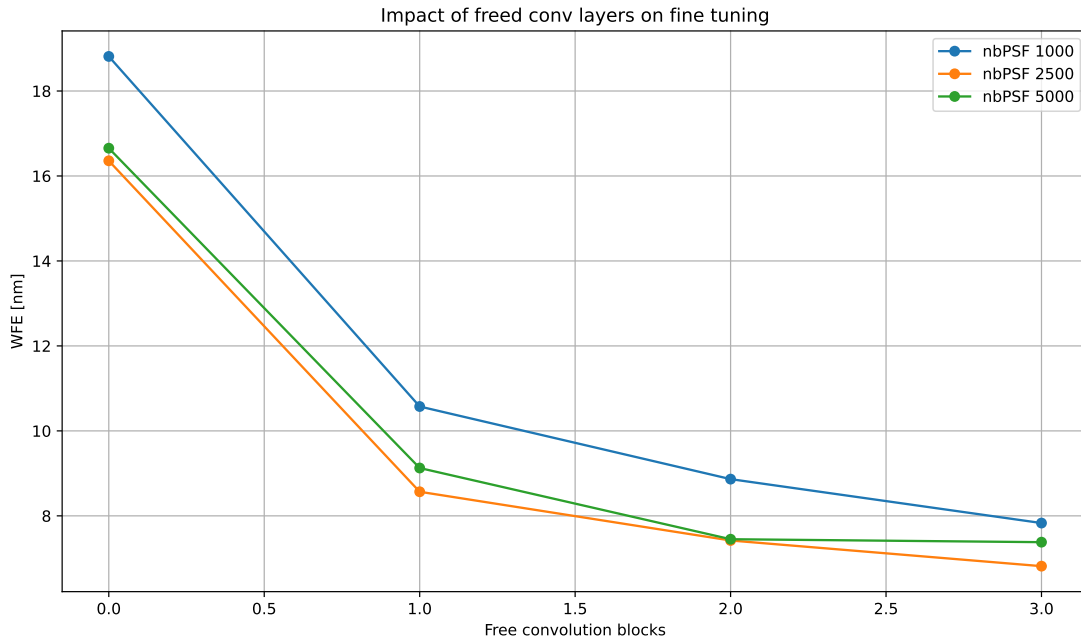


FIGURE 4.15. – Comparaison de l'erreur d'estimation des réseaux en fonction du nombre de couches convolutives libérées.

La figure 4.15 présente la valeur de la moyenne de la validation en fonction du nombre de couche convolutives libérées durant le *fine-tuning*. Finalement, on observe que l'erreur d'estimation descend au fur et à mesure que l'on libère les couches convolutives. L'écart entre données simulées et réelles était donc trop grand et les couches de sélection de caractéristiques doivent aussi être améliorées par le *fine-tuning*. On peut observer aussi que pour un second apprentissage avec 5000 PSFs, le WFE est sensiblement plus élevée par rapport au WFE correspondant à 2500 PSFs. Cet écart minime est due au calcul de la moyenne des losses prise sur un plus grand

nombre de données, néanmoins l’optimisation des hyperparametres lors de l’entraînement n’a pas présenté d’amélioration nette. A la suite de ces résultats, l’ensemble des paramètres (couches convolutives et entièrement connectées) du réseau seront donc réajustés.

4.4. Jeu de données optimal

Nous proposons maintenant d’étudier le nombre de données expérimentales optimal nécessaire pour atteindre les performances attendues pour le phasage d’AZIMOV. Pour ce faire, nous allons suivre la même méthodologie de *fine-tuning*, en variant le nombre de PSFs expérimentales de ré-entraînement afin d’estimer la quantité minimale de ces données nécessaires.

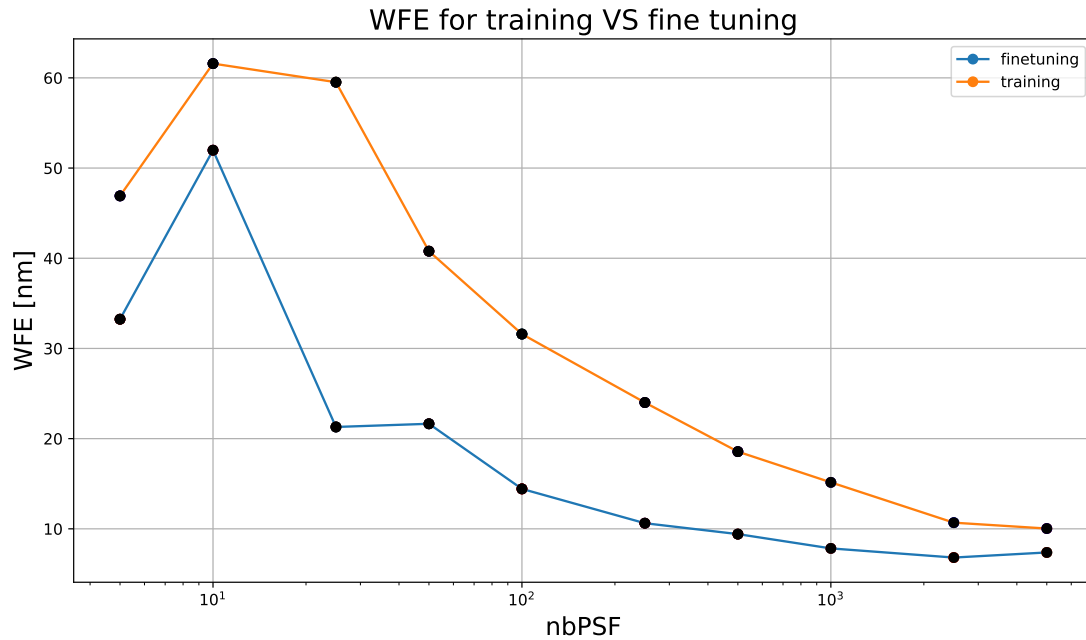


FIGURE 4.16. – Évolution de l’erreur d’estimation en fonction du nombre de données de fine-tuning.

La figure 4.16 montre l’évolution de la performance de l’estimation du réseau de neurones en fonction du nombre de données utilisées pour le *fine-tuning*. En bleu, on retrouve l’erreur d’estimation pour un réseau ré-entraîné en fonction du nombre de données de *fine-tuning*. En orange, on compare à un réseau entraîné une seule fois, uniquement sur le nombre de données correspondantes, montrant donc l’importance du premier entraînement sur données simulées. Le *fine-tuning* affiche très rapidement d’excellentes performances, permettant avec de très petits jeux de données expérimentales d’améliorer l’estimation de coefficients à la limite de diffraction requise, tout en généralisant sur les conditions d’observation des données expérimentales. Néanmoins on repère une variation de la tendance des courbes dans le cas de

4. Validation expérimentale – 4.4. Jeu de données optimal

peu de données d'apprentissage : ces variations sont principalement dues au tirage qui est fait sur les PSFs utilisées pour l'entraînement.

Conclusion

Idéalement, 2500 données expérimentales montrent le plus grand intérêt avec un front d'onde résiduel inférieur à 7 nm tout en gardant un nombre de données relativement petit. La limite de diffraction est atteinte pour 100 données expérimentales.

4. Validation expérimentale – 4.4. Jeu de données optimal

Dans cette partie, nous avons étudié notre méthode en utilisant des données expérimentales issues d'un banc optique développé à UK-ATC. Après une analyse des données et la création d'un jumeau numérique de ces dernières, nous avons examiné la réponse des réseaux de neurones à ces nouvelles données. Étant une méthode basée uniquement sur les données, l'approche par réseau de neurones est donc extrêmement sensible au jeu de données sur lequel elle est testée et montre une nette limitation entre les données expérimentales et les données simulées. Cependant, une première validation encourageante monte l'efficacité de notre méthode avec le réseaux R-Exp qui valide l'apprentissage sur le jeu de données de test purement expérimental. De plus, les réseaux de neurones R-Sim ont montré une excellente généralisation aux données expérimentales en effectuant un second entraînement sur un jeu bien moins conséquent de données expérimentales, réapprenant ainsi correctement l'identification des paramètres de la phase grâce au *fine-tuning*.

5. Etude de l'objet étendu

Dans ce chapitre, nous allons étudier l'analyse de front d'onde sur objets étendus par réseaux de neurones. Lors de l'observation de la Terre depuis l'espace, les objets étudiés présentent une grande variété et leurs caractéristiques peuvent différer considérablement. La complexité de la formation d'image, due à la convolution de l'objet par la PSF du système optique, ajoute une difficulté supplémentaire au problème inverse d'estimation de phase. De plus, comme pour l'étude des sources ponctuelles, les données disponibles réelles sont limitées et difficilement répétables, ce qui empêche la constitution de jeux de données conséquents pour notre méthodologie. Enfin, des limitations liées aux caractéristiques des objets seront également examinées.

Ce chapitre se décompose en cinq parties. La première présente un rappel sur la formation d'image pour un objet étendu, dans la seconde partie nous proposons une stratégie de phasage par Deep Learning. Ensuite, la troisième partie aborde l'estimation de la phase et les résultats de cette méthodologie. Enfin, les quatrième et cinquième parties traitent respectivement des limitations de l'analyse du front d'onde sur un objet étendu et du phasage en boucle fermée.

Sommaire

5.1. Contexte	161
5.1.1. Observation de la Terre en orbite basse	161
5.1.2. Base et limite de la formation d'image sur objet étendu	163
5.2. Stratégie de phasage par réseaux de neurones	165
5.2.1. Simulation d'objet	165
5.2.2. Stratégie naïve d'estimation : méthode directe	169
5.2.3. Renforcement de la stratégie d'apprentissage : PINN	173
5.3. Estimation de la phase	174
5.3.1. Vérification de la méthode sur point source	174
5.3.2. Application aux objets simulés	176
5.3.3. Diversité d'objet connu	176
5.4. Limitation de l'analyse de front d'onde	179
5.4.1. Contraste de l'objet	179
5.4.2. Sensibilité de la phase au contraste	182
5.5. Phasage en boucle fermée	183

5.1. Contexte

5.1.1. Observation de la Terre en orbite basse

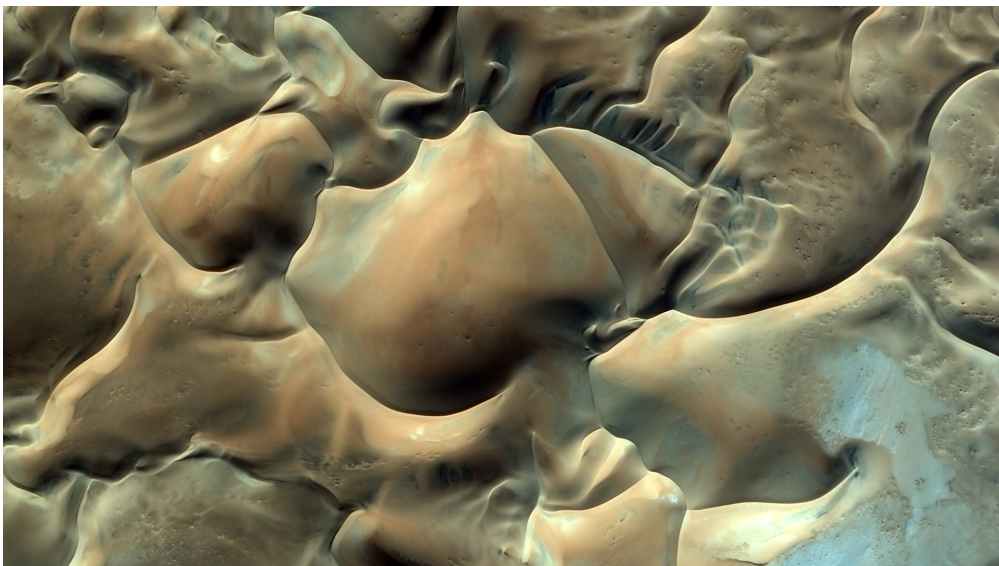
Les télescopes en orbite basse jouent un rôle dans la surveillance de notre planète. Placés à une altitude de 200 à 2 000 kilomètres, ces satellites circulent autour de la Terre à des vitesses impressionnantes d'environ 28000 km/h ce qui leur permet de compléter une orbite en seulement 90 à 120 minutes. Leur objectif principal est de collecter des données sur divers aspects de la Terre, allant de l'observation météorologique à la cartographie, en passant par la surveillance environnementale et les applications militaires. La largeur de la fauchée, c'est-à-dire la zone observée à la surface terrestre lors de chaque passage, peut varier considérablement, permettant ainsi une couverture détaillée et régulière de vastes régions. Actuellement, plusieurs centaines de satellites sont dédiés à l'observation de la Terre, chacun équipé de technologies de détecteurs optiques, des radars, ou encore des spectromètres. Ces télescopes en orbite sont essentiels pour l'analyse des changements climatiques, la gestion des ressources naturelles, la prévention des catastrophes naturelles, et même pour des applications économiques, offrant ainsi un aperçu inestimable de notre monde en constante évolution.

AZIMOV propose une solution innovante pour améliorer l'observation de la Terre depuis l'espace. Ce système est conçu pour être déployé à une altitude de 500 km, ce qui permet aux satellites de compléter une orbite autour de la Terre en environ 90 minutes. Grâce à son concept de télescope déployable, AZIMOV vise à obtenir une résolution au sol (GSD) d' 1 m , offrant des images à très haute résolution du sol. La fauchée de ces satellites, est de 5 km, permettant ainsi une couverture large et détaillée de la surface terrestre. On peut observer des exemples ce type d'images, acquises à Nadir par le télescope Pléiades :

5. Etude de l'objet étendu – 5.1. Contexte



(a) Image des White Sundays en Australie.



(b) Image du désert d'Algérie



(c) Image de la zone aeroportuaire de Tucson aux USA.

FIGURE 5.1. – Exemple d'images acquises en orbite basse.

On peut observer une diversité de paysages. En figure 5.1, on retrouve trois exemples de scènes observables : la côte, un désert et une zone aéroportuaire. D'autres paysages sont aussi observables et présentent différentes caractéristiques : des forêts, océans, villes, montagnes, glaciers. Les images présentées ici ont été traitées et déconvoluées pour en améliorer la qualité, de façon à représenter au mieux les objets étudiés. Dans le cadre de cette thèse, le terme "objet" désignera les petites sections extraites à partir des images montrées en figure 5.1 et d'autres images des bases de données. Ces objets seront ensuite convolués avec la PSF pour générer les "images" captées par le détecteur du télescope AZIMOV

5.1.2. Base et limite de la formation d'image sur objet étendu

On propose dans cette section de rappeler le formalisme et le processus de formation d'image d'un objet étendu. La formation d'image sur objet étendu décrite en partie 1.5 présente l'image dans le plan-focal comme la convolution de la PSF par l'objet étudié, ce qui représente donc la multiplication fréquence à fréquence de la FTO (vue en partie 4.2.1) par le contenu fréquentiel de l'objet.

$$I(\vec{r}) = \{o \otimes PSF\}(\vec{r}) \quad (5.1)$$

Prenons l'exemple d'un objet simple composé d'un sinus horizontal haute fréquence et d'un sinus vertical moyenne fréquence. Les fréquences spatiales horizontales et verticales sont choisies spécifiquement afin de permettre une interprétation des images.

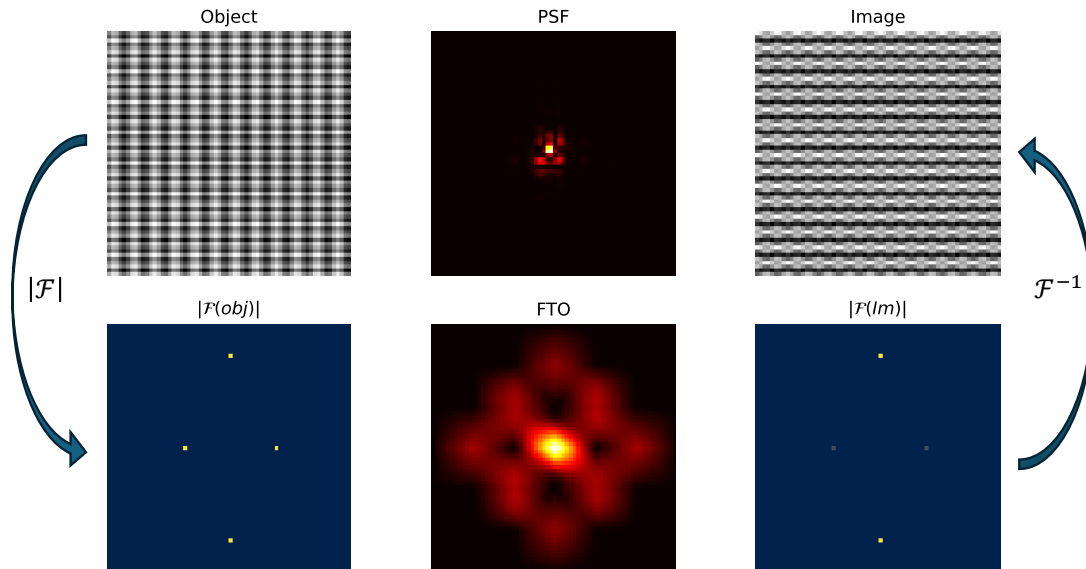


FIGURE 5.2. – Exemple de formation d'image d'un objet étendu.

On peut observer en figure 5.2 l'objet composé de deux sinusoïdes de fréquences dif-

férentes orientées horizontalement et verticalement. Fréquemment, ils sont donc représentés par des Dirac dans le domaine de Fourier où leur distance par rapport au centre de l'image caractérise la fréquence spatiale des sinusoïdes. L'orientation relative des Dirac correspond à l'orientation des sinusoïdes. Les deux Dirac les plus proches du centre représentent donc les sinusoïdes évoluant horizontalement dans l'objet, tandis que le sinus de plus haute fréquence est représenté par les Dirac plus proches du bord du domaine fréquentiel. La seconde colonne représente notre système optique : sur la première ligne, la PSF que nous avons étudiée tout au long du manuscrit, tandis que sur la seconde, nous pouvons observer la FTO. Cette FTO, s'annulant en certains points du domaine fréquentiel en raison de la forme de notre pupille et des aberrations de la phase, vient donc minimiser les fréquences spatiales de notre objet de façon plus ou moins prononcée selon la direction et l'amplitude de ces dernières. Dans l'image ainsi formée on retrouve alors majoritairement le sinus vertical de haute fréquence (car la FTO comporte des amplitudes fréquentielles élevées à cette fréquence et dans cette direction) alors que le sinus horizontal se retrouve presque annulé (très peu visible sur l'image). Le spectre de l'image ainsi formée $|\mathcal{F}(Im)|$ témoigne de l'atténuation du sinus horizontal et de la conservation du sinus vertical. On peut retrouver cet effet en simulant des lignes verticales et horizontales plus franches.

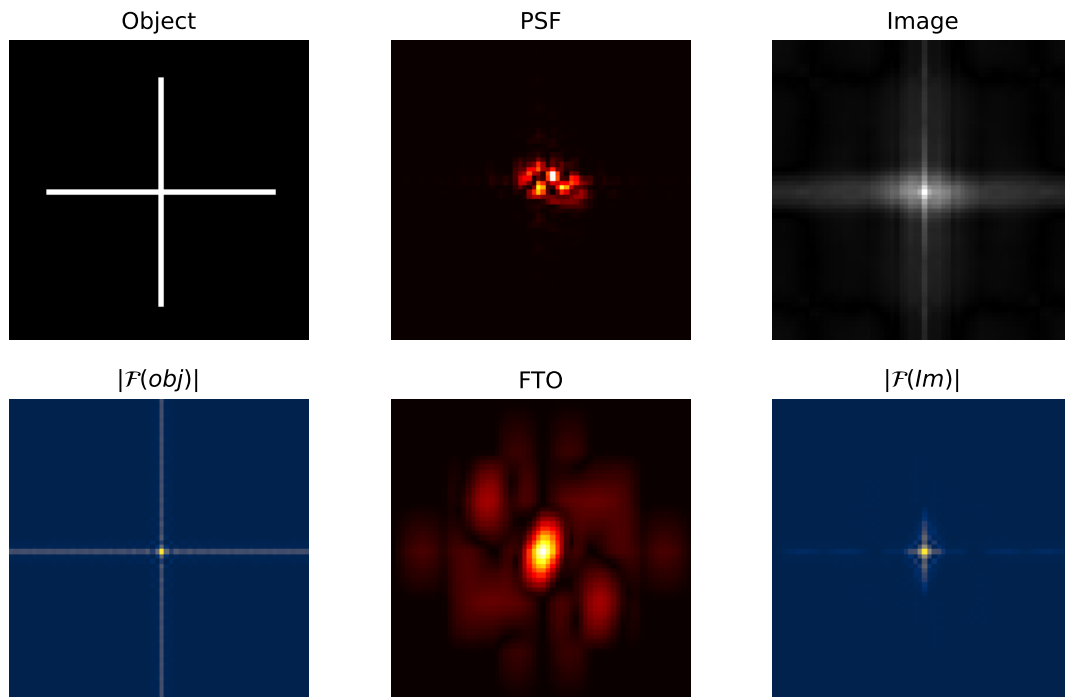


FIGURE 5.3. – Exemple de formation d'image de lignes.

Ici, les fréquences spatiales de la ligne verticale sont correctement retrouvées alors que celles de la ligne horizontale sont en partie atténuées. Cette annulation est alors dépendante de la PSF et de sa FTO ; une forme moins aberrée de la PSF aurait conduit

à un objet mieux imagé au plan-focal.

En somme ce cas simple permet de généraliser lors de la formation d'image d'objet réaliste : certaines fréquences spatiales de l'objet sont annulées par la FTO, et ne sont donc plus présentes dans l'image. Celle-ci peut dans ce cas, ne plus contenir aucune information sur ces fréquences spatiales. En effet, lors de l'analyse de front d'onde avec une seule image, on peut difficilement déterminer si ces fréquences spatiales faibles sont dues à la topologie de l'objet observé ou aux aberrations du système optique.

5.2. Stratégie de phasage par réseaux de neurones

Étudions maintenant le phasage à partir de l'image en plan-focal dans le cas d'objet étendu. J'étudie dans cette partie la génération d'objet réaliste observable puis la méthodologie pour le phasage par Deep Learning et ses limites. La méthodologie étant confirmée pour l'analyse d'un point source, je propose dans ce manuscrit d'étudier des objets proches d'un point source, tels qu'un objet brillant de forme approximativement circulaire centré sur l'image. Un exemple des ces objets peut être retrouvé

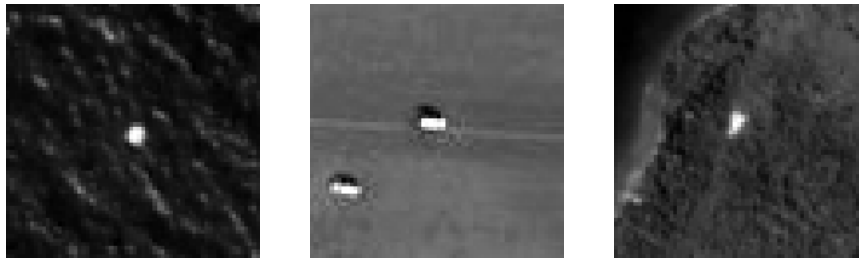


FIGURE 5.4. – Exemple d'objets observables.

5.2.1. Simulation d'objet

Nous avons donc maintenant un paramètre supplémentaire à notre problème : les scènes observées. Pour ce faire nous avons besoin d'une base de données de scènes réalistes du sol terrestre, à une résolution au moins égale à celle d'AZIMOV. Ces bases de données d'images sont relativement rares. Je propose donc 3 méthodes pour générer ces objets dans les trois prochains paragraphes.

Objet étendu non bruité gaussien

Dans un premier temps, un objet étendu simple est simulé, sur fond nul et de forme gaussienne avec des variations d'intensités au sein de l'objet. Cela permet d'explorer un cas nominal relativement simple pour attester de la faisabilité de l'analyse du front

5. Etude de l'objet étendu – 5.2. Stratégie de phasage par réseaux de neurones

d'onde. Ces objets étendus non bruités gaussiens, sont donc générés à partir d'un seuillage d'une distribution normale généralisée en 2 dimensions :

$$\mathcal{N} = \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)} e^{\left(-\left(\frac{|x-\mu|}{\alpha_x}\right)^\beta - \left(\frac{|y-\mu|}{\alpha_y}\right)^\beta\right)} \quad (5.2)$$

où Γ est la fonction gamma, $\mu \in \mathbb{R}$ est un paramètre de position, $\alpha_x, \alpha_y \in]0, \infty[^2$ sont des paramètres d'échelle et $\beta \in]0, \infty[$ est un paramètre de forme. Les valeurs des coefficients de la gaussienne sont tirées aléatoirement selon un tirage uniforme. La gaussienne est ensuite seuillée, puis un bruit de Poisson est appliqué une fois son amplitude multipliée par 200 afin de générer une non uniformité de l'amplitude de l'objet. Les valeurs trop petites sont ensuite imposées à 0 et notre objet étendu simulé est ensuite normalisé par la somme de ses intensités. Un exemple des objets générés est présenté ci-dessous.

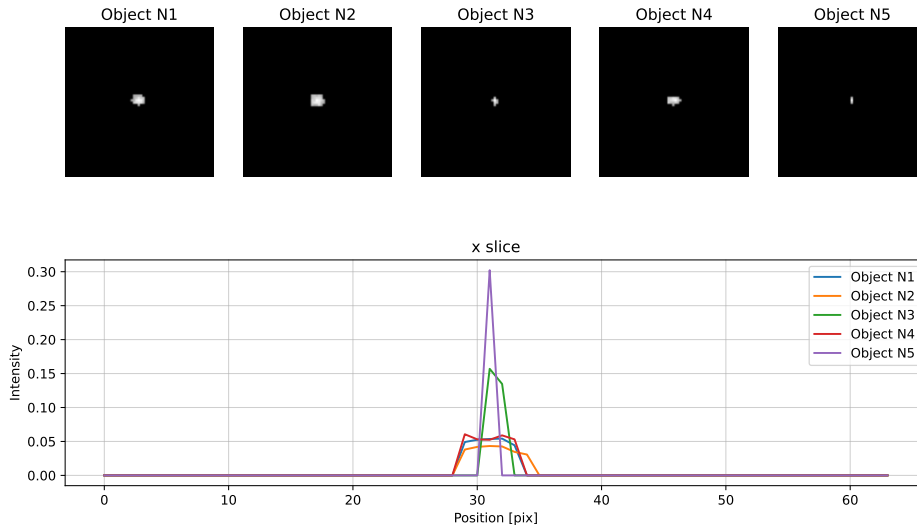


FIGURE 5.5. – Exemple d'objets étendus simulés.

Ces objets générés permettent l'application de notre méthode sur un modèle très simpliste. Dans le contexte de l'observation de la Terre, ces objets idéalisés n'existent pas mais leur forme est similaire. Néanmoins les objets réels sont nécessairement bruités par un fond non uniforme ou mélangés à d'autres informations spatiales de la scène étendue.

Variational Auto-encoders

La stratégie précédente consiste à générer des objets pour entraîner les NNs. Cette gé-

nération d'objet suit un modèle simple d'objet "quasi-ponctuels" mais cette stratégie est limitée. Modéliser un objet peut demander un très grand nombre de paramètres, qui nécessiteront un réseau de taille croissante. Ce dernier point sortant de notre ambition, nous choisissons une stratégie alternative consistant à étudier l'intérêt du Deep Learning pour générer lui-même des objets pertinents, en se basant sur des images réalistes du sol terrestre.

Un jeu de données a été utilisé afin d'identifier des objets étendus à partir d'observations réelles de la Terre par des satellites en orbite basse. Après un pré-traitement pour extraire les objets d'intérêt, une base de données d'environ 400 objets a été constituée. Cependant, cette base étant insuffisante pour effectuer un entraînement sur des réseaux de neurones, l'utilisation de générateurs de données plus réalistes est envisagée : comme mentionné en partie 2.3.4, les VAEs permettent de simuler des données de manière très fiable, reproduisant fidèlement les caractéristiques du jeu de données d'entraînement. Le VAE est une première approche de génération de données, d'autres approches sont possibles, mais le VAE a été choisi pour sa simplicité de mise en oeuvre. Pour rappel, le VAE est une méthode de Deep Learning permettant d'encoder des caractéristiques des données d'entrée dans un espace compact, appelé espace latent. Cet espace latent est défini par des distributions de ces caractéristiques (plutôt que des simples scalaires), permettant de générer des données non redondantes à partir de tirages de ces distributions.

Le VAE présenté figure 2.11, a permis à partir des 460 objets manuellement extraits des données, d'apprendre les caractéristiques fines des objets d'intérêt. L'espace latent est choisi arbitrairement à 8 dimensions et l'apprentissage a duré plus de 200 époques. Les coefficients pondérant la loss de reconstruction et la loss KL sont choisis tels que $k_{rec} = 0.99$ et $k_{KL} = 0.01$. C'est le ratio principalement utilisé lors de l'entraînement des VAEs, permettant de prioriser la reconstruction tout en étant suffisante pour guider la distribution latente vers une distribution gaussienne. Une fois le réseau correctement entraîné, des données peuvent ensuite être générées en effectuant des tirages aléatoire depuis l'espace latent à 8 dimensions. Afin d'assurer la qualité des images générées et d'assurer un minimum de contraste au niveau de l'objet étendu, un tri des images générées est effectuée. Il est constitué d'un seuillage autour de la somme et de l'écart type des pixels au centre de l'image est effectué. On peut observer les images générées par le VAE en figure 5.6.

5. Etude de l'objet étendu – 5.2. Stratégie de phasage par réseaux de neurones

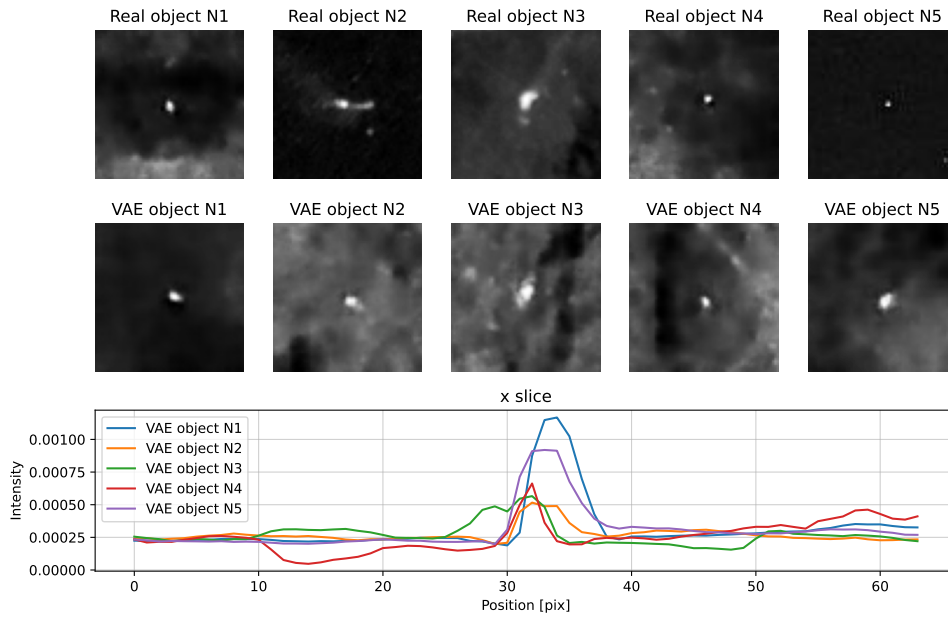


FIGURE 5.6. – Exemple d'objets étendus simulés par VAE. Les objets réels sont présentés à la première ligne tandis que ceux générés par le VAE sont présentés à la seconde. On retrouve une coupe du centre des objets à la dernière ligne.

La figure 5.6 montre un exemple de 5 objets générés par le VAE. On peut noter la différence nette de contraste entre les données générées à l'aide d'une gaussienne et les données générées par le VAE. Cette différence de contraste vient introduire une complexité supplémentaire à l'estimation de la phase, étant une manifestation plus basse fréquence qu'un objet étendu à fond nul.

Un dernier jeu de donnée vient introduire un cas moyennement contrasté : à partir des données issues du VAE, on propose d'augmenter localement le contraste autour de l'objet d'intérêt, en diminuant le fond de l'image.

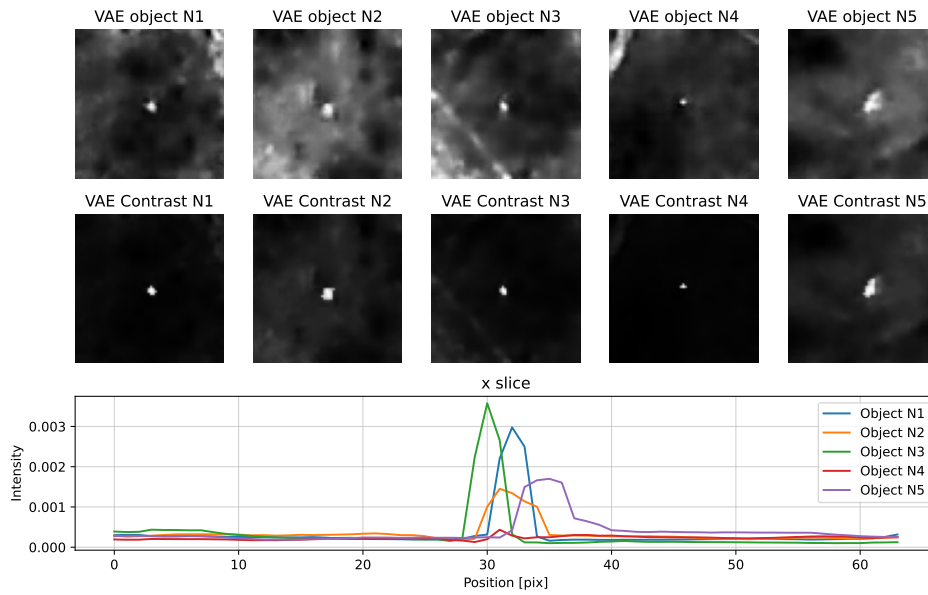


FIGURE 5.7. – Exemple d'objets étendus simulés par VAE présentés à la première ligne tandis que ceux générés par le VAE avec un contraste augmenté sont présentés à la deuxième ligne. On retrouve une coupe du centre des objets plus contrastés à la dernière ligne.

Les trois méthodes présentées ci-dessus permettent une génération illimitée et contrôlée des objets pour les jeux de données des réseaux. Voyons maintenant la stratégie adoptée pour assurer le phasage du télescope.

5.2.2. Stratégie naïve d'estimation : méthode directe

Dans cette section, nous allons étudier l'application directe de notre méthode à l'objet étendu : à partir de l'image seule du plan-focal, nous allons chercher à estimer directement les coefficients de la phase.

L'architecture utilisée pour effectuer l'inférence des paramètres de la phase est l'EfficientNet, présenté au chapitre 2, partie 2.3. En effet, les architectures utilisées précédemment sont rapidement limitées par la complexité de l'estimation : aucune des architectures Resnet présentées au chapitre 2 n'ont convergées durant leur apprentissage ; une architecture plus complète est donc choisie ici afin de modéliser plus finement la complexité du problème. L'un des points majeurs de l'architecture EfficientNet est qu'elle reste une architecture relativement légère (autour de 20 MB contre 10 MB pour le modèle le plus complexe de notre architecture personnalisée) permettant des performances plus complètes que nos architectures Resnet.

Conclusion

L'EfficientNet sera l'architecture utilisée tout au long de ce chapitre.

A partir d'un unique objet

Dans un premier cas nominal nous allons générer des données à partir d'un seul et unique objet. Seules les PSFs vont varier afin de modifier l'image d'entrée des réseaux. Nous allons nous intéresser à l'entraînement des réseaux pour des objets gaussiens à fond nul et pour un objet généré par VAE. On observe alors ci-dessous les courbes d'entraînement et de validation au cours des époques d'entraînement (voir partie 2.2).

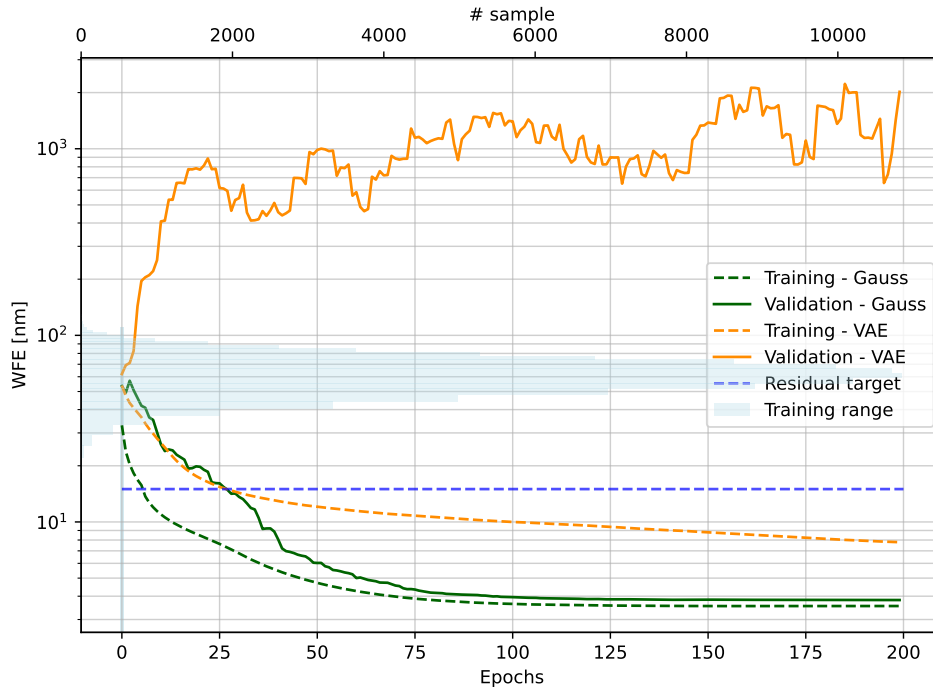


FIGURE 5.8. – Loss d'apprentissage et de validation en fonction du nombre d'époque pour un unique objet gaussien (courbes vertes) et un unique objet issu du VAE (courbes oranges), et distribution des WFE d'entrée (bleu clair).

On observe la qualité d'estimation de l'EfficientNet avec l'approche directe pour des données comportant une variété de PSF et un seul et même objet. On remarque notamment que, dans le cas d'un objet gaussien, l'apprentissage s'effectue correctement, avec une convergence des courbes d'entraînement et de validation proche d'une valeur autour de 3 nm, des performances sensiblement similaires aux réseaux NN3 et NN5 (§) étudiés pour le point source. Cependant, lorsqu'un objet issu du VAE est utilisé comme donnée d'entrée, le réseau ne parvient pas à effectuer un apprentissage convenable : la loss d'apprentissage diminue pendant que la loss de validation stagne

aléatoirement. Malgré la recherche poussée des hyperparamètres, l'apprentissage du réseau ne se généralise pas à des données non vues pendant l'entraînement. En détail, cela signifie que le réseau apprend uniquement les caractéristiques de ses données d'entraînement (convergence de la loss d'apprentissage) plutôt que les spécificités des images reliant la PSF aux erreurs de phasage (divergence de la loss de validation).

Conclusion

On observe donc déjà, dans un cas simpliste, une limitation de l'approche directe de la méthode.

L'entraînement des deux réseaux a suivi le même nombre d'époques, une optimisation des hyperparamètres, possède la même architecture, et les mêmes PSFs. La seule différence réside dans l'objet observé qui apparaît être limitant pour l'analyse du front d'onde.

Généralisation aux objets

Intéressons nous maintenant à la généralisation sur divers objet. Ainsi pour chaque donnée d'entraînement et de test, un objet différent est convolué par une PSF différente. L'objectif est alors la généralisation des réseaux sur l'ensemble des objets et des PSFs. On propose donc d'étudier le résultat de l'entraînement des réseaux.

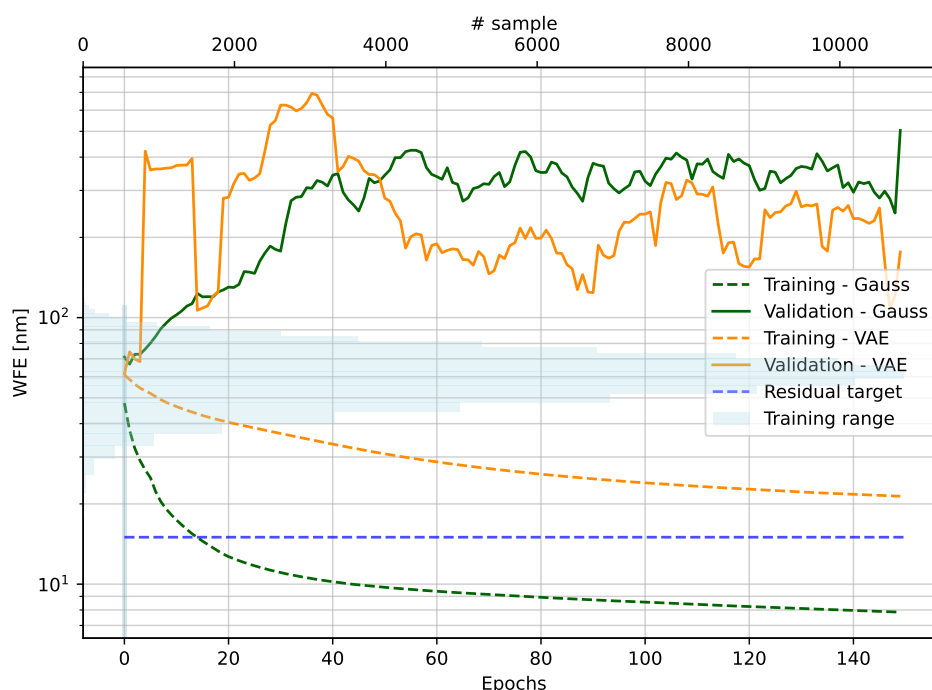


FIGURE 5.9. – Loss d'apprentissage et de validation en fonction des époques pour 100k objets différents gaussien (courbes vertes) et issus du VAE (courbes oranges).

Une réelle limitation se manifeste ici : les réseaux de neurones ne parviennent pas à généraliser sur les objets. En effet, le paragraphe précédent confirme la généralisation de l'apprentissage sur les PSFs avec un objet unique sur objet gaussien. Durant l'apprentissage du réseaux, l'ajout de diversité d'objet dans le jeu de donnée contraint le réseau à ne pas apprendre les caractéristiques de la phase contenues dans l'image, mais la distribution des données elle mêmes. Les réseaux ne parviennent donc pas à valider leur apprentissage sur les jeux de données de validation (image et donc objets qu'ils n'ont pas vus pendant l'entraînement), et l'estimation diverge en conséquence. Des tests ont aussi été réalisé sans succès avec des jeux de données plus conséquent.

Conclusion

Fort de ces exemples, la stratégie directe employée pour estimer le front d'onde directement à partir de l'image semble nettement limitée par l'ajout de complexité dû à la diversité des objets.

Une nouvelle méthode est donc mise en oeuvre pour compenser les limitations identifiées.

5.2.3. Renforcement de la stratégie d'apprentissage : PINN

La variété d'objets et de PSFs ajoute des inconnues au problème étudié. Si l'on veut étendre notre étude à la généralisation d'objets et de PSFs, le réseau doit alors apprendre à traiter des objets et des PSFs différentes pour chaque donnée d'entraînement. Nous proposons d'améliorer la stratégie d'apprentissage en ajoutant une erreur de reconstruction basée sur l'information plan-focale en plus de l'erreur de front d'onde WFE. Cette seconde partie de la loss, inspirée des travaux de Wang ([Wan+20]), que l'on appellera $loss_{phy}$, est définie par la différence quadratique entre l'image d'entrée et l'image reconstruite à partir de l'objet et de l'estimation des coefficients de la PSF. La loss totale utilisée pour l'objet étendu (*Extended Object* (EO)), $loss_{EO}$ est alors définie telle que

$$loss_{EO} = \alpha \sum_{k=1}^4 \sqrt{\sum_{i=1}^3 (c_i^k - \hat{c}_i^k)^2} + (1 - \alpha) |I - \hat{I}|^2 \quad (5.3)$$

où α est un coefficient pondérateur entre les deux loss et $\hat{I}$ est l'image reconstruite à partir des coefficients estimés $\hat{c}_i^k$. On peut trouver figure 5.10 un schéma de l'architecture du réseau de neurones durant l'apprentissage :

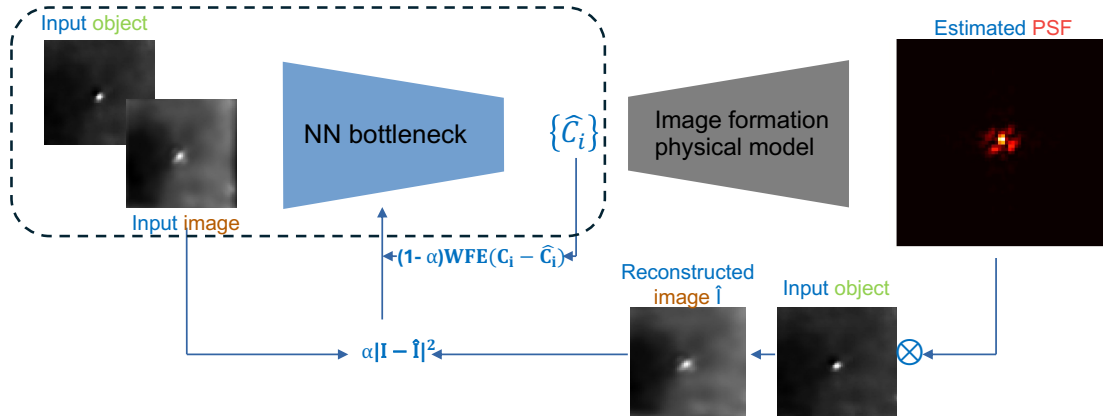


FIGURE 5.10. – Schéma de la stratégie d'apprentissage du réseau de neurones pour l'objet étendu.

On propose donc de réutiliser le modèle de formation d'image utilisé lors de la génération d'image pour reconstruire des PSFs à partir des coefficients estimés par notre réseau de neurones. Ces PSFs, ensuite convoluées par l'objet générant l'image, retournent une image estimée $\hat{I}$ qui, une fois comparée à l'image d'entrée, ajoute la contribution de la $loss_{phy}$ à la loss totale d'apprentissage. Cette stratégie permet donc d'ajouter le modèle physique de formation d'image à l'apprentissage du réseau de neurones. Le point majeur de cet ajout est qu'il permet, lors de la backpropagation, de rétropropager les gradients à travers le modèle physique et de le prendre en compte

lors de l'ajustement des paramètres internes du réseau de neurones à chaque époque. Cette méthodologie s'apparente aux réseaux de neurones informé par la physique (*Physics Informed Neural Network* (PINN), [Cuo+22]).

Une approche directe par apprentissage supervisé apprend une représentation des données à partir de leur label en minimisant $J(\theta)$:

$$J(\theta) = \operatorname{argmin}_{\theta} \|\mathcal{L}(f_{\theta}(x), \mathbf{c})\| \quad (5.4)$$

Ce qui fonctionne lorsqu'on maîtrise son jeu de données et que les données de test sont très semblables aux données d'entraînement. Cependant, dans le cas de l'objet étendu, on souhaite que l'apprentissage se fasse en partie sur le modèle physique de la formation d'image établi. On optimise donc le réseau autour de ce modèle physique en plus du front d'onde :

$$J(\theta) = \operatorname{argmin}_{\theta} \left\| \left(\alpha \mathcal{L}_{phy}(\mathcal{I}(f_{\theta}(x), O), I) + (1 - \alpha) \mathcal{L}_{WFE}(f_{\theta}(x), \mathbf{c}) \right) \right\| \quad (5.5)$$

où $\mathcal{I}$ représente le modèle de formation d'image qui prend en entrée les coefficients estimés par le réseau $f_{\theta}(x)$ et l'objet étudié O afin de générer une image estimée. Dans la partie modèle physique, l'erreur de front d'onde n'apparaît pas et un modèle non supervisé pourrait être étudié aussi. Cependant, l'ajout de la loss de front d'onde se justifie par les erreurs de λ sur les pistons et les erreurs numériques de recouvrement des tip et des tilts. De plus, cet ajout venant complexifier drastiquement la complexité calculatoire lors de l'apprentissage, il n'est pas contraignant lors du mode opératoire du réseau. En effet, la partie de la loss correspondant à la reconstruction de l'image est uniquement ajoutée lors de l'entraînement. Une fois le réseau entraîné, ce dernier effectue une inférence, et se comporte comme un réseau de neurones bottleneck classique en utilisant uniquement la partie en pointillé de la figure 5.10 en tant qu'estimateur de phase.

5.3. Estimation de la phase

Dans cette section nous allons détailler les résultats obtenus sur objet étendu par notre méthode PINN. Dans un premier temps, l'analyse de front d'onde sera supposée à objet connu (*Known Object* (KO)), avec une diversité d'objet et de PSF pour la génération d'image.

5.3.1. Vérification de la méthode sur point source

Nous souhaitons vérifier la faisabilité de la méthode sur un point source similaire au chapitre 3 permettant d'identifier directement la PSF en plan-focal. L'entrée du réseau est alors un tenseur composé d'une PSF et d'un point source dans le cas d'un objet connu. Les courbes d'entraînement convergent bien moins rapidement lorsque l'objet est connu.

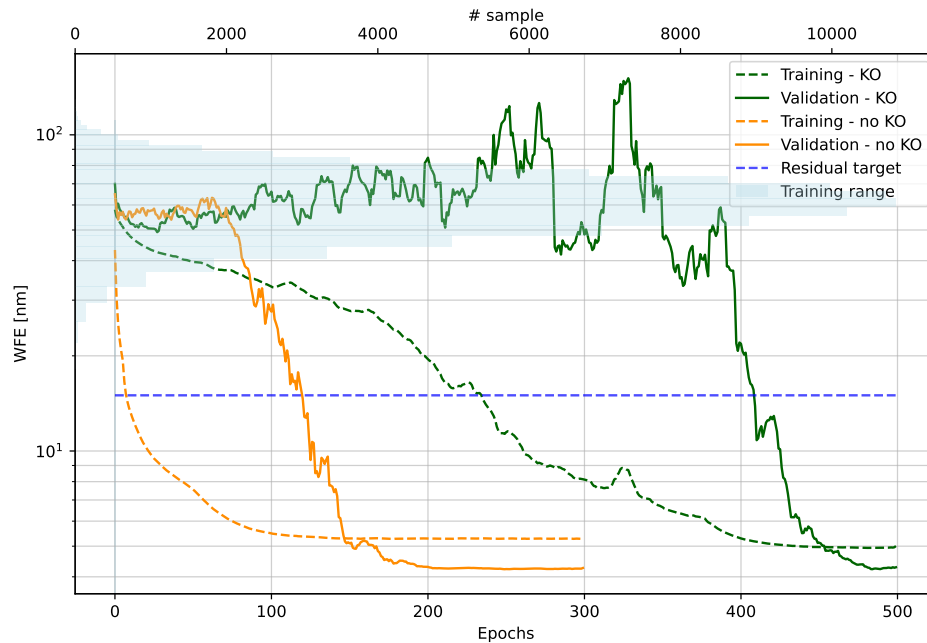


FIGURE 5.11. – Courbes d’entraînement sur point source à objet non connu (no KO) et objet connu (KO). L’apprentissage du réseau à objet non connu s’arrête à 300 époques afin d’éviter le sur-apprentissage.

On observe les courbes d’apprentissage d’entraînement et de validation avec en fond bleu la représentation de la distribution des erreurs de front d’onde en entrée des réseaux. La courbe d’entraînement à objet connu converge doucement vers sa valeur asymptote de 6 nm , tandis que la courbe de validation apparaît bien plus variable jusqu’à chuter après l’ajustement du learning rate à l’époque 350. Comparé au cas sans objet connu, l’entraînement et notamment la validation semble plus complexe dans notre méthode car la WFE n’est plus la seule métrique prise en compte lors du calcul de la loss, la $loss_{\text{phy}}$ intervient et vient perturber l’estimation directe des coefficients. De plus, l’ajout d’un second canal en entrée des réseaux vient ajouter des paramètres à ce dernier, et donc rallonger l’apprentissage des coefficients impliquant une augmentation du nombre d’époque lors de l’entraînement.

Conclusion

Le réseau converge bien vers des valeurs de front d’onde résiduel en dessous des 15 nm des spécifications optiques, validant alors le processus d’entraînement figure 5.10 sur un objet complètement maîtrisé.

5.3.2. Application aux objets simulés

Une fois notre méthode validée sur un objet ponctuel, nous proposons d'étudier notre méthode sur les objets plus complexes vus partie 5.6. Nous proposons donc d'étudier l'objet étendu étape par étape, en commençant par l'étude d'un seul objet connu sur l'ensemble du jeu de données d'entraînement et de test. Puis nous allons faire varier les objets pour chaque image. Sur l'ensemble des jeux de données, les objets et les PSFs varient pour chacune des images.

5.3.3. Diversité d'objet connu

Nous proposons maintenant d'entraîner un réseau de neurones avec comme donnée d'entrée une paire d'image composée d'un tenseur comportant image et objet, pour l'entraînement et l'opération.

Objet gaussien

Dans un premier temps, nous étudions la performance des réseaux de neurones pour un objet gaussien à fond nul. On s'intéresse aux performances à objet connu et à objet inconnu. Les deux réseaux sont entraînés sur 300 époques avec un learning rate qui décroît à partir de la 150ème époque.

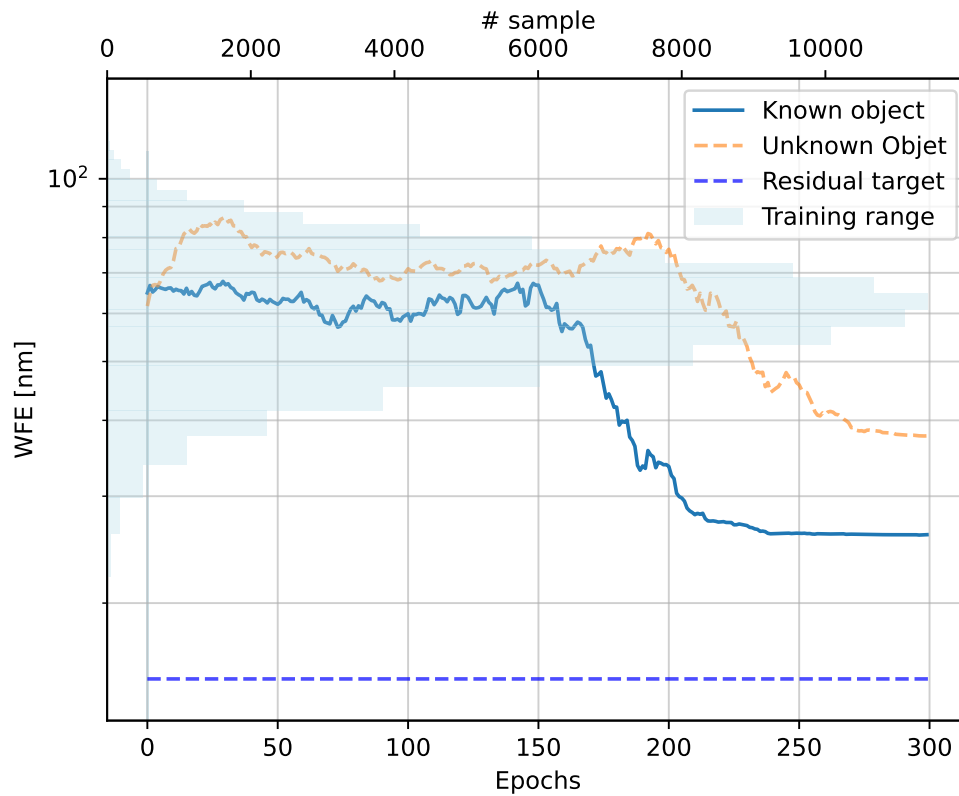


FIGURE 5.12. – Courbe de validation lors de l'apprentissage pour un objet étendu gaussien sur fond nul.

On peut observer des variations de la loss de validation conséquentes dues à l'apprentissage sur l'objet et sur la PSF. Cependant, dans les deux cas d'objet connu ou inconnu, l'erreur de front d'onde tend à être réduite à partir de la 150ème époque, époque où le learning rate décroît. Cependant, réduire le learning rate quelques époques en amont n'a pas amélioré l'estimation des réseaux, la 150ème époque semble être optimale pour l'apprentissage.

Conclusion

La limite de diffraction n'est pas atteinte même dans le cas d'un objet peu étendu sur un fond uniforme, l'analyse de front d'onde est bien plus complexe lorsque l'objet s'étale.

En effet, certaines fréquences spatiales sont minimisées voire même annulées par la formation d'image, ces dernières ne peuvent donc pas être imagées et ne contiennent donc aucune information sur les aberrations du front d'onde.

Objet créé par le VAE

Pour des objets plus réalistes, générés à partir du VAE, on propose le même type

d'entraînement. Deux types de données sont comparées : les données issues du VAE et ces mêmes données où le contraste local de l'objet a été augmenté.

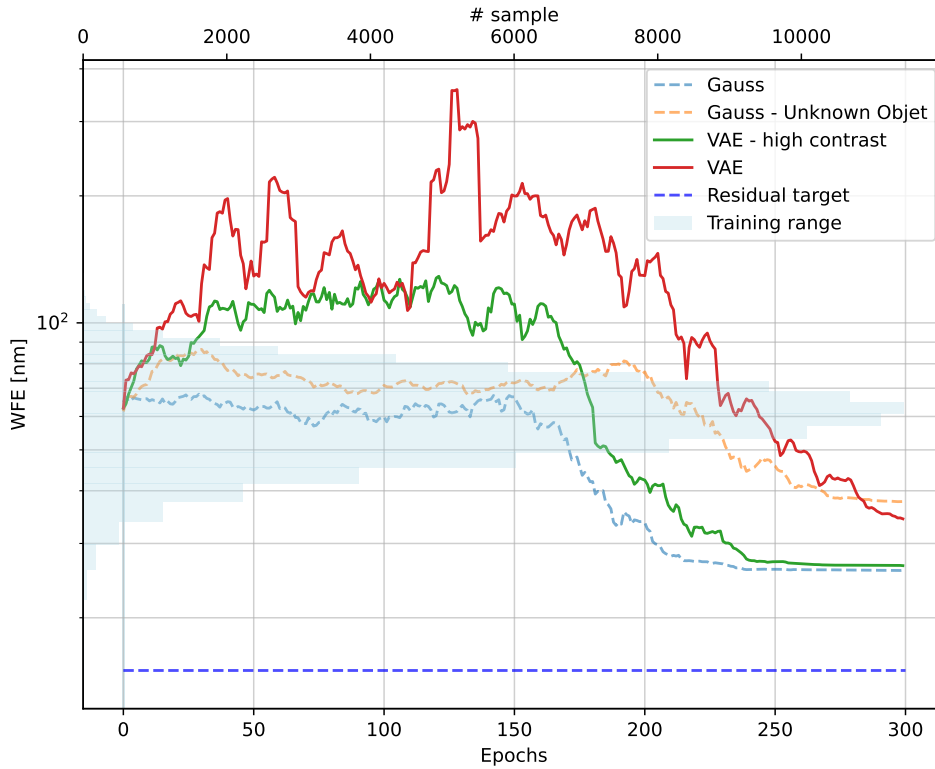


FIGURE 5.13. – Courbe de validation pour un objet étendu gaussien sur fond nul (courbes en pointillées) et sur objet issu du VAE (courbe en trait plein).

Sur la figure 5.13 on peut observer les courbes des précédents résultats sur objet gaussien sur fond nul (en pointillé) et les courbes de validation pour objets générés à partir du VAE (ligne pleine). On observe ainsi le résultat des loss de validation pour les données issues du VAE et celle des données du VAE à contraste augmenté. Les loss des réseaux se stabilisent autour de 300 époques. La courbe pour les données du VAE converge directement après la 300 ème époque. Augmenter le nombre d'époque ne n'améliore pas significativement l'erreur d'estimation de la courbe du VAE : l'entraînement à donc été arrêté à 300 époques. Le front d'onde est donc estimé et peut être réduit si il est corrigé par les systèmes d'optique active. Cependant la performance attendue pour la limite de diffraction n'est pas atteinte car la valeurs des loss est supérieure aux 15 nm. Néanmoins, en comparant donc les performances entre la courbe rouge et la courbe verte représentant respectivement les objets générés par le VAE et les objets générés par le VAE avec un contraste augmenté.

Conclusion

On peut conclure que la qualité de l'estimation du front d'onde ne dépend donc pas uniquement de la forme des objets ou des aberrations d'entrée mais bien du contraste entre l'objet et le fond qui le compose.

5.4. Limitation de l'analyse de front d'onde

5.4.1. Contraste de l'objet

On propose donc d'étudier la robustesse des deux réseaux issus des données VAE aux différences de contraste entre les images à différents fronts d'onde incidents. Tout d'abord on définit le contraste c tel que

$$c = \frac{\max(I) - \min(I)}{\max(I) + \min(I)} \quad (5.6)$$

On peut alors définir des niveaux de contraste faible, moyen et élevé; correspondant au niveau moyen de contraste respectivement des données du VAE, des données du VAE à contraste augmenté et aux objets gaussiens :

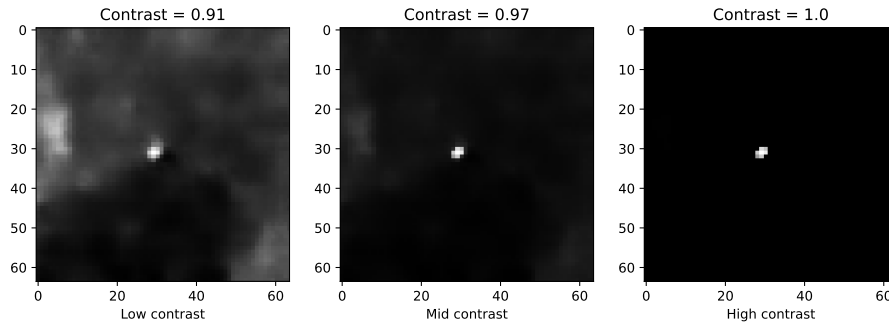


FIGURE 5.14. – Exemple de contraste des objets étudiés.

Nous allons maintenant estimer les coefficients de 10 images par niveau d'aberration et moyenner le résultat à contraste fixe, puis faire varier le contraste et le niveau d'aberrations des PSFs afin d'étudier la qualité d'estimation des réseaux en fonction du contraste et de l'amplitude du front d'onde.

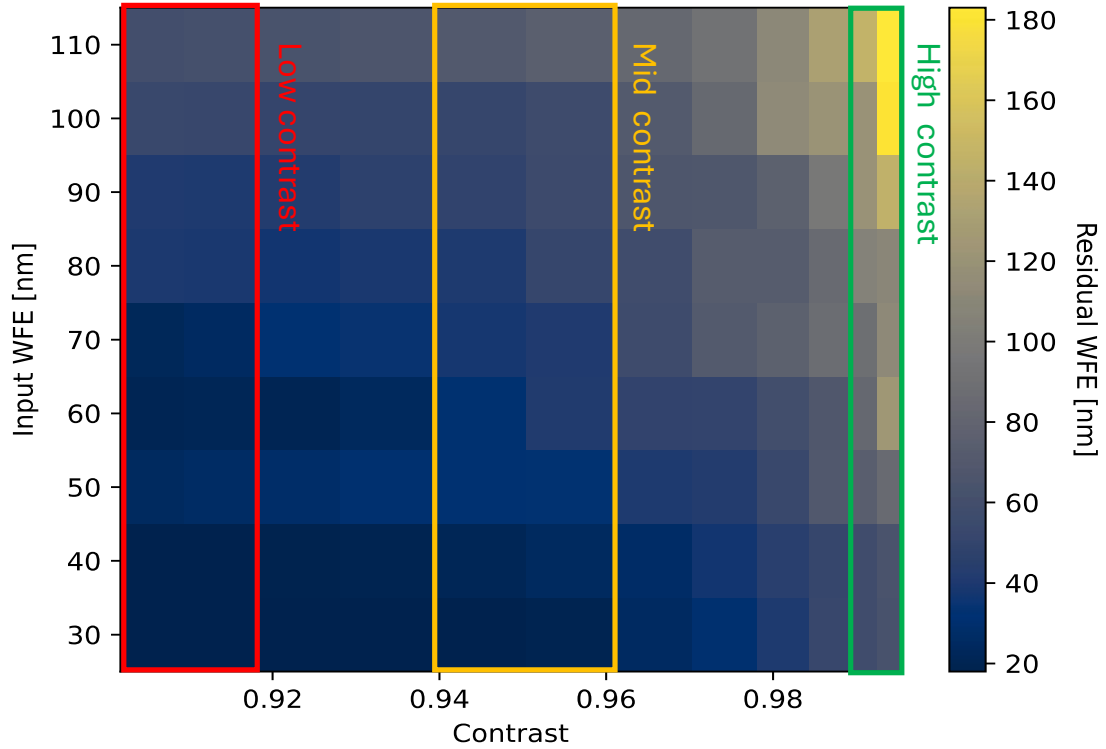


FIGURE 5.15. – Étude de la qualité d'estimation du PINN (colorbar) en fonction du contraste de l'image et du front d'onde incident pour le réseau entraîné sur les données issues du VAE.

Ici la première colonne correspond aux données mêmes du VAE, à faible contraste. Les zones de moyen et fort contraste sont aussi représentées respectivement en orange et en vert. Le réseau une fois entraîné permet de réduire l'erreur front d'onde mais ne permet pas d'atteindre la limite de diffraction. Évidemment, au plus le WFE incident est élevé au plus l'erreur d'estimation augmente. Cependant, lorsque le contraste augmente, la qualité d'estimation du réseau augmente en conséquence, atteignant des performances meilleures que sur les données comportant le même contraste que celui d'entraînement (colonne de gauche).

On peut alors s'intéresser aux résultats produits par le réseau entraîné sur les données du VAE à contraste augmenté et estimer sa robustesse au contraste. On effectue alors le même test que ci-dessus.

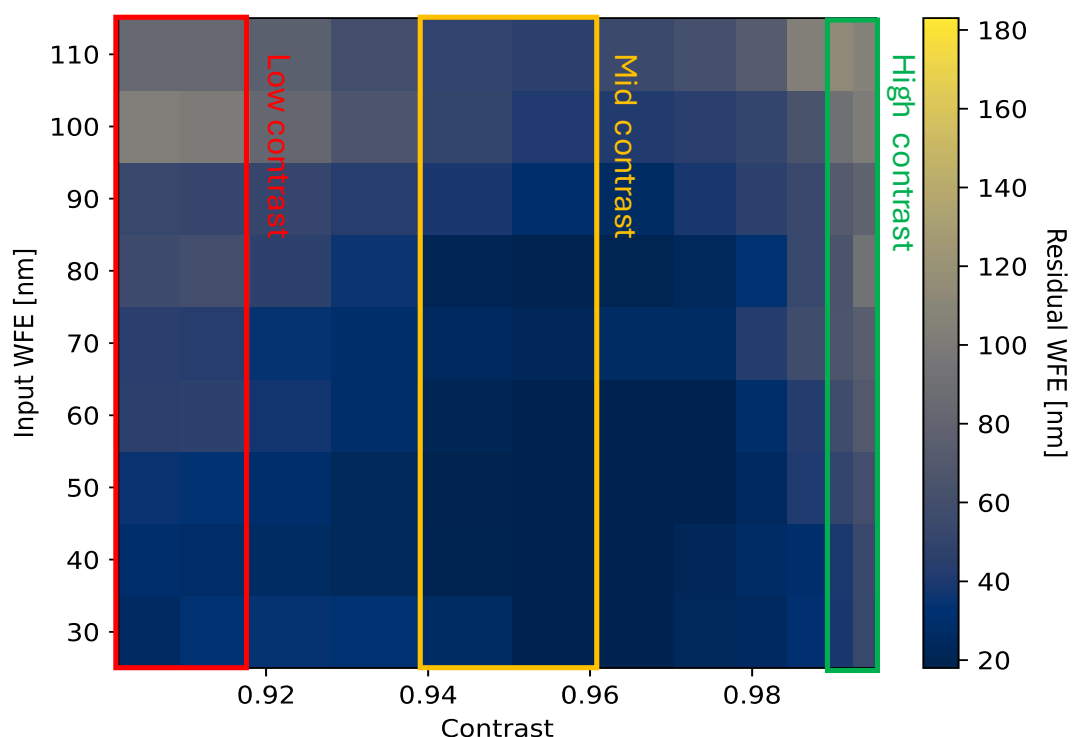


FIGURE 5.16. – Étude de la qualité d'estimation du PINN (colorbar) en fonction du contraste de l'image et du front d'onde incident pour le réseau entraîné sur les données issues du VAE à contraste augmenté.

La même barre de couleur que sur la figure 5.15 est utilisée en figure 5.16. On remarque que l'estimation du réseau est plus stable en fonction du contraste. La colonne de gauche représente les données issues du VAE directement. Le contraste de chaque objet est ensuite augmenté jusqu'à atteindre le contraste utilisé pour l'entraînement du réseau (autour d'un contraste de 0.96) puis est augmenté encore jusqu'à arriver au contraste maximum représentant un objet brillant sur fond nul. Les résultats d'estimation du réseau semblent assez homogène en fonction du contraste. Cependant, en comparant avec les performances du réseau entraîné sur les données VAE, le réseau présenté ici montre une performance bien plus robuste au contraste et estime correctement le front d'onde, sans pour autant atteindre la limite de diffraction en moyenne.

Conclusion

L'estimation du réseau est la meilleure pour des données au contraste autour du contraste utilisé pour l'entraînement de ce dernier. Néanmoins ces deux résultats montrent une nette dépendance au contraste de l'objet pour l'analyse de front d'onde.

5.4.2. Sensibilité de la phase au contraste

La difficulté des réseaux à identifier correctement la phase à partir de l'image est donc en partie due à la perte de sensibilité de l'image aux variations des PSFs lors de faible contraste de l'objet. Si l'on s'intéresse à la variation de l'intensité des pixels de l'image en étudiant l'écart type σ_I de la différence entre une image de référence générée à partir d'une PSF à la limite de la diffraction et une PSF avec un WFE plus ou moins importe. On peut définir la sensibilité S

$$S = \sigma_I (|I_{\text{abb}} - I_{\text{ref}}|) \quad (5.7)$$

On peut alors étudier alors l'allure de la sensibilité comme définie en fonction du contraste et du front d'onde incident :

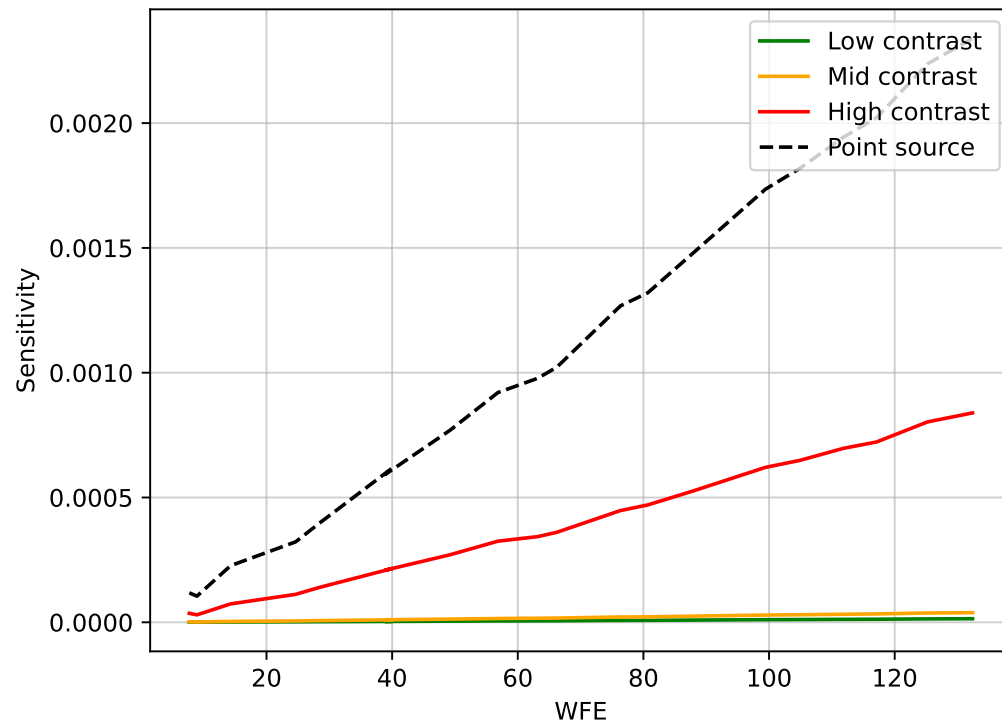


FIGURE 5.17. – Courbe de l'évolution de la sensibilité en fonction du contraste. La courbe verte représente les variations pour un petit contraste, la courbe orange pour un contraste moyen, et enfin la courbe rouge pour un contraste élevé. La courbe en pointillé représente le cas pour un point source, objet dont le contraste est maximal.

Les courbes pleines de la figure 5.17 représentent la sensibilité de l'intensité d'une image en fonction du front d'onde incident. La sensibilité est très faible lorsque le

contraste est peu marqué (courbe verte), ce qui signifie qu'un léger changement dans le front d'onde conduit à une variation de l'intensité des pixels de l'image très faible voire quasiment nulle. Cela réduit donc la capacité des réseaux (et de n'importe quelle méthode mono-image) à identifier finement les variations du front d'onde par rapport à l'objet. À l'inverse, lorsqu'un objet plus contrasté est imagé, les variations d'intensité des images sont plus prononcées, ce qui permet de mieux apprécier les détails de la phase à partir de l'image.

Conclusion

Ainsi, ces variations de contraste, notamment dans le cas d'objets réels ou simulés par le VAE, limitent l'estimation correcte du front d'onde.

Néanmoins, l'étude sur point source a montrée que les performances peuvent être améliorées lors du passage en boucle fermée de notre méthode. C'est ce que nous allons étudier dans la prochaine partie.

5.5. Phasage en boucle fermée

L'objet limitant l'estimation des erreurs de phasage jusqu'aux spécifications du cahier des charges d'AZIMOV, un scénario de phasage en orbite face à la surface de la Terre peut être proposé en boucle fermée. En effet, le télescope défilant autour de la Terre observe des objets variés sur sa surface tandis que les erreurs de phasage restent presque constantes. Ce scénario permet de fermer une boucle de phasage où, à chaque itération, l'objet change et est convolué par la PSF résiduelle de l'itération précédente. Un schéma du processus de boucle fermé est présenté figure 5.18.

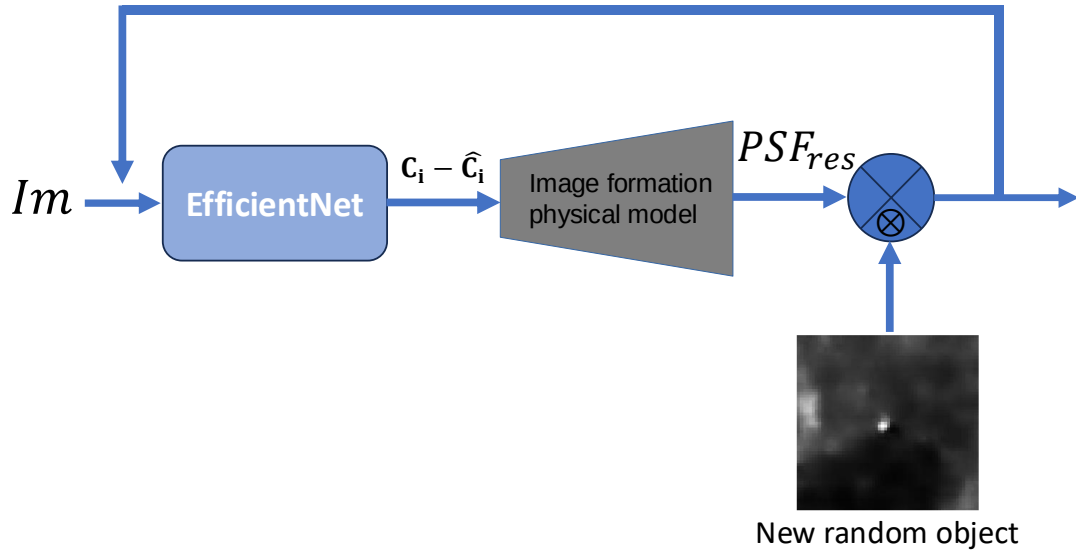


FIGURE 5.18. – Schéma de la boucle fermée sur objet étendu.

Le schéma de la figure 5.18 présente l'étude de la boucle fermée proposée. A chaque itération, les coefficients de la PSF $\hat{c}_i$ sont estimés. La PSF résiduelle PSF_{res} est ensuite formée à l'aide de notre modèle de formation d'image, puis un nouvel objet tiré aléatoirement puis convolué à la PSF résiduelle. L'image ainsi formée est utilisée une en entrée du réseau de neurones afin d'en estimer les coefficients de la PSF. On peut appliquer cette méthode aux réseaux entraînés sur des données VAE et des données VAE à contraste augmenté :

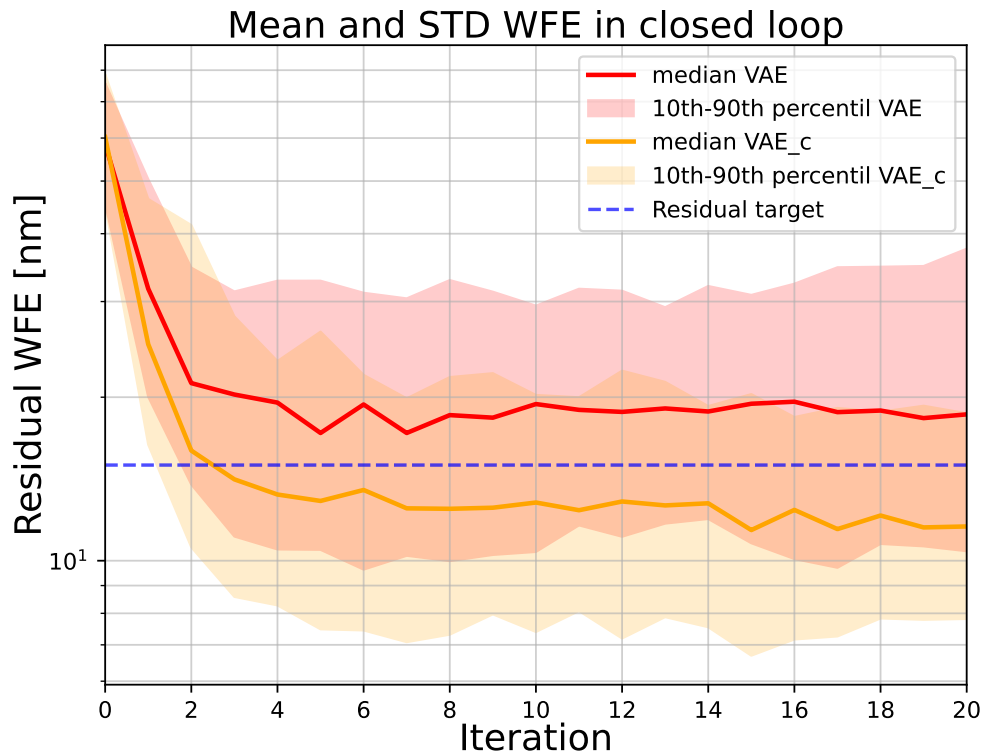


FIGURE 5.19. – Exemple de boucle fermée pour le phasage temps réel d'AZIMOV.

La figure 5.19 présente le front d'onde résiduel moyen dans une boucle fermée où l'objet varie à chaque itération. En rouge sont présentés les résultats du réseau entraîné sur les données à bas contraste (VAE), en orange le réseau entraîné sur les données à moyen contraste (VAE_c). On peut observer que le front d'onde est correctement identifié et est réduit au fur et à mesure des itérations. Cependant, les spécifications en termes d'estimation du front d'onde ne sont pas atteintes dans le cas de données à faible contraste, la valeur minimum du front d'onde résiduel moyen étant d'une vingtaine de nanomètres. Néanmoins, lorsque le contraste est plus élevé sur les objets observés, la limite de diffraction est atteinte en moyenne.

Toutefois, la dispersion des mesures ne permet pas de garantir un phasage stable malgré une erreur de front d'onde minime.

Conclusion

Dans le scénario le moins avantageux, le front d'onde est au minimum réduit de moitié, l'estimation est donc avantageuse pour le front d'onde total.

5. Etude de l'objet étendu – 5.5. Phasage en boucle fermée

En somme, l'analyse du front d'onde sur un objet étendu se révèle bien plus complexe que l'étude d'un point source en raison de l'ajout d'inconnues provenant de l'objet. Les premières limitations apparaissent en termes de formation d'image, où la forme particulière de notre pupille limite l'imagerie de certaines fréquences spatiales de l'objet allant jusqu'à les annuler et les rendant ainsi inexploitable. La méthode directe fonctionnant sur point source montre rapidement ses limites lors de l'étude d'un objet étendu : la complexité de l'objet ou encore les variations de ce dernier rendent le problème trop complexe pour exploiter correctement cette méthode. Une méthode incorporant la physique de la formation d'image durant l'entraînement permet de stabiliser ce dernier et de généraliser l'estimation du front d'onde aux objets et aux PSFs, au prix d'une estimation de moindre qualité ne respectant pas les limitations du cahier des charges. Le contraste de l'objet d'étude perturbe l'estimation de la phase en minimisant la sensibilité de l'évolution des intensités de l'image à différentes amplitudes de front d'onde. Néanmoins, cette caractéristique étant dépendante de l'objet, le test de la méthode PINN en boucle fermée assure une minimisation de l'erreur de front d'onde et peut être considéré comme une méthode de phasage convenable.

Conclusion et perspectives

Dans le domaine de l'observation de la Terre puis l'espace, l'utilisation d'un petit télescope déployable permet de découpler les applications d'imagerie. Celui-ci permet en effet de coupler un haut taux de revisite aux précisions d'imagerie apportées par la haute résolution angulaire. Le prix à payer pour utiliser un télescope segmenté est de devoir garantir un alignement parfait de son miroir primaire par l'ajustement actif des segments.

Nous avons vu que les méthodes à l'état de l'art du phasage en plan-focal permettent une estimation cohérente du front d'onde, moyennant des processus itératifs et donc un coût calculatoire élevé, peu compatible d'une petite plateforme spatiale. L'essor de l'intelligence artificielle vient apporter une corde de plus à notre arc, venant offrir une approche novatrice, rapide et robuste pour mesurer ces erreurs de phasage. Le coût calculatoire étant alors transformé en apprentissage au sol et donc déplacé en amont de l'opération, les réseaux s'inscrivent comme une méthode efficace et légère de résolution du problème inverse. Nous avons proposé au chapitre 2 comment utiliser ces réseaux pour faire de l'analyse de front d'onde.

Comme démontré dans le chapitre 3, l'étude du Deep Learning pour le phasage sur un point source offre des résultats extrêmement concluants. En passant par l'architecture interne des réseaux, nous avons vu que les limitations physiques de l'AFOPF à une seule image ne pouvaient être contrecarrées. Cependant, apporter une diversité d'amplitude en brisant la centro-symétrie de notre télescope permet de résoudre l'ambiguïté de signe propre à l'analyse de front d'onde à partir de l'image plan focal. Forts de cette asymétrie, la méthode directe d'apprentissage supervisé des réseaux de neurones à partir des PSFs peut s'articuler correctement. La robustesse des réseaux a ensuite été prouvée dans des conditions plus complexes : bruit et aberrations de plus hauts ordres. Les méthodes employées ont montré en outre un avantage significatif par rapport aux méthodes à l'état de l'art : le temps d'exécution. Les réseaux étant (une fois entraînés) déterministes, ils dominent les méthodes itératives de l'état de l'art notamment en termes de temps de calcul. Si cet avantage n'a qu'un intérêt limité dans le cas d'un télescope en orbite basse, soumis aux seuls gradients thermiques évoluant lentement au cours d'une orbite, cet atout permet cependant un phasage en temps réel et l'exploitation d'une boucle fermée du phasage. Une ouverture sur des problèmes connexes tel que le phasage de télescope comportant plus de segments et des résultats préliminaires sur l'étude du mode "Pétale" de l'ELT ont montré efficacité mais aussi certaines limites de l'application de notre méthode.

Les méthodes développées en simulation au chapitre 3 ont été validées expéri-

mentalement dans le chapitre 4. Nous avons pu valider la méthode sur données expérimentales en complexifiant le modèle. Par ailleurs, nous avons pu mettre en avant l'importance du ratio entre données expérimentales et données simulées. L'apprentissage peut se faire sur données simulées comme sur données expérimentales. Ceci a permis alors une approche "à la carte" du phasage de télescope en fonction des données disponibles, et validant l'applicabilité des réseaux de neurones légers sur des données réelles.

Enfin, nous avons abordé dans le chapitre 5 l'étude d'un objet étendu. Ce vaste sujet a été abordé en surface et a rapidement montré les limites de la méthode directe (développée pour traiter le point source) d'estimation du front d'onde. Cela a néanmoins permis la mise en place d'une méthode plus sophistiquée d'entraînement des réseaux, en incorporant le modèle physique de la formation d'image au sein même de l'entraînement du réseau et de permettre à ces derniers de généraliser leur apprentissage sur diverses PSFs et divers objets. Cependant, notre méthode se retrouve très limitée par le contraste de l'objet, caractéristique essentielle pour la sensibilité des images étudiées et donc pour l'estimation de la phase par réseau de neurones.

En conclusion, le deep learning est une méthode novatrice et efficace pour l'analyse du front d'onde. Il s'intègre de plus en plus dans les travaux des chercheurs en optique pour les grands télescopes. Il s'avère être un outil puissant et efficace pour améliorer les performances actuelles. Ces méthodes doivent être utilisées comme des outils complémentaires aux méthodes actuelles bien étudiées afin de combiner le meilleur des deux mondes entre physique et intelligence artificielle, et ainsi atteindre des performances au-delà des attentes dans le futur. De manière générale, les travaux présentés dans cette thèse constituent une avancée innovante, permettant l'application de méthodes d'apprentissage automatique pour le phasage de petits télescopes segmentés. Ces recherches offrent une première étude approfondie sur le cadre d'application et les limites de ces méthodes dans le contexte du phasage de télescope.

Perspectives

Ce travail de thèse a permis d'identifier plusieurs pistes restant à explorer. Je souhaite les développer ici succinctement.

L'analyse sur une source ponctuelle s'est montrée très concluante dans un cas monochromatique. Ce cas nécessaire est cependant peu réaliste dans un cas d'observation de la Terre, où la lumière renvoyée par la surface terrestre est large bande. La généralisation de la méthode proposée à un cas à large bande spectrale nous paraît donc très pertinent. Ce cas nécessite de complexifier les modèles lors de l'apprentissage afin d'entraîner un réseau sur des images à large bande, ou sur des images dans différentes bandes spectrales. Cet ajout, au-delà de représenter un gain en représentativité, permet également de façon sûre d'élargir la dynamique de mesure du piston entre les segments et de dépasser la limite imposée par la longueur d'onde.

Dans le cas de l'étude sur des données expérimentales, nous avons validé l'aspect

mesure de front d'onde. Une démonstration de la capacité de correction complète incluant une correction en temps réel sur le démonstrateur de UK-ATC permettrait de valider la robustesse de la méthode et d'assurer un phasage complet et stable du télescope.

Une perspective intéressante à la méthode développée sur source ponctuelle serait son application à l'imagerie d'exoplanètes. Dans ce cas, on souhaite détecter les quelques photons émis par une exoplanète à proximité d'une étoile bien plus brillante (un facteur allant jusqu'à quelques centaines de milliards dans le cas du futur observatoire HWO). Pour optimiser la détection de ces photons, il est nécessaire de supprimer optiquement la lumière émise par l'étoile. Ceci se fait en utilisant un coronographe, composant optique permettant d'éteindre un point source lumineux. Pour atteindre les niveaux d'extinction requis pour HWO, il est nécessaire de mesurer dans une image coronographique tous les résidus de front d'onde. Une image coronographique correspond à une formation d'image particulièrement non-linéaire, et l'utilisation des méthodes d'apprentissage machine paraît toute indiquée. La mesure de défauts de phasage démontrée dans ce manuscrit est une première étape prometteuse.

Enfin l'étude de la mesure de surface d'onde sur objets étendus offre également de nombreuses perspectives intéressantes. Tout d'abord, la simulation d'objets pourrait être améliorée par des méthodes plus complexes, telles que les réseaux génératifs adverses (*Generative Adversarial Networks*). Le fait que les aberrations de champ d'un télescope comme celui d'AZIMOV ne varient pas (ou peu) dans le champ de vue permettrait également de sélectionner plusieurs objets dans une même image, offrant ainsi une plus grande diversité d'objets pour une même estimation des paramètres de la PSF. L'utilisation de plusieurs canaux en entrée des réseaux de neurones permettrait en effet de combiner ces différents canaux pour potentiellement obtenir une estimation plus précise.

Par ailleurs, l'étude des objets étendus s'est limitée à des objets à forme circulaires et à fort contraste. Des améliorations peuvent être apportées à ce niveau sachant la dépendance au contraste pour l'AFOPF reste un facteur limitant. Néanmoins, d'autres régions à haut contraste peuvent également être étudiées, comme celles à bords francs (rivières, grands bâtiments, ombres, etc.). Ces régions, porteuses d'informations sur une direction de la phase, peuvent compléter notre étude sur les sources ponctuelles étendues.

Bibliographie

- [Act+12] D Scott ACTON, J Scott KNIGHT, A CONTOS et al. « Wavefront sensing and controls for the James Webb Space Telescope ». In : Space Telescopes and Instrumentation 2012. T. 8442. SPIE. 2012, p. 877-887 (cf. p. [33](#)).
- [Act+22] D Scott ACTON, Scott KNIGHT, Maria CARRASQUILLA et al. « Phasing the Webb telescope ». In : Space Telescopes and Instrumentation 2022. T. 12180. SPIE. 2022, p. 329-352 (cf. p. [33-35](#)).
- [AAD21] Collins E AKUMU, Emmanuel O AMADI et S DENNIS. « Application of Drone and WorldView-4 Satellite Data in Mapping and Monitoring Grazing Land Cover and Pasture Quality : Pre- and Post-Flooding ». In : Land 10.3 (2021), p. 321 (cf. p. [16](#)).
- [ASB08] C ALVAREZ DIEZ, F SHAO et J BILLE. « Pyramid and Hartmann–Shack wavefront sensor with artificial neural network for adaptive optics ». In : Journal of Modern Optics 55.4-5 (2008), p. 683-689 (cf. p. [37](#)).
- [AOE20] Torben ANDERSEN, Mette OWNER-PETERSEN et Anita ENMARK. « Image-based wavefront sensing for astronomy using neural networks ». In : Journal of Astronomical Telescopes, Instruments, and Systems 6.3 (2020), p. 034002-034002 (cf. p. [37](#), [38](#)).
- [Ber+22] A BERTROU-CANTOU, E GENDRON, G ROUSSET et al. « Confusion in differential piston measurement with the pyramid wavefront sensor ». In : Astronomy & Astrophysics 658 (2022), A49 (cf. p. [45](#)).
- [Bla02] Amandine BLANC. « Identification de réponse impulsionnelle et restauration d'images : apport de la diversité de phase ». Thèse de doct. Paris 11, 2002 (cf. p. [30](#)).
- [CTP00] David A CARRARA, Brian J THELEN et Richard G PAXMAN. « Aberration correction of segmented-aperture telescopes by using phase diversity ». In : Image reconstruction from incomplete data. T. 4123. SPIE. 2000, p. 56-63 (cf. p. [27](#)).
- [Cha+23] Charleston CHAUVET, Morgan GRAY, Benoit NEICHEL et al. « Estimating the ELT's Differential Pistons with Deep Learning ». In : Adaptive Optics for Extremely Large Telescopes 7th Edition. 2023 (cf. p. [135](#), [136](#)).

- [Cuo+22] Salvatore CUOMO, Vincenzo Schiano DI COLA, Fabio GIAMPAOLO et al. « Scientific machine learning through physics-informed neural networks : Where we are and what's next ». In : Journal of Scientific Computing 92.3 (2022), p. 88 (cf. p. 174).
- [Dam+12] Marcos A van DAM, Rodolphe CONAN, Antonin H BOUCHEZ et al. « Design of a truth sensor for the GMT laser tomography adaptive optics system ». In : Adaptive Optics Systems III. T. 8447. SPIE. 2012, p. 482-492 (cf. p. 33).
- [Dea+06] Bruce H DEAN, David L ARONSTEIN, J Scott SMITH et al. « Phase retrieval algorithm for JWST flight and testbed telescope ». In : Space telescopes and instrumentation I. T. 6265. SPIE. 2006, p. 314-330 (cf. p. 35).
- [DKC18] Dennis DOLKENS, Hans KUIPER et Victor Villalba CORBACHO. « The deployable telescope : a cutting-edge solution for high spatial and temporal resolved Earth observation ». In : Advanced Optical Technologies 7.6 (2018), p. 365-376 (cf. p. 18, 19).
- [EV21] David ESCOBAR et Esteban VERA. « Wavefront sensing using deep learning for Shack Hartmann and pyramidal sensors ». In : 2021 IEEE CHILEAN Conference (CHILECON). IEEE. 2021, p. 1-5 (cf. p. 37).
- [Fra+19] Carlo de FRANCHIS, Enric MEINHARDT-LLOPIS, Jean MICHEL et al. « An automatic and modular hyperspectral stereo mosaicking pipeline for SPOT-6 and SPOT-7 ». In : IEEE Transactions on Geoscience and Remote Sensing 57.6 (2019), p. 3825-3834 (cf. p. 16).
- [Fuk80] Kunihiro FUKUSHIMA. « Neocognitron : A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position ». In : Biological cybernetics 36.4 (1980), p. 193-202 (cf. p. 79).
- [Gon82] Robert A GONSALVES. « Phase retrieval and diversity in adaptive optics ». In : Optical Engineering 21.5 (1982), p. 829-832 (cf. p. 26).
- [Guo+06] Hong GUO, Nina KORABLOVA, Qiushi REN et al. « Wavefront reconstruction with artificial neural networks ». In : Optics Express 14.14 (2006), p. 6456-6462 (cf. p. 37).
- [He+16] Kaiming HE, Xiangyu ZHANG, Shaoqing REN et al. « Deep residual learning for image recognition ». In : IEEE conference on computer vision and pattern recognition. 2016, p. 770-778 (cf. p. 87).
- [Her+18] Jörg HERBEL, Tomasz KACPRZAK, Adam AMARA et al. « Fast point spread function modeling with deep learning ». In : Journal of Cosmology and Astroparticle Physics 2018.07 (2018), p. 054 (cf. p. 37, 38).

- [HS+52] Magnus Rudolph HESTENES, Eduard STIEFEL et al. Methods of conjugate gradients for solving linear systems. T. 49. 1. NBS Washington, DC, 1952 (cf. p. [124](#)).
- [Hin+10] Philip M HINZ, Antonin BOUCHEZ, Matt JOHNS et al. « The GMT adaptive optics system ». In : Adaptive Optics Systems II. T. 7736. SPIE. 2010, p. 139-150 (cf. p. [33](#)).
- [Hoc98] Sepp HOCHREITER. « The vanishing gradient problem during learning recurrent neural nets and problem solutions ». In : IJUFKBS 6.02 (1998), p. 107-116 (cf. p. [87](#)).
- [JWG13] Shuanggen JIN, Jun WANG et Peng GONG. « Land-cover mapping using a hybrid classification approach and multi-source GeoEye-1 data ». In : Remote Sensing 5.12 (2013), p. 6254-6273 (cf. p. [17](#)).
- [KGB18] Andreas KÄÄB, Luc GIROD et Etienne BERTHIER. « Glacier elevation and mass changes over High Mountain Asia based on multi-temporal DEMs from stereo satellite data ». In : The Cryosphere 12 (2018), p. 649-663 (cf. p. [16](#)).
- [KAD94] Richard L KENDRICK, Daniel S ACTON et AL DUNCAN. « Phase-diversity wave-front sensor for imaging systems ». In : Applied Optics 33.27 (1994), p. 6533-6546 (cf. p. [28](#)).
- [KB14] Diederik P KINGMA et Jimmy BA. « Adam : A method for stochastic optimization ». In : arXiv preprint arXiv :1412.6980 (2014) (cf. p. [83](#)).
- [KW13] Diederik P KINGMA et Max WELLING. « Auto-encoding variational bayes ». In : arXiv preprint arXiv :1312.6114 (2013) (cf. p. [90](#)).
- [Koz+00] Lester J KOZLOWSKI, John T MONTROY, Craig A CABELLI et al. « Visible and infrared detectors at Rockwell Science Center ». In : Optical and IR Telescope Instrumentation and Detectors. T. 4008. SPIE. 2000, p. 1240-1253 (cf. p. [32](#)).
- [KSH12] Alex KRIZHEVSKY, Ilya SUTSKEVER et Geoffrey E HINTON. « Imagenet classification with deep convolutional neural networks ». In : Advances in neural information processing systems 25 (2012) (cf. p. [82](#), [85](#)).
- [Leb+13] Nicolas LEBEGUE, Thierry TOUTIN, Laurent BLONDE et al. « Geometric Evaluation of Pléiades Tri-Stereo Satellite Images ». In : ISPRS. T. XL-1/W1. Copernicus GmbH. 2013, p. 243-248 (cf. p. [16](#)).
- [Lee+97] David J LEE, Byron M WELSH, Michael C ROGGMANN et al. « Diagnosing unknown aberrations in an adaptive optics system by use of phase diversity ». In : Optics letters 22.13 (1997), p. 952-954 (cf. p. [28](#)).

- [Lev+24] Nicolas LEVRAUD, Vincent CHAMBOULEYRON, Jean François SAUVAGE et al. « A strategy for sensing the petal mode in the presence of AO residual turbulence with the pyramid wavefront sensor ». In : Astronomy & Astrophysics 682 (2024), A84 (cf. p. [45](#), [135](#)).
- [LDS98] Mats G LOFDAHL, Alan L DUNCAN et Goran B SCHARMER. « Fast phase diversity wavefront sensing for mirror control ». In : Adaptive Optical System Technologies. T. 3353. SPIE. 1998, p. 952-963 (cf. p. [28](#)).
- [Lof+98] Mats G LOFDAHL, Rick Lee KENDRICK, Alex HARWIT et al. « Phase diversity experiment to measure piston misalignment on the segmented primary mirror of the Keck II telescope ». In : Space Telescopes and Instruments V. T. 3356. SPIE. 1998, p. 1190-1201 (cf. p. [27](#)).
- [LS94] Mats G LOFDAHL et Goran B SCHARMER. « Application of phase-diversity to solar images ». In : Image reconstruction and restoration. T. 2302. SPIE. 1994, p. 254-267 (cf. p. [28](#)).
- [Mai+20] Erin R MAIER, Phil HINZ, Denis DEFRÈRE et al. « Implementing multiwavelength fringe tracking for the Large Binocular Telescope Interferometer's phase sensor, PHASECam ». In : Journal of Astronomical Telescopes, Instruments, and Systems 6.3 (2020), p. 035001-035001 (cf. p. [32](#)).
- [Mar13] Frantz MARTINACHE. « The asymmetric pupil Fourier wavefront sensor ». In : Publications of the Astronomical Society of the Pacific 125.926 (2013), p. 422 (cf. p. [52](#)).
- [MB74] Richard A MULLER et Andrew BUFFINGTON. « Real-time correction of atmospherically degraded telescope images through image sharpening ». In : JOSA 64.9 (1974), p. 1200-1210 (cf. p. [28](#)).
- [Mur+14] Kiran MURTHY, Michael SHEARN, Byron D SMILEY et al. « SkySat-1 : very high-resolution imagery from a small satellite ». In : Sensors, systems, and next-generation Satellites XVIII. T. 9241. SPIE. 2014, p. 367-378 (cf. p. [71](#)).
- [Nis+19] Yohei NISHIZAKI, Matias VALDIVIA, Ryoichi HORISAKI et al. Deep learning wavefront sensing. New York : Opt. Express 27, 2019 (cf. p. [37](#), [38](#)).
- [Nol76] Robert J NOLL. « Zernike polynomials and atmospheric turbulence ». In : JOsA 66.3 (1976), p. 207-211 (cf. p. [55](#)).
- [Orb+21] Gilles ORBAN DE XIVRY, Maxime QUESNEL, PO VANBERG et al. « Focal plane wavefront sensing using machine learning : performance of convolutional neural networks compared to fundamental limits ». In : Monthly Notices of the Royal Astronomical Society 505.4 (2021), p. 5702-5713 (cf. p. [37](#), [38](#)).

- [Os+14] J OSBORN, Dani GUZMÁN, FJ de COS JUEZ et al. « First on-sky results of a neural network based tomographic reconstructor : Carmen on Canary ». In : Adaptive Optics Systems IV. T. 9148. SPIE. 2014, p. 1541-1546 (cf. p. 37).
- [PF18] Scott W PAINE et James R FIENUP. « Machine learning for improved image-based wavefront sensing ». In : Optics letters 43.6 (2018), p. 1235-1238 (cf. p. 37, 38).
- [Pow64] Michael JD POWELL. « An efficient method for finding the minimum of a function of several variables without calculating derivatives ». In : The computer journal 7.2 (1964), p. 155-162 (cf. p. 124).
- [Que+22] Maxime QUESNEL, G Orban de XIVRY, Gilles LOUPPE et al. « A deep learning approach for focal-plane wavefront sensing using vortex phase diversity ». In : arXiv preprint arXiv :2210.00632 (2022) (cf. p. 38).
- [Qui+18] Fernando QUIRÓS-PACHECO, David SCHWARTZ, Kaushik DAS et al. « The Giant Magellan Telescope phasing strategy and performance ». In : Ground-based and Airborne Telescopes VII. T. 10700. SPIE. 2018, p. 181-191 (cf. p. 33).
- [Rag96] Roberto RAGAZZONI. « Pupil plane wavefront sensing with an oscillating prism ». In : Journal of modern optics 43.2 (1996), p. 289-293 (cf. p. 30).
- [RTF22] Heriniaina F RAJA OBERISON, Joseph S TANG et James R FIENUP. « Machine learning wavefront sensing for the James Webb Space Telescope ». In : Space Telescopes and Instrumentation 2022. T. 12180. SPIE. 2022, p. 2210-2216 (cf. p. 37, 38).
- [Rod81] François RODDIER. « V the effects of atmospheric turbulence in optical astronomy ». In : Progress in optics. T. 19. Elsevier, 1981, p. 281-376 (cf. p. 22).
- [Ros+22] Fabio ROSSI, Cédric PLANTET, Anne-Laure CHEFFOT et al. « Machine learning techniques for piston sensing ». In : Adaptive Optics Systems VIII. T. 12185. SPIE. 2022, p. 1669-1674 (cf. p. 37, 38).
- [RF22] Gérard ROUSSET et Thierry FUSCO. « Optique adaptative : correction des effets de la turbulence atmosphérique sur les images astronomiques ». In : Comptes Rendus. Physique 23.S1 (2022), p. 1-52 (cf. p. 29, 30).
- [Sch+22] Noah SCHWARTZ, William BRZOZOWSKI, Zeshan ALI et al. « 6U CubeSat deployable telescope for optical Earth observation and astronomical optical imaging ». In : Space Telescopes and Instrumentation 2022. T. 12180. SPIE. 2022, p. 1097-1108 (cf. p. 39, 40, 43).
- [Sch+20] Noah SCHWARTZ, William BRZOZOWSKI, Maria MILANOVA et al. « High-resolution deployable CubeSat prototype ». In : Space Telescopes and Instrumentation 2020. T. 11443. SPIE. 2020, p. 565-577 (cf. p. 39, 40, 42).

- [Sha71] Roland V SHACK. « Production and use of a lenticular Hartmann screen ». In : Spring Meeting of Optical Society of America, 1971. T. 656. 1971 (cf. p. 30).
- [SZ14] Karen SIMONYAN et Andrew ZISSERMAN. « Very deep convolutional networks for large-scale image recognition ». In : arXiv preprint arXiv :1409.1556 (2014) (cf. p. 85).
- [Sut+13] Ilya SUTSKEVER, James MARTENS, George DAHL et al. « On the importance of initialization and momentum in deep learning ». In : Proceedings of the 30th International Conference on Machine Learning. Sous la dir. de Sanjoy DASGUPTA et David MCALLESTER. T. 28. Proceedings of Machine Learning Research 3. Atlanta, Georgia, USA : PMLR, 17–19 Jun 2013, p. 1139-1147. URL : <https://proceedings.mlr.press/v28/sutskever13.html> (cf. p. 83).
- [Swa+18] Robin SWANSON, Masen LAMB, Carlos CORREIA et al. « Wavefront reconstruction and prediction with convolutional neural networks ». In : Adaptive Optics Systems VI. T. 10703. SPIE. 2018, p. 481-490 (cf. p. 37).
- [Tam+22] Roberto TAMAI, Bertrand KOEHLER, Michele CIRASUOLO et al. « Status of the ESO's ELT construction ». In : Ground-based and Airborne Telescopes IX. T. 12182. SPIE. 2022, p. 437-453 (cf. p. 14).
- [TL19] Mingxing TAN et Quoc LE. « Efficientnet : Rethinking model scaling for convolutional neural networks ». In : International conference on machine learning. PMLR. 2019, p. 6105-6114 (cf. p. 88).
- [Tha+16] Niranjana A THATTE, Fraser CLARKE, Ian BRYSON et al. « The E-ELT first light spectrograph HARMONI : capabilities and modes ». In : Ground-based and Airborne Instrumentation for Astronomy VI. T. 9908. SPIE. 2016, p. 595-605 (cf. p. 30).
- [VCM17] S VIEVARD, F CASSAING et LM MUGNIER. « Large amplitude tip/tilt estimation by geometric diversity for multiple-aperture telescopes ». In : JOSA A 34.8 (2017), p. 1272-1284 (cf. p. 43, 128).
- [Wan+20] Fei WANG, Yaoming BIAN, Haichao WANG et al. « Phase imaging with an untrained neural network ». In : Light : Science & Applications 9.1 (2020), p. 77 (cf. p. 173).
- [Wan+21] Yirui WANG, Fengyi JIANG, Guohao JU et al. « Deep learning wavefront sensing for fine phasing of segmented mirrors ». In : Optics Express 29.16 (2021), p. 25960-25978 (cf. p. 37, 38).
- [Wer05] Paul J WERBOS. « Applications of advances in nonlinear sensitivity analysis ». In : System Modeling and Optimization. Springer. 2005, p. 762-770 (cf. p. 84).

- [Wu09] Huaiqin WU. « Global stability analysis of a general class of discontinuous neural networks with linear growth activation functions ». In : Information Sciences 179.19 (2009), p. 3432-3441 (cf. p. [81](#)).
- [Xu+19] Zhenxing XU, Ping YANG, Ke HU et al. « Deep learning control model for adaptive optics systems ». In : Applied optics 58.8 (2019), p. 1998-2009 (cf. p. [37](#)).
- [Zou+19] Fangyu ZOU, Li SHEN, Zequn JIE et al. « A sufficient condition for convergences of adam and rmsprop ». In : Proceedings of the IEEE/CVF Conference on computer vision. 2019, p. 11127-11135 (cf. p. [83](#)).

ANNEXES

A. Analyse des paramètres d'un réseau de neurones convolutif

Dans un premier temps, on rappelle l'architecture utilisée pour l'inférence des 4 paramètres (tip et tilt pour chacune des deux pupilles).

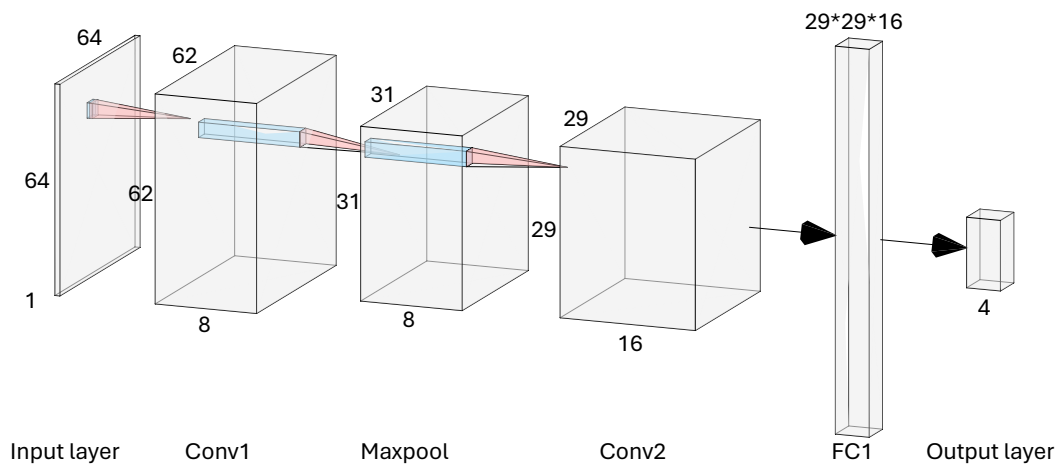
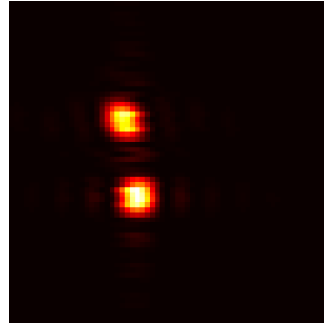


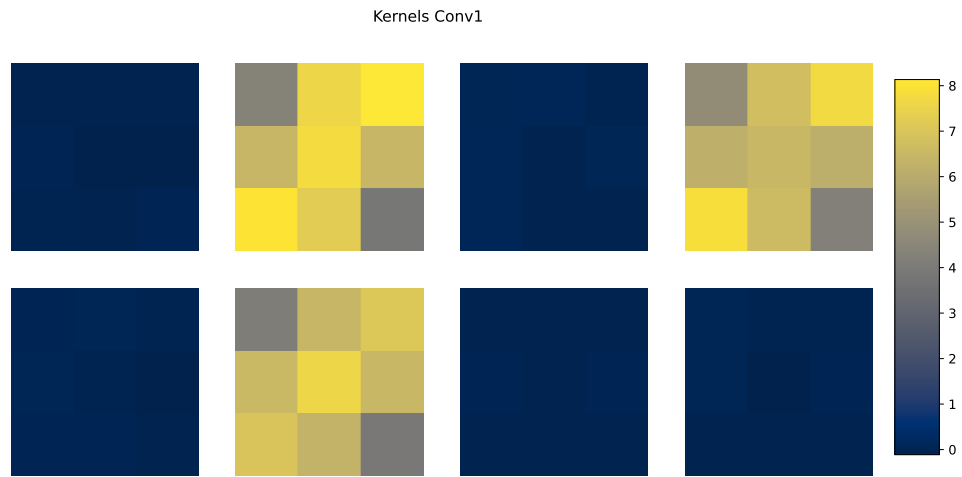
FIGURE A.1. – Architecture du réseau de neurones convolutif.

L'architecture présente la couche d'entrée (l'image de la PSF, 64×64), puis une couche convolutive composée de 8 filtres 3×3 , retournant 8 images 62×62 . Ces images passent ensuite par une fonction d'activation ReLu puis sont transformées en images 31×31 par l'opération de maxpooling. Les 8 images de sortie du maxpooling sont concaténées afin de passer par la prochaine couche convolutive, qui retourne alors 16 imagerie 29×29 . Ces imagerie sont alors mises sous forme de vecteur afin de passer à travers la couche entièrement connectée qui donne alors la couche de sortie composée de 4 neurones, correspondant aux coefficients à estimer. Maintenant que le processus d'inférence a été montré, on propose de prendre comme exemple une PSF et d'étudier les transformations effectuées au sein du réseau présenté afin de comprendre les étapes de l'estimation.

A. Analyse des paramètres d'un réseau de neurones convolutif



(a) PSF d'entrée.

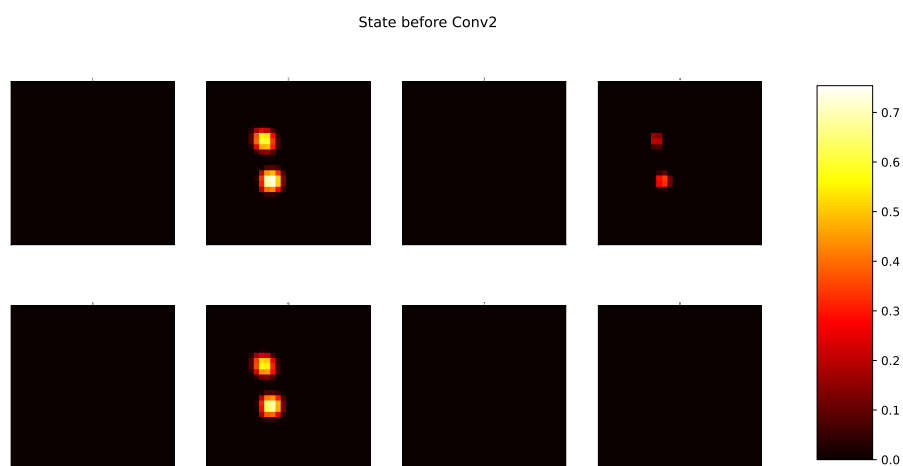


(b) Kernels de convolution de la couche Conv1.

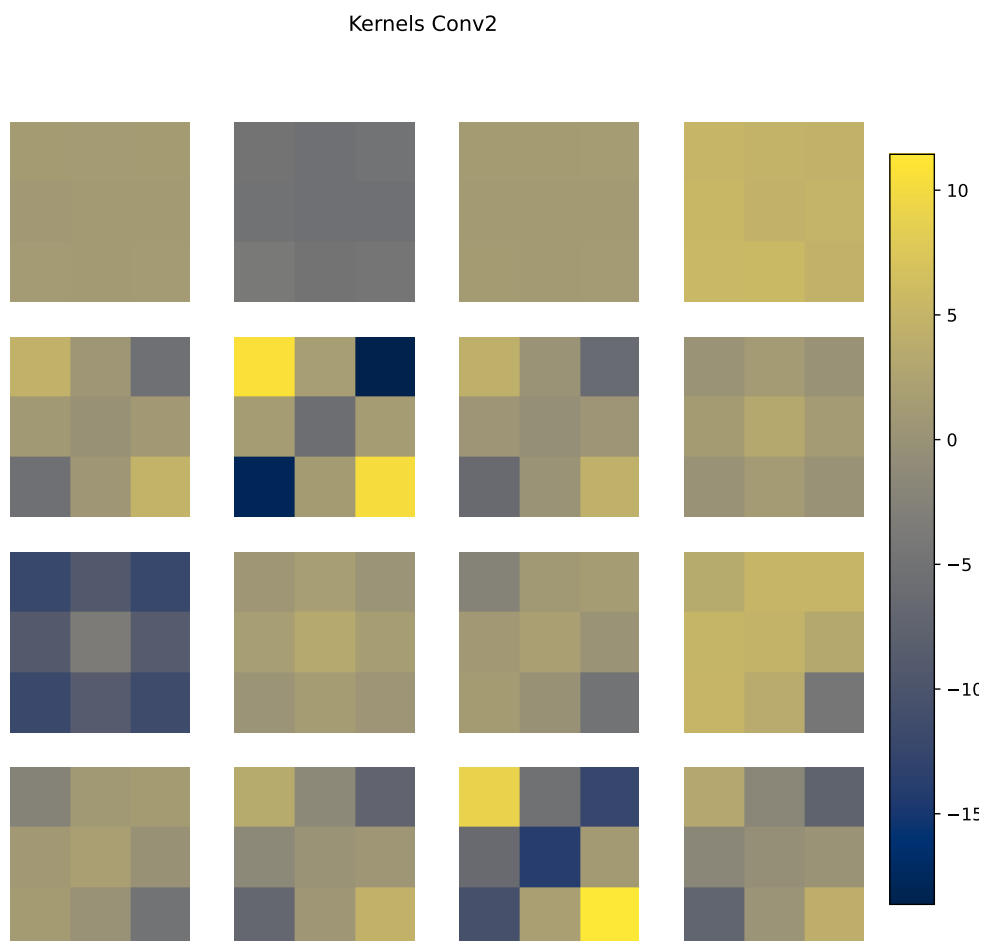
FIGURE A.2. – Exemple de la PSF d'entrée et des Kernels de la Conv1.

La figure A.2 montre la PSF d'entrée et les filtres de convolution appris par la couche Conv1. On peut remarquer directement que 5 filtres semblent de très petite intensité, donc que la couche convolutive semble posséder trop de filtres par rapport aux dimensions du problème. Cependant, des tests non concluants ont été effectués en réduisant le nombre de filtres de la Conv1 à 3. Néanmoins, cette première couche de convolution permet d'extraire de l'information spatiale utile sur les PSFs, et retourne alors la couche suivante après l'application de la couche d'activation et de maxpooling :

A. Analyse des paramètres d'un réseau de neurones convolutif



(a) Imageries en entrée de la couche Conv2.



(b) Kernels de convolution de la couche Conv2.

FIGURE A.3. – Exemple des imageries en entrée de la Conv2 et des kernels de la Conv2.

A. Analyse des paramètres d'un réseau de neurones convolutif

Les kernels présentés en figure A.3 vont être appliqués à chacune des imageries de la première figure de la figure A.3. De même que pour la première couche convolutive, certains kernels de très petite amplitude semblent être inutiles à l'inférence des paramètres ; cependant, des tests avec deux couches convolutives de respectivement 3 et 6 filtres n'ont pas montré de résultats concluants. Un jeu de données plus conséquent couplé à une optimisation des hyperparamètres pourrait cependant afficher des résultats satisfaisants avec ce nombre optimal de filtres. Néanmoins, on peut noter un motif à symétrie de révolution au niveau des filtres de convolution, chacun étant d'amplitude différente. L'extraction de caractéristiques des imageries semble suivre un motif bien défini mais difficilement interprétable. Pour la prochaine couche, 16 filtres sont appliqués par sous-imageries, le résultat de chacun des filtres est ensuite concaténé de façon à ce que la somme des images issues des i -èmes filtres corresponde à l'imagerie i .

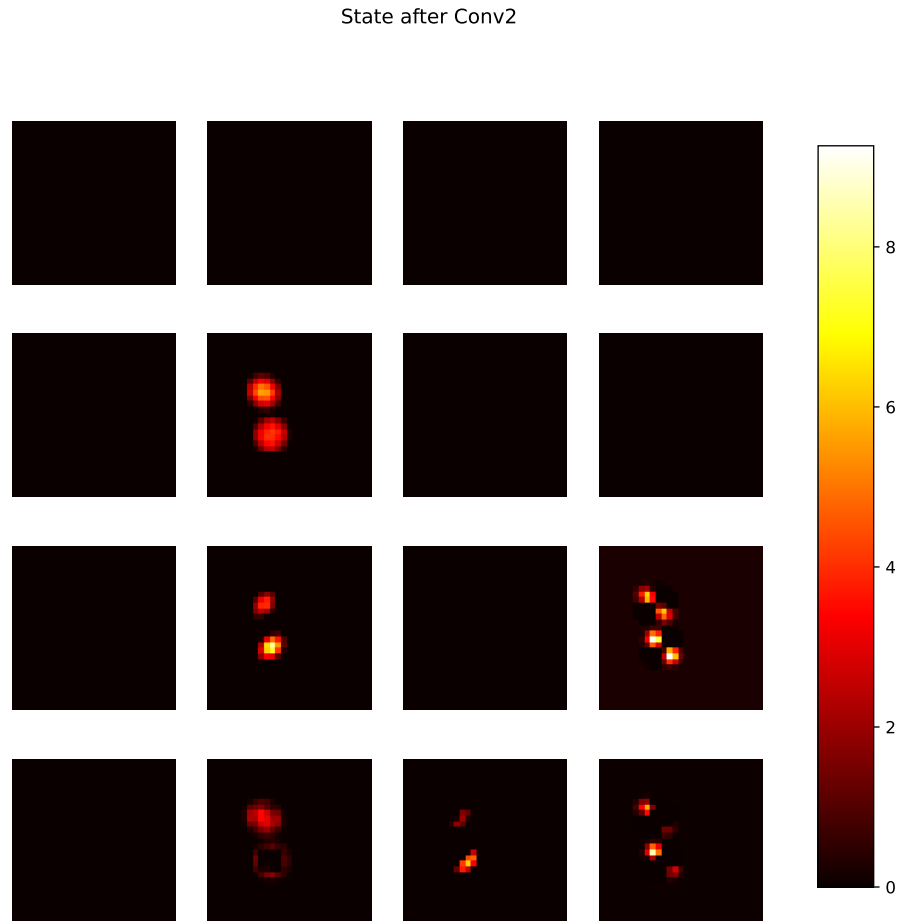


FIGURE A.4. – Imageries de sortie de la couche Conv2.

A. Analyse des paramètres d'un réseau de neurones convolutif

La figure A.4 présente l'exemple des 16 imageries de sortie de la couche Conv2 une fois la couche d'activation calculée, appliquée aux imageries concaténées par filtre. Ici l'extraction de caractéristiques à partir des images est calculée, on retrouve un set de 16 imageries, qui vont ensuite être linéarisées avant de passer par une couche entièrement connectée qui viendra alors estimer les coefficients de tips et de tilts. Pour rappel, la couche entièrement connectée contient $16 \times 29 \times 29$ neurones. Cependant, 4 coefficients sont à estimer, chaque neurone comprend alors 4 poids et 4 biais distincts. Tous les poids de la couche entièrement connectée ne sont pas forcément exploitables :

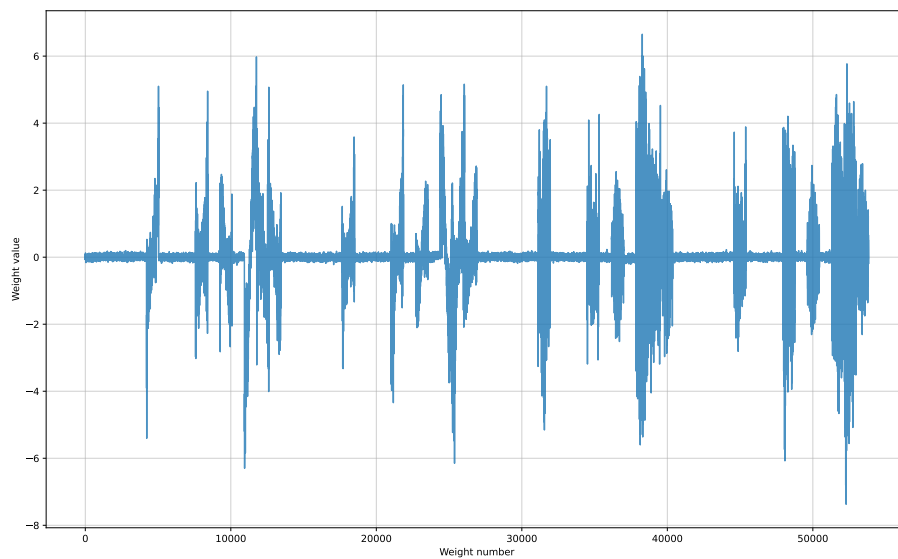


FIGURE A.5. – Ensemble des poids de la couche entièrement connectée.

Les poids dont la valeur est proche de 0 sur la figure A.5 représentent majoritairement des poids non appris par le réseau, considérés alors comme du bruit. En repérant alors uniquement les neurones contenant du signal autre que du bruit, et en représentant les poids en 2 dimensions, on obtient alors les cartes des poids suivantes :

A. Analyse des paramètres d'un réseau de neurones convolutif

Weights FC1

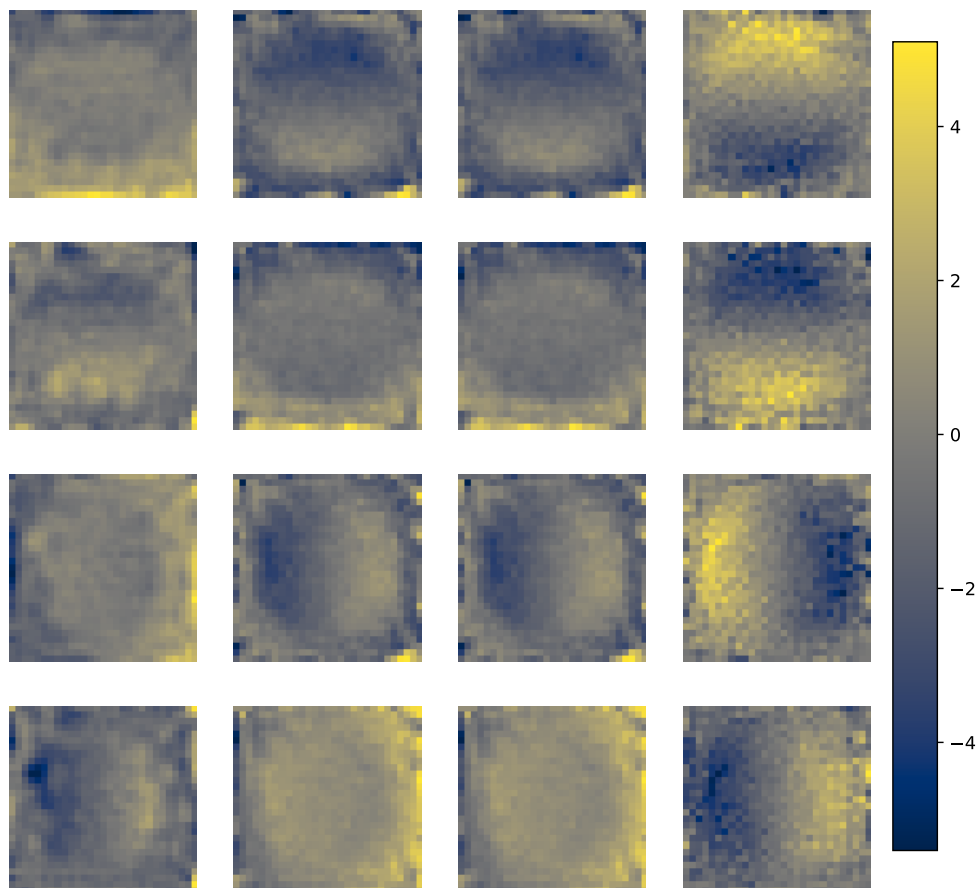


FIGURE A.6. – Poids de la FC1 contenant du signal exploitable.

On retrouve des motifs très familiers ici : les motifs du centre de gravité pour les tips et tilts sont présents sur chacune des cartes de poids. Une fois appliqué aux imagerie contenant du signal significatif, on peut alors obtenir la dernière carte d'intensité :

A. Analyse des paramètres d'un réseau de neurones convolutif

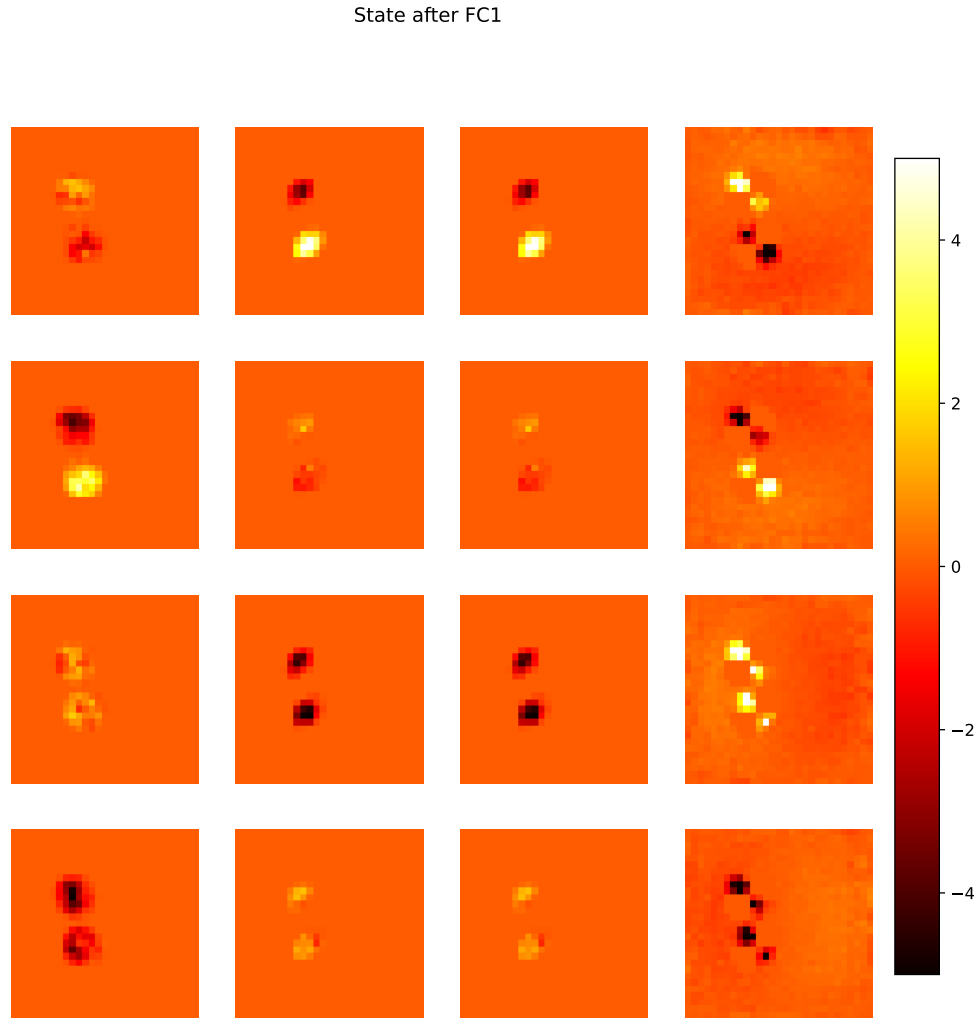


FIGURE A.7. – Multiplication pixel à pixel des poids avec les imagerie correspondantes.

La figure A.7 présente le résultat de la multiplication pixel à pixel entre les poids exploitables et les imagerie correspondantes réordonnées en images 2 dimensions. Afin de retrouver les coefficients de la phase à partir de ces images, il suffit alors de sommer l'ensemble des pixels des images d'une même ligne. À en déduire par l'allure des poids de la figure A.6, les coefficients sont donc estimés dans l'ordre Tip1, Tip2, Tilt1, Tilt2, correspondant à l'ordre des coefficients établit lors de la mise en place de la méthode.

Table des figures

1.	Exemple de télescope spatiaux avec étude de leurs résolution par rapport à leur taux de revisite. Extrait de	16
2.	Exemple de cubesat. Extrait de	18
3.	Exemple de tâche d'Airy en échelle racine carrée.	21
4.	Schéma de l'influence de l'atmosphère au niveau d'une orbite basse.	23
5.	Exemple de PSF aberrée (gauche) et de la FTO correspondante (figure de droite en trait plein), comparée à la FTO théorique du télescope parfait (figure de droite en pointillé).	24
6.	Illustration de l'influence d'un focus positif et négatif sur l'image de la PSF sur le détecteur.	26
7.	Schéma de la mise ne place de la diversité de phase. Extrait de la thèse d'Amandine Blanc.	27
8.	Schéma du système d'optique adaptative. Extrait de [RF22].	29
9.	Exemple d'ASOs. Extrait de [RF22].	30
10.	Sensibilité de la Pyramide modulée à une discontinuité de phase (largeur du spider). Sans modulation la pyramide reste très peu sensible aux discontinuités. Extrait de la thèse de N.Levraud.	31
11.	Illustration des simulations de l'impact sur les sous-PSFs de l'ajustement du focus du JWST. Extrait de Acton [Act+22].	33
12.	Illustration du centrage itératif des sous-PSFs du JWST, corrigeant le tip et le tilt de chacun des segments. Extrait de Acton [Act+22].	34
13.	Illustration des méthodes de correction des pistons du JWST. Extrait de Acton [Act+22].	35
14.	Illustration de la diversité de phase à 6 longueurs d'onde appliquée au JWST. Extrait de Acton [Act+22].	35
15.	Illustration des différentes méthodes d'analyse du front d'onde (IS pour <i>Image Sharpening</i> , PD pour <i>Phase Diversity</i> , ES pour <i>Edge Sensor</i> , HDFS pour <i>Holographic Dispersed Fringe Sensor</i> , AO pour <i>Adaptive Optics</i>).	36
16.	Télescope spatial deployable AZIMOV. Extrait de Schwartz [Sch+20].	40
17.	Schéma optique du télescope spatial deployable AZIMOV. Extrait de Schwartz [Sch+20] et [Sch+22].	40
18.	Exemple de la stratégie de phasage d'AZIMOV. Extrait de Schwartz [Sch+22].	43
19.	Exemple du petaling de l'ELT. Les zones grisées représentent les araignées de support du M2. Extrait de la thèse de N.Levraud.	45
1.1.	Représentation des 21 premiers modes de Zernike. Extrait de	49

1.2. Illustration de pupille circulaire et de forme d'AZIMOV, avec et sans obstruction centrale avec leurs PSF et FTO correspondantes.	50
1.3. Schéma du système optique d'AZIMOV.	51
1.4. Illustration du modèle choisi de la pupille AZIMOV symétrique et asymétrique.	52
1.5. Comparaison de l'impact d'un focus sur une pupille symétrique et asymétrique sur les PSFs (en échelle racine).	54
1.6. Illustration de l'impact du choix des matériaux sur la déformation du miroir par des gradients thermiques. Extrait du rapport de stage de L. Bouron (2021).	58
1.7. Exemple d'aberration de haut ordres, additionnelle au Piston-Tip-Tilt sur la pupille AZIMOV asymétrique.	59
1.8. Exemple de PSFs typique d'AZIMOV, pour les différents scenario évoqués et leur FTO correspondante. La première image (haut gauche) représente la PSF parfaite du télescope, la PSF à la limite de la diffraction. La seconde image représente la PSF dans le cas du phasage fin. La figure en haut à droite représente la PSF soumise aux erreurs de phasage de déploiement. La figure en dessous représente les moyennes azimutales des FTOs correspondantes.	62
1.9. Exemple de distribution du WFE sur un dataset de 10^5 PSFs. A gauche la distribution du régime de phasage fin, avec un WFE moyen de 64 nm . A droite la distribution pour le régime de déploiement avec un WFE moyen de 320 nm	63
1.10. Exemple de la relation entre piston et intensité des pixels (ici le pixel central de l'image est choisi). A gauche 2 PSF similaires sont présentées. La PSF du bas représente une PSF générée avec le piston du segment 1 à une valeur de $+\lambda \text{ nm PV}$ par rapport au piston 1 de la PSF du haut. Le graphique de droite représente l'évolution de l'intensité du pixel central de la PSF en fonction du piston 1.	65
1.11. Exemple d'objet (colonne de gauche) et d'images correspondante (colonne de droite).	68
1.12. Exemple d'objet observable et du contenu spectral une fois imagé par le télescope non phasé.	70
1.13. Exemple d'objet point source étendu.	71
2.1. Illustration d'un perceptron (gauche) et comparaison avec un neurone biologique (droite). Extrait des cours de M.Gray.	77
2.2. Illustration d'un MLP. Dans cet exemple, les poids sont tirés aléatoirement et sont représentés en couleur et en intensité en fonction de leur amplitude. De gauche à droite : couche d'entrée, couche cachée, couche cachée, couche de sortie.	78
2.3. Illustration de la convolution (produit d'Hadamard) d'une image par un kernel. Extrait de FIDLE.	79
2.4. Exemple de max pooling. Extrait de FIDLE.	81

2.5. Exemple des fonctions d'activation usuelles.	82
2.6. Schéma de l'architecture VGG-Net utilisée.	86
2.7. Schéma d'une skip connection. Extrait de AI summer.	87
2.8. Schéma de l'architecture ResNet utilisée.	88
2.9. Exemple d'un bloc MBConv. Extrait de ResearchGate.	89
2.10. Exemple convolution transposée. Extrait de FIDLE.	92
2.11. Principe de l'architecture d'un VAE.	92
3.1. Architecture d'un réseau de neurones entièrement connectée à une seule couche d'entrée entièrement connecté à un seul neurone de sortie sans couche cachée.	99
3.2. Exemple de PSF en entrée et carte de poids	100
3.3. Architecture d'un réseau à une couche entièrement connectée et deux neurones de sortie.	101
3.4. Carte des poids de la couche entièrement connectée.	102
3.5. Répartitions des 1000 PSFs d'entraînement.	103
3.6. Exemple de PSF en entrée et carte de poids	104
3.7. Carte des poids de la couche entièrement connectée et zoom sur la bande centrale.	104
3.8. Exemple de multiplication terme à terme de la carte de poids et de deux PSFs (échelle racine).	105
3.9. Carte des poids de la couche entièrement connectée.	106
3.10. Exemple de PSF en entrée du réseaux, I , et des PSFs estimées (reconstruites $\hat{I}$ à partir des coefficients estimés par le réseaux).	107
3.11. Illustration de la minimisation de la moyenne du WFE par $tip_1 = tip_2$ et $tilt_1 = tilt_2$	108
3.12. Vue du dessus de l'erreur pour les tip-tilts égaux. Un seuillage est effectué à une erreur de 500 nm pour apprécier la dynamique des plus petites erreurs de front d'onde.	109
3.13. Architecture pour le cas 4 coefficients composée de 2 couches convolutives (le 3eme bloc représente la couche de max-pooling uniquement) et une couche entièrement connectée.	110
3.14. Numérotation des segments de la pupille AZIMOV.	112
3.15. Décomposition de l'erreur de la phase reconstruite en fonction de sa partie paire (rouge) et impaire (bleu).	113
3.16. Front d'onde résiduel en fonction de la surface d'asymétrie de la pupille pour la loss d'entraînement et la loss de validation.	114
3.17. Comparaison entre VGG-net (courbes orangées) et Resnet (courbes vertes) entraînés sur données brutes et testés sur données non bruitées (trait plein) et données bruitées (pointillés).	115
3.18. Comparaison de la qualité d'estimation du piston pour des réseaux entraînés avec des jeux de données correspondant à différentes distributions uniformes de piston.	117

3.19. Histogramme des erreurs de piston pour le réseau entraîné sur des PSFs dont les pistons varient entre $[-75, 75]$ nm.	118
3.20. Comparaison de la qualité d'estimation des réseaux entraînés sans bruit (pointillé) ou entraînés avec des données bruitées (trait plein) en fonction du SNR.	119
3.21. Robustesse des réseaux en fonction du SNR. La courbe verte foncé illustre la qualité de l'estimation du NN5; la courbe verte clair celle du NN3.	121
3.22. Robustesse des réseaux a des aberrations de plus haut ordre. En vert les courbes des deux scénarios identifiés. En rouge les réseaux entraînés sans HO testé avec 60 nm de haut ordres.	122
3.23. Comparaison de la propagation de bruit de la Diversité de Phase à $c^{PD} = 10^{-15}$ (courbe noire), de l'Image Sharpening (courbe marron) et des réseaux NN3 (courbe verte claire) et NN5 (courbe verte foncée).	126
3.24. Comparaison du temps de calcul et de la performance des 3 méthodes.	127
3.25. Schéma de la stratégie de phasage complet du télescope.	128
3.26. Résultats des réseaux NN3 et NN5 en boucle fermée.	129
3.27. Exemple de la pupille et d'une carte d'OPD, avec la PSF correspondante pour la pupille du JWST.	131
3.28. Distribution du front d'onde PV total sur la pupille JWST.	132
3.29. Amplitude des erreurs d'estimation pour un front d'onde $\sigma = 30$ nm PV.	133
3.30. Amplitude des erreurs d'estimation pour un front d'onde $\sigma = 70$ nm PV.	134
3.31. Illustration de la pupille ELT segmentée par les spiders et PSF associée à la sous-ouverture haut-gauche du Shack-Hartmann. Extrait de [Cha+23]	136
3.32. Résultats d'estimation des pistons différentiels par réseaux de neurones présenté pour les 2 sous pupilles du haut.	138
4.1. Schéma de la maquette optique d'AZIMOV. Extrait du reporting du banc de UK-ATC.	141
4.2. Image du banc. Extrait du reporting du banc de UK-ATC.	141
4.3. Distribution des coefficients des PSFs expérimentales.	143
4.4. Exemple de PSFs expérimentales.	144
4.5. Illustration de la FTO AZIMOV comparée à une FTO sur pupille circulaire sans obstruction centrale.	145
4.6. Ratio de la position relative β en fonction de l'échantillonnage α pour 3 angles θ	145
4.7. Ratio de la position relative β en fonction de l'échantillonnage α pour 3 angles θ	146
4.8. Estimation de l'échantillonnage $\hat{\alpha}$	147
4.9. Comparaison des PSFs expérimentales (gauche), issues du simulateur (milieu), et jumeaux numérique (droite) pour les régimes de phasage fin (haut) de déploiement (bas).	149
4.10. Résultats de l'estimation des réseaux R-Exp et R-Sim sur les jeux de données de test de phasage fin.	151

4.11. PSF expérimentale (gauche), reconstruite avec les bons coefficient (milieu), reconstruite à partir des coefficients estimés (gauche).	152
4.12. Résultats de l'estimation des réseaux R-Exp et R-Sim sur les jeux de données de test du régime de déploiement.	153
4.13. Comparaison des estimations pour données simulées et expérimentales pour 100 PSFs.	154
4.14. Comparaison de l'erreur d'estimation des réseaux en fonction du nombre de donnée utilisée pour le <i>fine-tuning</i>	155
4.15. Comparaison de l'erreur d'estimation des réseaux en fonction du nombre de couches convolutives libérées.	156
4.16. Évolution de l'erreur d'estimation en fonction du nombre de données de fine-tuning.	157
5.1. Exemple d'images acquises en orbite basse.	162
5.2. Exemple de formation d'image d'un objet étendu.	163
5.3. Exemple de formation d'image de lignes.	164
5.4. Exemple d'objets observables.	165
5.5. Exemple d'objets étendus simulés.	166
5.6. Exemple d'objets étendus simulés par VAE. Les objets réels sont présentés à la première ligne tandis que ceux générés par le VAE sont présentés à la seconde. On retrouve une coupe du centre des objets à la dernière ligne.	168
5.7. Exemple d'objets étendus simulés par VAE présentés à la première ligne tandis que ceux générés par le VAE avec un contraste augmenté sont présenté a la deuxième ligne. On retrouve une coupe du centre des objets plus contrasté à la dernière ligne.	169
5.8. Loss d'apprentissage et de validation en fonction du nombre d'époque pour un unique objet gaussien (courbes vertes) et un unique objet issus du VAE (courbes oranges), et distribution des WFE d'entrée (bleu clair).	170
5.9. Loss d'apprentissage et de validation en fonction des époques pour 100k objets différents gaussien (courbes vertes) et issus du VAE (courbes oranges).	172
5.10. Schéma de la stratégie d'apprentissage du réseau de neurones pour l'objet étendu.	173
5.11. Courbes d'entraînement sur point source à objet non connu (no KO) et objet connu (KO). L'apprentissage du réseau à objet non connu s'arrête à 300 époques afin d'éviter le sur-apprentissage.	175
5.12. Courbe de validation lors de l'apprentissage pour un objet étendu gaussien sur fond nul.	177
5.13. Courbe de validation pour un objet étendu gaussien sur fond nul (courbes en pointillées) et sur objet issu du VAE (courbe en trait plein).	178
5.14. Exemple de contraste des objets étudiés.	179

5.15. Étude de la qualité d'estimation du PINN (colorbar) en fonction du contraste de l'image et du front d'onde incident pour le réseau entraîné sur les données issues du VAE.	180
5.16. Étude de la qualité d'estimation du PINN (colorbar) en fonction du contraste de l'image et du front d'onde incident pour le réseau entraîné sur les données issues du VAE à contraste augmenté.	181
5.17. Courbe de l'évolution de la sensibilité en fonction du contraste. La courbe verte représente les variations pour un petit contraste, la courbe orange pour un contraste moyen, et enfin la courbe rouge pour un contraste élevé. La courbe en pointillé représente le cas pour un point source, objet dont le contraste est maximal.	182
5.18. Schéma de la boucle fermée sur objet étendu.	184
5.19. Exemple de boucle fermée pour le phasage temps réel d'AZIMOV. . . .	185
A.1. Architecture du réseau de neurones convolutif.	198
A.2. Exemple de la PSF d'entrée et des Kernels de la Conv1.	199
A.3. Exemple des imagerie en entrée de la Conv2 et des kernels de la Conv2.	200
A.4. Imagerie de sortie de la couche Conv2.	201
A.5. Ensemble des poids de la couche entièrement connectée.	202
A.6. Poids de la FC1 contenant du signal exploitable.	203
A.7. Multiplication pixel à pixel des poids avec les imagerie correspondantes.	204

Acronymes

AFOPF

Analyse de Front d'Onde Plan-Focal. [25](#)

AO

Adaptive Optics. [28](#)

ASO

Analyseur de Surface d'Onde. [28](#)

AZIMOV

AutomatiZatIon of MOVable optics. [39](#)

BCE

BinaryCross-Entropy. [90](#)

DFS

Dispersed Fringe Sensor. [33](#)

DL

Diffraction Limit. [20](#)

DM

Deformable Mirror. [28](#)

ELT

Extremely Large Telescope. [14](#)

EO

Extended Object. [173](#)

ES

Edge Sensors. [32](#)

FTO ou OTF

Fonction de Transfert Optique. [24](#)

GMT

Giant Magellan Telescope. [32](#)

GSD

Ground Sampling Distance. [18](#)

HO

Higher Order aberrations. [57](#)

IHRA

Imagerie à Haute Résolution Angulaire. [14](#)

IS

Image Sharpening. [28](#)

JWST

James Webb Space Telescope. [15](#)

KO

Known Object. [174](#)

M1

Miroir primaire. [51](#)

ML

Machine Learning. [36](#)

MLP

Multi-Layer Perceptron. [37](#), [77](#)

MSE

Mean Square Error. [90](#)

OPD

Optical Path Difference. [57](#)

PD

Phase Diversity. [26](#)

PINN

Physics Informed Neural Network. [174](#)

PSD

Densité Spectrale d'Energie, Power Spectral Density. [70](#)

PSF

Point Spread Function. [20](#)

PTT

Piston, Tip, Tilt. [49](#)

PV

Peak to Valley. [117](#), [131](#)

R-Sim

Réseau entraîné sur données Simulées. [150](#)

R-Exp

Réseau entraîné sur données Expérimentales. [150](#)

ReLu

Regularized Linear Unit. [82](#)

RTC

Real Time Computer. [28](#)

SH

Shack-Hartmann. [30](#)

SLM

Spatial Light Modulator. [142](#)

SNR

Signal to Noise Ratio. [59](#)

TLU

Threshold Logic Unit. [76](#)

VAE

Variational Auto-Encoder. [85](#)

WFE

WaveFront Error. [56](#)

Article



Phasing segmented telescopes via deep learning methods: application to a deployable CubeSat

MAXIME DUMONT,^{1,2,3,*} CARLOS M. CORREIA,^{2,4} JEAN-FRANÇOIS SAUVAGE,^{1,3}
NOAH SCHWARTZ,⁵ MORGAN GRAY,³ AND JAIME CARDOSO^{2,6}

¹DOTA, ONERA, F-13661 Salon Cedex AIR, France

²Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal

³Aix-Marseille Université, CNRS, CNES, LAM, Marseille, France

⁴Center for Astrophysics and Gravitation, Instituto Superior Técnico, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal

⁵UK Astronomy Technology Centre, Edinburgh EH9 3HU, UK

⁶INESCTEC, Porto, Portugal

*maxime.dumont@lam.fr

Received 19 September 2023; revised 22 January 2024; accepted 22 January 2024; posted 25 January 2024; published 20 February 2024

Capturing high-resolution imagery of the Earth's surface often calls for a telescope of considerable size, even from low Earth orbits (LEOs). A large aperture often requires large and expensive platforms. For instance, achieving a resolution of 1 m at visible wavelengths from LEO typically requires an aperture diameter of at least 30 cm. Additionally, ensuring high revisit times often prompts the use of multiple satellites. In light of these challenges, a small, segmented, deployable CubeSat telescope was recently proposed creating the additional need of phasing the telescope's mirrors. Phasing methods on compact platforms are constrained by the limited volume and power available, excluding solutions that rely on dedicated hardware or demand substantial computational resources. Neural networks (NNs) are known for their computationally efficient inference and reduced onboard requirements. Therefore, we developed a NN-based method to measure co-phasing errors inherent to a deployable telescope. The proposed technique demonstrates its ability to detect phasing errors at the targeted performance level [typically a wavefront error (WFE) below 15 nm RMS for a visible imager operating at the diffraction limit] using a point source. The robustness of the NN method is verified in presence of high-order aberrations or noise and the results are compared against existing state-of-the-art techniques. The developed NN model ensures its feasibility and provides a realistic pathway towards achieving diffraction-limited images. © 2024 Optica Publishing Group

<https://doi.org/10.1364/JOSAA.506182>

1. CONTEXT AND MAIN CHALLENGES

The angular resolution of telescopes is fundamentally limited by the size of their aperture for a given wavelength. For example, a telescope with a diameter of 30 cm has a ground sampling of 1 m at a distance of 400 km or 2 m at 800 km at visible wavelengths ($\lambda = 800$ nm). In the fields of Earth observation and astronomy, high-angular-resolution images are crucial for maximizing the scientific output and return on investment, which has driven the demand for larger apertures. However, launching large monolithic pupils into space is expensive, and designing and manufacturing such telescopes pose significant challenges [1]. An effective solution to increase aperture size is to fragment the aperture into smaller segments. By folding the telescope inside a compact volume during launch and deploying it in orbit, both the collecting power and angular resolution can be increased. This volume optimization and gain in mass translate directly into cost savings, enabling the deployment of a constellation with multiple high-angular-resolution platforms and therefore achieving higher revisit rates (i.e., temporal resolution).

Improving both the revisit rate and angular resolution finds applications in Earth climate monitoring, defense, and security for which the highest commercially available spatial resolutions today are given by the satellite Pleiades [with a ground sampling distance (GSD) of 70 cm, at a rate of one image per day] as well as Pleiades-Neo (30 cm GSD color images). For example, flood monitoring (accounting for over 40% of the natural disasters) currently provides information either over large areas or with limited spatial resolutions and a revisit every two days (e.g., Copernicus with a 5 km resolution), or instead with a revisit time at best of a few days yet with a high spatial resolution (e.g., Pléiades, Sentinel-1, and Sentinel-2). Solar system exploration can potentially also benefit from this improvement: although the HRI camera of Deep Impact has an aperture larger than 30 cm, the typical size is 10 to 20 cm (NAC camera on Rosetta, Cassis camera on TGO, New Horizon camera).

The objective of this study is to enable, in ESA's own parlance *very very high-angular-resolution* imaging in the visible spectrum for Earth observation from space at a low cost by

quantifying the performance of a NN-based phasing algorithm during pre-launch, deployment, and exploitation phases. To achieve this, we use AZIMOV [2] as a reference design for our study. AZIMOV is a satellite concept featuring a 30 cm wide, diluted aperture telescope onboard a 6U CubeSat. It is capable of reaching a 1 m GSD in the visible (i.e., $\text{GSD} = \lambda / D * z = 800 \cdot 10^{-9} / 0.3 \times 400000 \approx 1 \text{ m}$) from LEO with a swath of over 4 km. Figure 1 illustrates a simplified concept of the payload. This technology allows the telescope's aperture to surpass the size of the supporting platform, thereby enhancing the telescope's light-gathering power and improving its achievable resolution.

To achieve the theoretical angular resolution for a given segmented pupil, the individual mirror segments must be co-phased with a precision typically on the order of a fraction of the imaging wavelength (usually $\lambda/14$ [3]). Compensation for static and dynamic phase aberrations is required for each segment. High-precision actuators are utilized to control the position of the deployable primary mirror segments in piston and tip-tilt [4] (PTT). For visible wavelength imaging at the metric resolution limit (we select $\lambda = 800 \text{ nm}$ as the study wavelength), the wavefront error (WFE) after correction should be less than 50 nm RMS. However, due to actuator response and dynamics, a portion of the WFE error is independent of sensing. From a dedicated study of this error budget, Sauvage *et al.* [5] allocate $\epsilon_{\text{WFS}} < 15 \text{ nm RMS}$ to the sensing and reconstruction contribution. Our study proposes a solution to perform accurate WFS onboard a CubeSat, using directly a single focal plane image and deep learning.

This paper is organized as follows: we first provide an introduction to the study and the motivation behind it. Second we delve into the concepts of image formation in the context of wavefront sensing. The methodology employed in our research is detailed in Section 4. Then, Section 5 presents the results obtained from our experiments, analyzing the performance

and robustness of the proposed methods. Section 6 discusses the implementation of neural networks in a complete phasing process specifically tailored for the AZIMOV system. Finally, Section 7 concludes the paper.

2. PHASING SEGMENTED MIRROR TELESCOPES

The phasing of segmented mirror telescopes is a critical aspect of their operation, as it directly impacts the quality of their optical performance. Ground-based telescopes such as the segmented, twin Keck telescopes correct for tip-tilt errors using a Shack-Hartmann wavefront sensor (SH-WFS) [6]. However, the SH-WFS is known to be insensitive to global piston error. To overcome this issue, a micro-lenslet array is placed at the junction of two segments on the primary mirror, and allows to identify the differential piston between two segments [6]. More recently, for the Giant Magellan Telescope, and following roughly the same principle, a differential piston between segments will be identified using a holographic dispersed fringe sensor (HDFS) [7]. It employs a holographic optical element to disperse the incoming light into multiple wavelength channels. By analyzing the interference patterns created by these channels, the HDFS can accurately measure the differential piston. For space-borne telescopes such as the JWST [8], the phasing strategy is divided into three parts: i) segment location-identification, to identify the coarse position of each segment; ii) co-phasing to correct the global phasing error of the 18 segments; iii) finally, wavefront monitoring and maintenance to ensure a correct phasing throughout the life of the mission. The coarse and fine phasing is ensured by a dispersed fringe SH-WFS and weak lens hardware [9]. In this context, this step-by-step approach has inspired our work. However, we propose an innovative approach to co-phase segmented telescopes by using a single focal plane image as the observable, and deep learning to tune our NN-based algorithm. Single-image approaches can greatly simplify the phasing process and enhance the optical performance of future segmented telescopes. Our work focuses specifically on the primary mirror phasing from the point spread functions (PSFs) obtained while imaging a point source object. The use of a point source is a key element in our approach: first to understand the behavior and NN specifics to process focal plane images and estimate wavefront errors from them, and second, to be able to generalize to extended objects. Lastly, the phasing process of AZIMOV is not defined yet. One strategy could be to undertake the coarse phasing imaging of a distant unresolved star and then rotate the telescope towards the Earth to proceed with the fine phasing and capture images for further ends. The use of NN on point sources is therefore central to many applications. Our work aims to employ a computationally efficient phasing method taking into account the typical CubeSat limitations especially in terms of volume and computing power. In particular, we demonstrate the potential of using NN in contrast to more classical optimization methods, focusing on execution time as a proxy to computational complexity. In this context, we use the time required for NN inference to identify a suitable NN model capable of reaching the diffraction limit with minimal computational burden. The AZIMOV payload is currently being designed, and requirements for the

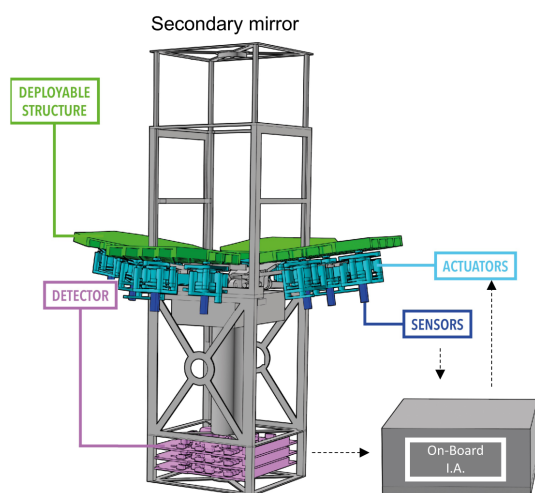


Fig. 1. Simplified illustration of the deployable payload concept: deployable structures (primary and secondary mirrors, baffle), actuators to adjust mirror positions, sensors to measure mirror positions, detector to assess image quality, and onboard artificial intelligence to control and adjust mirror positions to reach diffraction-limited image quality (i.e., active optics). From Schwartz [2].

onboard computer (OBC) are still being finalized. Among other aspects, the OBC will need to run the compute intensive co-phasing algorithm, the data-intensive image processing algorithm (including algorithms such as image registration and super-resolution), and other housekeeping tasks. The satellite being in LEO, it will receive a varying thermal load (approx. every 90 min): multiple updates per orbit will be required to maintain good phasing performance. It is important to minimize the co-phasing algorithm complexity and therefore execution time. The exact execution time will naturally depend on the selected OBC.

A. Classical Wavefront Sensing Approaches

Various methods have been proposed for measuring and estimating wavefront error, broadly classified into two categories: direct wavefront sensing and focal plane methods. Direct wavefront sensing relies on a relatively simple algorithm, at a cost of an optical path specific to the wavefront sensing with dedicated hardware, such as the JWST's dispersed fringe Shack-Hartmann WFS [8].

An alternative method involves indirect measurement, where the wavefront is deduced from the focal plane image captured by the detector. One approach is image sharpening, as proposed in the work of Schwartz [2,10], which entails acquiring a sequence of images and iterating starting from an initial guess of the aberrations. For each of these images, an image quality metric, such as maximum intensity or image contrast, is computed and then used in a non-linear iterative optimization algorithm to estimate the aberration. In contrast, model-based focal plane wavefront sensing methods typically rely on introducing diversity (whether in phase, e.g., phase diversity [11], wavelength, or amplitude) to resolve sign-ambiguity [12] of even modes. A best fit is sought to the focal plane images where from the wavefront aberration parameters are estimated. The algorithms are typically more complex than those used in direct wavefront sensing due to the relationship between the image intensity in the focal plane and the electric field in the pupil plane. Nevertheless, considering the constraints posed by Earth observation and the non-repeatability of scenes, the goal is to conduct focal plane wavefront sensing using only a single image. Data-driven, deep-learning-based methods can also be applied successfully in focal plane wavefront sensing (FPWFS). One of the main advantages lies in the fact that once the network is trained, the model directly outputs the estimated wavefront without requiring iterations. The computation gains time and, consequently, efficiency in calculations. The network's parameters have been learned during the training phase, where all the computational burden is handled before launch. Therefore, once onboard, the process is relatively fast and not contingent on the uncertainty of the number of iterations. The training steps build numerically and iteratively the NN model from the data. Table 1 compares qualitatively the three methods cited above.

B. Motivation for New Phasing Approaches for Space-Borne Segmented Telescopes

Recently, access to space has shown significant changes characterized by the emergence of the "NewSpace" paradigm, which

Table 1. Feature Comparison of FPWFS Methods

	Deep Learning Method	Image Sharpening	Phase Diversity
Data driven	Yes	No	No
Model-free	Yes	Yes	No
Iterative	No	Very	Yes
Onboard computing	Deterministic	Optical optimization	Numerical optimization
Off-board computing	Stochastic	None	None
Initial guess	No	Yes	Yes

emphasizes cost-effective access to space and introduces novel constraints such as limited volume allocation. Therefore, this necessitates the development of innovative methodologies, such as those shown in this paper. By leveraging a single focal plane image, we estimate and subsequently correct low-order phase aberration piston and tip-tilt using neural networks (NNs) and a space-borne, four-petal deployable CubeSat as the reference design. CubeSats are not only constrained in volume but also in computing power. Therefore, the complexity of the model will be considered in this study, as it plays a crucial role in minimizing optical aberration control for successful implementation on small satellites.

C. Previous Work on NN for Focal Plane Wavefront Sensing

The utilization of deep learning techniques has been proposed in previous studies as a powerful tool for FPWFS, often aiming for performance levels similar to or beyond the state of the art [13–15]. The majority of NN methods for FPWFS employ a single defocused image, as demonstrated in studies such as [13,15–19]. Others have used an optical preconditioner (overexposure, defocus, scatter) to improve the performance of the WFS [20]. Finally, others have used a pair of in- and out-of-focus images to identify Zernike coefficients from a pair of PSFs. [21]. These works have shown excellent performance for wavefront sensing using deep NN architectures, with Xception and Inception v3 being the most widely used. Xception utilizes a depthwise separable convolution approach, which separates spatial and channel-wise convolution operations. These architectures have been, for example, employed by Paine [15], Rajaoberison [17], Andersen [18], Nishizaki [20], and Orban [19]. Other popular architectures such as Resnet 50 and VGG-16, have also been tested by Paine [15] and Orban [19]. Architectures like EfficientNet-B4 (used by Quesnel [21]), Bi-GRU for sequential data (employed by Wang [16]), and Dense NN (used by Rossi [13]) also have demonstrated effectiveness for focal plane wavefront sensing. In these studies, two types of focal plane wavefront sensing have been identified. The majority of them focus on Zernike mode identification for a monolithic pupil imager (Paine [15], Andersen [18], Nishizaki [20], Orban [19], and Quesnel [21]), while others aim at phasing a segmented pupil by identifying low-order modes (Wang [16] focuses on piston and tip-tilt; Rossi [13] and Rajaoberison [17] focus on piston only). The major shortcoming of these studies is that the NN architecture requires hundreds of thousands to hundreds of millions of parameters (e.g., Resnet 50 and VGG-16 require,

respectively, 23 million and 138 million parameters). Our goal, apart from dealing with specific segmentation errors, is to minimize the number of parameters in order to optimize the inference time. For this reason we need to explore new suitable architectures with smaller dimensionalities while keeping a bottleneck architecture such as VGG-net and Resnet.

3. FORWARD MODEL

A. Image Formation

The image of a point source, or PSF, can be described as

$$\text{PSF} = |\text{FT}(P e^{j\phi})|^2 = |\text{FT}(P e^{j(\phi_{\text{PTT}} + \phi_{\text{HO}})})|^2, \quad (1)$$

where FT is the Fourier transform, and P is the pupil transmission function, with one inside the pupil boundaries and zero outside. ϕ is the phase function, which can be decomposed as $\phi = \phi_{\text{PTT}} + \phi_{\text{HO}}$ corresponding, respectively, to the PTT phase and the higher-order (HO) phase. Also, $\phi = \frac{2\pi\Delta}{\lambda}$, where Δ is the corresponding optical path difference (OPD) and the wavelength $\lambda = 800$ nm.

The AZIMOV prototype (Fig. 1) is composed of a pupil with four deployable petals. The precise design of the AZIMOV pupil is not finalized today. The global shape is four segments with an approximately square shape, but the final design might evolve due to the reduced volume constraints. In particular the four segments might not be symmetric in the final design. In this paper, we propose a simplified pupil composed of four squared segments where two segments are 10% cropped on the edge in order to break the pupil centro-symmetry as shown in Fig. 2. The asymmetric pupil Fourier wavefront sensor (APF-WFS) in Martinache's work shows great potential for small aberrations on an asymmetric pupil. Pope [22] shows an interest in using Martinache's work for segmented pupil, retrieving a piston value for an interval of $\frac{\lambda}{3.5}$. However, it has not been studied for large tip-tilt errors, a subject discussed in this paper. This method appears to be faster than our approach (only requiring a FFT) but lacks the ability to estimate large aberrations and only works on a point source—essential considerations for AZIMOV. In our case, cropping 10% of the segments is chosen as a trade-off between the loss of the collecting area and the correct wavefront estimation: the more asymmetric the pupil is, the better the even parts of the phase will be identified [12]. However, a limit is shown in the ratio of even part identification over the percentage of the cropped area. For instance, doubling the cropped area (from 10% to 20% of two segments) improves the estimation by 25%. For our space-borne application, the pupil design and the need to use the images from the science detector directly compel us to use instead an amplitude diversity as shown in Fig. 2.

To address the specifics of a segmented pupil, we define our modal basis with three modes $B_{i=1-4}^{k=1-4}$ on each of the four segments, i.e.,

$$\phi = \sum_{i=1}^3 \sum_{k=1}^4 c_i^k B_i^k, \quad (2)$$

where the modes are

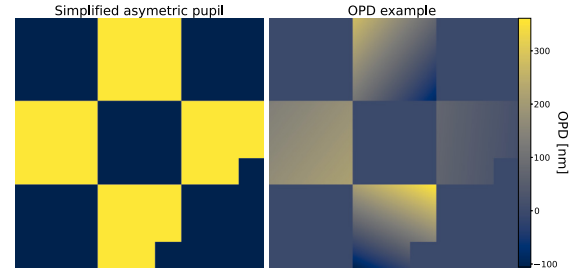


Fig. 2. Example of the pupil and OPD map.

$$\begin{aligned} B_1^k(x, y) &= 4, \\ B_2^k(x, y) &= 4\sqrt{3}x, \\ B_3^k(x, y) &= 4\sqrt{3}y, \end{aligned} \quad (3)$$

representing, respectively, piston, tip, and tilt on segment k . Clearly these modes are orthogonal across different segments and their inner product:

$$\langle B_i, B_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}. \quad (4)$$

The expansion coefficients in Eq. (3) are easily found when establishing an orthonormalized basis with unitary root-mean-square (RMS) over the whole, un-cropped, pupil in Fig. 2:

$$\frac{1}{4} \sqrt{\frac{1}{4} \int_{-1}^1 \int_{-1}^1 M_i M_j dx dy} = 1 \quad \text{if } i = j, \quad (5)$$

where the $1/4$ in front of the square root represents one-fourth of the pupil, i.e., the relative area of one segment, and the integral is computed on normalized units; and M_i represents the phase mode of one squared pupil. For instance M_1 corresponds to a piston on a squared segment, when the three others are set to zero. Equation (5) allows to identify normalization coefficients in Eq. (3).

Here we allow ourselves a slight lack of mathematical rigor in that the integration bounds, on account of the removed portion shown in Fig. 2, would not be exactly 1 and -1 . Since the amount of amplitude diversity needed to lift the even mode sign ambiguity is an open optimization parameter, we gloss over this detail and assume the same normalization for each and all segments.

The WFE definition for segmented pupils differs from the definition over monolithic pupils as shown in Eq. (6):

$$\text{WFE}(\{c_i^k\}) = \sum_{k=1}^4 \sqrt{\sum_{i=1}^N c_i^{k2}}. \quad (6)$$

This WFE is described by a quadratic sum of the phase coefficients over each segment. Provided the segments are non-overlapping, the total RMS is the sum of individual RMS values and not their quadratic sum.

An example of the PSFs studied is shown Fig. 3. These PSFs, paired with their Zernike coefficients, are directly used for the NN training.

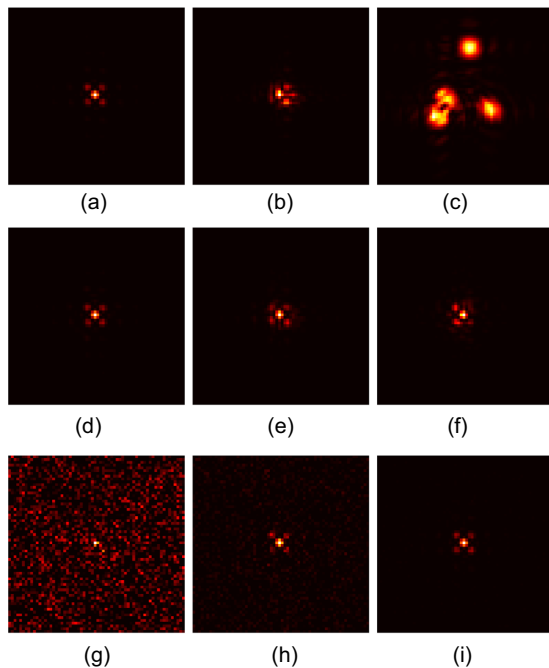


Fig. 3. Illustration of AZIMOV PSFs obtained in different configurations of WFE and signal to noise ratio (SNR). First row is PSF disturbed by PTT aberrations only. (a) PSF diffraction limit, (b) PSF $\text{WFE}_{\text{PTT}} = 97$ nm, (c) PSF $\text{WFE}_{\text{PTT}} = 454$ nm. Second row is PSF disturbed both by PTT and HO aberrations. (d) PSF diffraction limit + 30 nm HO, (e) PSF $\text{WFE}_{\text{PTT}} = 96$ nm $\text{WFE}_{\text{HO}} = 30$ nm, (f) PSF $\text{WFE}_{\text{PTT}} = 97$ nm $\text{WFE}_{\text{HO}} = 60$ nm. Last row is the noise influence depending on the SNR. (g) PSF diffraction limit, SNR = 10, (h) PSF diffraction limit, SNR = 30, (i) PSF diffraction limit, SNR = 100.

B. Data Generation

The training of NNs necessitates a substantial volume of data to effectively learn from diverse realizations. Unfortunately such an amount of real data is not available yet; therefore, data are simulated. All the generated imagery data are Nyquist sampled with a field of view (FoV) of 64×64 pixels at $\lambda = 800$ nm. We selected this field of view (FoV) based on the anticipated amplitude of the wavefront error (WFE) to be measured. The PSF's global intensity is normalized to one. Figure 4 presents the two studied distributions, one for large WFE errors corresponding to the initial mirror segments at the deployment (i.e., coarse phasing), and the other for smaller WFE errors for the final fine phasing step. The distribution of the 12 piston-tip-tilt coefficients follows a Gaussian distribution centered on 0 nm with $\sigma_{\text{fine}} = 10$ nm or $\sigma_{\text{coarse}} = 70$ nm. However, identification of pistons, under a single-image FPWFs setting, is constrained to a range between $-\frac{\lambda}{4}$ and $\frac{\lambda}{4}$ or their estimate can suffer from the so-called lambda ambiguity. We constrained piston values on a $\frac{\lambda}{2}$ width interval so that no piston can be identified at $+\frac{\lambda}{2}$ while another would be identified as $-\frac{\lambda}{2}$ since it will lead to the same PSF. Therefore, for the coarse phasing, the piston values are randomly drawn following the same distribution as for the fine phasing ($\sigma_{\text{coarse}}^{\text{piston}} = \sigma_{\text{fine}}^{\text{piston}}$) so that the same NN model studied in the following section can be used also in Section 6.

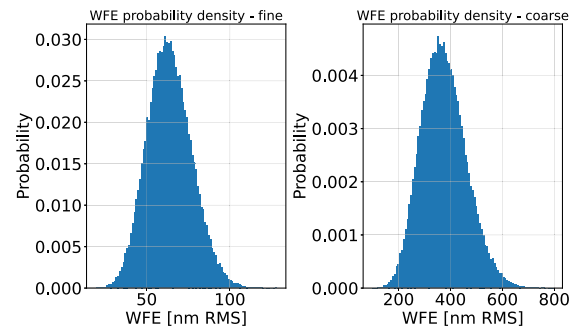


Fig. 4. Probability density of WFE over the full pupil for the training set calculated over 10^6 samples using Eq. (6). The figure on the left corresponds to the fine phasing. On the right, the distribution of the coarse phasing, studied in Section 6.

The capture range of the depicted distribution enables a range of minimum and maximum values within approximately $[-50, 50]$ nm RMS, corresponding to $[-200, 200]$ nm peak to valley (PV) for the piston and $[-350, 350]$ nm PV for tip-tilt. This amplitude is reached after an initial phasing step using the ELASTIC method [23], or our embedded onboard capacitive displacement sensors [2].

The determination of an optimal dataset size, depending on the complexity of the problem and the complexity of the learning algorithm (e.g., nonlinear relationships between input and output features), is controversial as there is currently no definitive methodology for its precise determination. We generated a dataset comprising 10^5 samples randomly split: 90% of the samples are allocated to the training set, while the remaining 10% are assigned to the validation. An additional 10^4 samples are generated exclusively for the testset so that the NN model can be tested with unseen data. All the datasets follow the same coefficient distribution.

C. Disturbance Sources

1. Noise

In a low-flux regime, the image mostly suffers from readout noise, estimated at 2.1 electrons (e.g., using a Sony Pregius-S IMX530 CMOS detector with 5320×4600 pixels) whereas photon noise dominates in high-flux cases. These are the two noises considered in the study, while other sources of noise (background, gain, drift) have been chosen to be ignored. Therefore, we define the signal to noise ratio (SNR) as

$$\text{SNR} = \frac{f}{\sqrt{f + \sigma_{\text{ron}}^2}}, \quad (7)$$

where f is the star flux in numbers of photons and $\sigma_{\text{ron}} = 2.1e^-$ the readout noise. The photon noise is described by a Poisson distribution $\mathcal{P}(f)$ for each pixel. The readout noise is a Gaussian noise drawn at $\mu = 0$, $\sigma = \sigma_{\text{ron}}$ added to the pixel value. The corresponding noisy images can be found in Fig. 3.

2. Higher-Order Aberrations

Higher-order aberrations introduce additional distortions to the wavefront on top of the PTT. However, their impact is different

from the PTT: although they degrade the PSF and therefore the delivered resolution, they cannot be compensated for. We carry out an analysis of their impact in the system: how they degrade the PTT estimation, and if it is needed to include them in the training set for PTT. These aberrations are added in addition to the PTT in the wavefront. They follow a $1/f^2$ power spectrum, typically encountered with good quality optics [24] and are orthogonal to PTT aberrations. The magnitude of HO aberrations is approximately 30 nm RMS over the full pupil, or in other words $\sum 1/f^2 = 30$ nm RMS, consistent with current estimates for AZIMOV.

4. METHODS

A. Deep Learning

Deep learning is a type of machine learning method able to extract representations or model directly from data with a complex architecture of learning successive layers with a high level of abstraction (increasingly meaningful representations). These architectures leverage a combination of diverse non-linear transformations to accomplish their objectives. The NN architecture has a lot of influence on the output accuracy and is composed of different layers, of which the two primary types are noted below.

- *Fully connected layers* [25]: layers composed of neurons. A neuron computes a weighted sum of its input, plus a bias and then applies a non-linear activation function. The model parameters are therefore the neurons' weights and biases.
- The *convolutional layer* [25], particularly adapted for image processing, computes the convolution of an image with one or several kernels. In this case, the neural network optimizes all the convolution kernels and biases.

The model architecture is created by adding layers one to another. Choosing the right architecture depends on the more important criterion of the project: it can be either the accuracy of the estimation, the model size, or the inference speed. In our work we aim to find the right balance between performance (the precision of WFE estimation) and model size (the size of the NN algorithm that needs to be embedded onboard the CubeSat). The number of model parameters depends on the choice of NN architecture. Decreasing it avoids overfitting and makes the model size smaller. Since our input data consists of 2D images, convolutional NNs have fewer parameters compared to fully connected layers and can capture pixel neighborhood information effectively. The simplest NN architecture to infer coefficients from an image is the VGG-net [26] or the Resnet [27] architecture, which are optimal in terms of number of parameters. Historically, the Resnet architecture is known for the “*skip connection*,” allowing the loss gradient to float easily through the deepest layer of the model during backpropagation. To achieve the best trade-off between performance and time complexity, we consider these two highly suitable architectures: VGG-net [26] and Resnet [27].

We therefore propose a NN baseline composed of five convolutional layer stacks as VGG-net architecture (two convolutional layers followed by one max-pooling layer) and three fully connected layers with a ReLu activation function. In the case of the Resnet, two skip connections are added to the

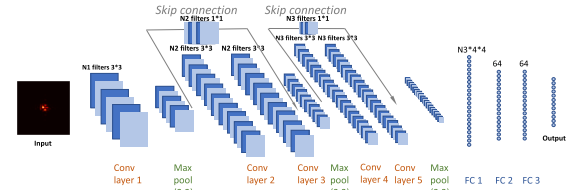


Fig. 5. Adopted Resnet architecture.

architecture (each skip connection is a convolutional layer with a 1×1 filter): the first one starts after the first layer to merge with the output of the third layer, and the second one after the third layer to merge after the fifth layer. The architecture is presented in Fig. 5 and has been chosen in order to find the smallest architecture suitable for reaching the diffraction limit requirement. Indeed such architecture goes up to 10^6 number of parameters whereas the smallest Resnet found in the literature, the Resnet-18, reaches 10^7 parameters, 10 times more than our heavier network.

N1, N2, and N3 correspond to the number of filters of the convolutional layers. This number of filters impacts the internal parameters and changes the size of the NN model (see Table 2); all model parameters are stored in a double precision float64 format as it is the default format in the Pytorch library. The number of parameters also impacts the computation time and influences the model's performance. Increasing the number of parameters can improve performance until the point of overfitting occurs, leading to limited generalization on new data.

The training step relies on a minimization algorithm of a cost function commonly known as the loss function. Conceptually, a loss function is a way of prioritizing which error to fix from our training samples, so that NN parameter updates result in adjustment to the output, decreasing the loss. The minimization algorithm optimizes the NN parameters (weights, biases, and filters) to minimize the loss function. Usually, the root mean square error is used, but in our study, the WFE over the pupil is computed as the loss function:

$$\text{Loss}_{\text{WFE}} = \text{WFE} (c_i^k - \hat{c}_i^k), \quad (8)$$

where c_i^k is the true coefficient for mode $i \in [1 \dots 3]$ and segment k , and $\hat{c}_i^k$ is the estimated coefficient. This loss, as mentioned above, allows the model to adjust accurately the model parameters according to the optical criterion on the pupil.

Table 2. Resnet Model Sizes Depending on the Number of Convolutional Filters^a

Label	N1	N2	N3	# Parameters	Model Size [MB]
NN1	4	8	16	$13 \cdot 10^3$	0.134
NN2	8	16	32	$31 \cdot 10^3$	0.278
NN3	16	32	64	$94 \cdot 10^3$	0.94
NN4	32	64	128	$326 \cdot 10^3$	2.6
NN5	64	128	256	$1.2 \cdot 10^6$	11

^aThis table approximates VGG-net sizes as well, since Resnet architecture has only slightly more internal parameters due to the skip connection that needs to link the number of filters from the input layer to its output layer.

B. Classical Phase Diversity and Image Sharpening

Other more classical FPWFS methods have also been proposed for the phasing of segmented telescopes, namely, phase diversity and image sharpening. The whole problem of phase diversity (PD) lies in the estimation from the focal plane image and the defocused image of the unknowns, which are the phase and the object [28]. However, in our case, the object is not to be estimated as it is a point source. Since our system does not include any additional optical components and aims to maintain the integrity of the focal plane image without any defocus, it is not possible to acquire a defocused image. Therefore, we propose a single-image phase diversity that minimizes the distance between the real PSF and the fitted one by adjusting the 12 phase coefficients. A criterion based on a quadratic pixel-to-pixel difference between the real PSF and the fitted one is minimized iteratively using the conjugate-gradient method [29]. A Tikhonov regularization is added to the criterion, in order to avoid the divergence of the phase coefficient, particularly at low SNR. The conjugate-gradient has shown the best performance in time, number of iteration, and WFE among others minimization methods.

Another method is image sharpening (IS). By taking focal plane images, the algorithm iteratively maximizes an image maximum intensity centered window by adjusting iteratively the PTT. Classical metrics are the maximum intensity in the center of the image, contrast, or the sum of the pixel squared. Alternative optimization methods are also possible and may lead to faster and improved estimates. Powell optimization methods [30] have shown the best efficiency in terms of noise robustness and computing time. To perform IS, we consider an exposure time of 0.1 ms. This exposure time is taking into account in the method calculus time as an image has to be acquired at each iteration of the algorithm. The computation time of generating the PSFs is deduced from the total computing time.

5. SAMPLE NUMERICAL RESULTS

A. Model Architecture

First, both types of model architectures (VGG-net and Resnet) are trained using identical hyperparameters (*batch size, learning rate, epochs*) and their inference performances are compared. The performances of the Resnet architecture are observed to systematically outperform the VGG-net even after fine-tuning both model hyperparameters (i.e., other models are trained) enhancing performance with respect to the architecture and number of parameters. The optimal set of hyperparameters for the Resnet includes a progressive scheduled learning rate (LR) depending on the epoch e :

$$LR(e) = \begin{cases} 10^{-4} & \text{if } e < 70 \\ 10^{-6} + (10^{-4} - 10^{-6})\exp^{-(e-70)*0.0125} & \text{otherwise.} \end{cases}$$

Specifically, a batch size of 32 PSFs is used during 300 learning epochs.

Figure 6 presents the performance comparison between the two architectures mentioned above for various model sizes. Performance is compared using the same test dataset and optimal sets of hyper-parameters, in terms of *residual WFE*, assuming a perfect correction based on the algorithm estimates.

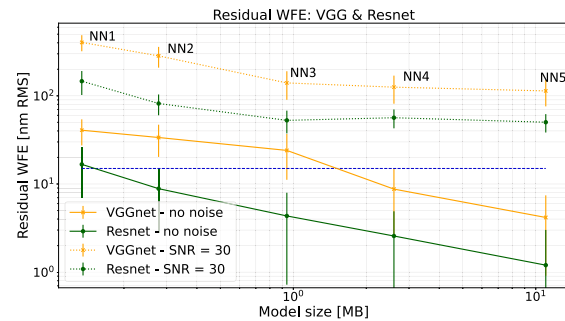


Fig. 6. Performance comparison of estimation between the VGG-net architecture (yellow line) and the Resnet architecture (green line) as a function of the model size. Dotted-line corresponds to model estimation for data at SNR = 30. The dashed blue line represents the residual target to reach the diffraction limit at 15 nm RMS. VGG-net size is rounded at the same value as Resnet for the graphic's readability.

Both NN models are trained with noiseless data and tested with either noiseless data or data at SNR = 30 (the noise study will be developed in Subsection 5.B). As expected, larger models lead to better performance, but at the cost of increased computational complexity. The best performance reaches about 1 nm of WFE, far below the residual target at 15 nm. However, when facing a low SNR, both NN models show weak performance where the best performance achieved is 50 nm of WFE for the heaviest Resnet. Improvement on noisy data is shown in Subsection 5.B. Whatever the model size, the Resnet outperforms the VGG-net. Additionally, Resnet has demonstrated faster convergence and lower loss function values compared to VGG-net. The skip connection in the Resnet architecture enables more effective parameter updates in the top layers and accelerates the convergence of the loss function. Based on the NN performance, the Resnet architecture is preferable. The requirement of 15 nm RMS is fulfilled by nearly all Resnet models regardless of the model size. However, in presence of noisy data, the methodology should be improved to reach the diffraction limit.

Subsequently, only the Resnet architecture will be examined in the rest of the paper. It is important to note that the simulated data used for evaluation lacks noise diversity and higher-order aberrations. Therefore, it is crucial to assess the robustness of the models by introducing noise and HO aberrations into the samples.

B. Model Robustness to Noisy Data

As mentioned above, images suffer from detector and photon noise. For the training step, the noise is added “on-the-fly” meaning that during each epoch, the noise is added to each noiseless PSF so that the noise realization is always different at each epoch for each PSF. Two scenarios are identified.

- Scenario 1: each NN model is trained with a noiseless dataset and is tested with a noisy dataset. Each single NN model is then tested over different SNRs. Several model sizes are tested. The aim is to evaluate the error made on noisy datasets by a model that has been trained with a noiseless dataset.

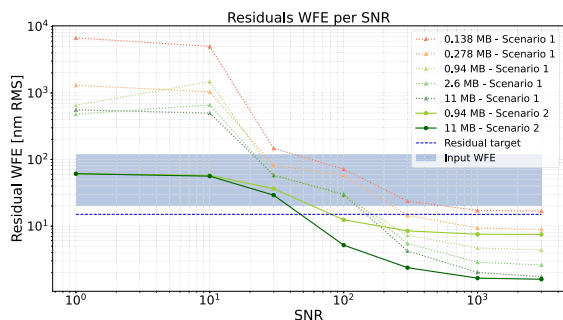


Fig. 7. Resnet robustness to noise: scenario 1 (dotted lines) and scenario 2 (solid lines). For the scenario 2 curves, each point represents one NN model, trained and tested on the corresponding SNR.

- Scenario 2: each NN model is trained with noisy samples at a given SNR. Each NN model is then tested on noisy data at the same SNR as for the training. Two NN model sizes are tested. The aim is also to evaluate the error made on noisy datasets when the NN models have been trained with noisy datasets.

We conducted a performance comparison between the two scenarios based on residual WFE after correction. Figure 7 illustrates the residual wavefront error depending on the SNR. The dotted lines represent NN models trained without noise, while the solid lines represent NN models trained with noisy PSFs (each point on the solid line corresponds to a NN model trained for a specific SNR). The performance of the models unsurprisingly improves as their size increases. The diffraction limit, defined by $WFE < \text{Residual target}$, is not achieved by the smallest models (dashed red line) even at high SNR. For low SNR, where the detector noise prevails over the photon noise, scenario 1 outputs an estimation error larger than the initial WFE: this would lead to an increase of the WFE if used in a closed loop. On the other hand, the scenario 2 network's estimation stays in the mean of the input WFE distribution: during training, the models minimize their loss function by setting all the phase coefficients to zero as the PSF geometry cannot be properly identified in the noise. At high SNR, in both scenarios, the largest model (NN5) converges towards the same WFE value. However, the NN3 scenario 2 model exhibits a slight limitation, achieving a WFE slightly below 10 nm RMS, whereas the same model trained without noise performs better.

An alternative approach is to train models using a range of SNRs drawn uniformly in the interval [30,100] to assess their performance.

Figure 8 illustrates a comparison of wavefront estimation quality between models trained on a fixed SNR and those trained on varying SNRs. Remarkably, the models trained on various SNRs show great performance, reaching the diffraction limit more rapidly at $\text{SNR} = 50$ for the NN5 and $\text{SNR} = 70$ for the NN3. Surprisingly, even at high SNRs, the models trained on varying SNRs demonstrate comparable performance to the models trained at fixed SNR. However, in a very-low-flux regime, performance is worsening the wavefront. Globally speaking, a SNR of at least 50 is required to reach the diffraction limit with a relatively small NN model. This is somehow similar to phase diversity.

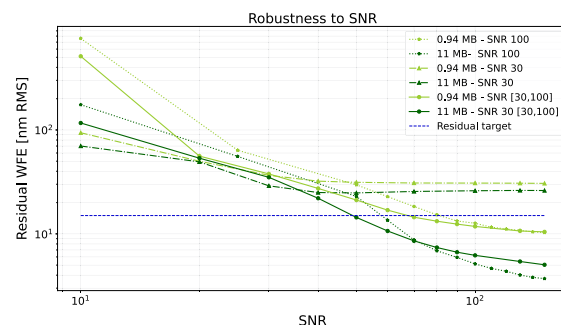


Fig. 8. Resnet robustness to SNR. Dark green curve illustrates the estimation quality of the NN5; light green shows NN3 performances.

C. Models Robustness to Higher-Order Aberrations

To validate the performance of the neural network, the model is also trained considering the presence of higher-order aberrations on the primary mirror. In the same way as for the noise robustness discussed in Subsection 5.B, we distinguish between two scenarios.

- Scenario 3: each NN model is trained with a dataset where higher-order aberrations are not considered, and is tested on PSFs generated with PTT and higher-order aberrations.
- Scenario 4: each NN model is trained with samples generated from PTT and higher-order aberration phase maps. The model is then tested over data that suffers from the same higher-order amplitude.

For these two scenarios the 12 PTT coefficients follow the same distribution. They are the only coefficients inferred by the NN model since the HO aberrations cannot be compensated for. The *residual WFE* refers here to the PPT part of the phase (i.e., excluding HO aberrations).

Figure 9 shows that even in the presence of unexpected HO aberrations, the inference step achieves a performance better than the diffraction limit requirement. The difference between scenario 3 and scenario 4 is only of a few nanometers, indicating that all model sizes exhibit robustness against small changes in the focal plane image due to HO aberrations. Surprisingly, the most robust model is for a size of 1MB, where the difference between scenario 3 and scenario 4 is the smallest, approximately 0.3 nm RMS. In contrast, the heaviest model begins to overfit the data: it infers with the best accuracy the raw data but when new data is introduced for testing, the performance slightly degrades by a few nanometers, while still reaching the diffraction limit. The inference accuracy worsens when stronger HO aberrations are added to the phase map, as shown by the red curve Fig. 9, which represents the inference performance when 60 nm RMS of WFE is added on top of PTT. However, the performance continues fulfilling the accuracy requirement and adding 60 nm RMS on the phase map introduces only a few nanometers of estimation errors.

D. Comparison to Classical Minimization Methods

Having explored the robustness of different model sizes in term of estimation accuracy, we now turn our attention to comparing

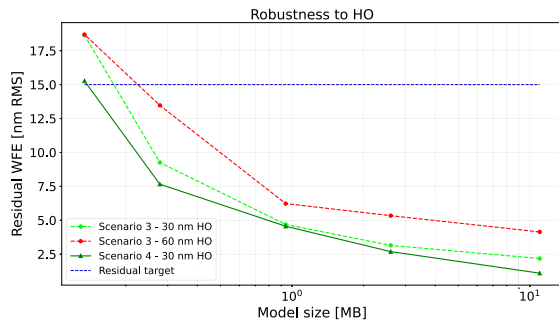


Fig. 9. Resnet performance as a function of model size for images including higher-order aberrations. The dashed light green line corresponds to scenario 3 (i.e., no HO aberrations included in the training) for 30 nm RMS of HO. The red line corresponds to scenario 3 for 60 nm RMS of HO aberrations. Solid green line, scenario 4 results for a NN trained with PSF aberrated with PTT and HO.

our method with phase diversity and image sharpening (IS) in terms of computation time and performance. Phase diversity relies on a numerical iterative minimization of a quadratic criterion over the image pixels, involving a computing time depending on the number of iterations of the algorithm. Each iteration requires several Fourier transforms as the model PSF is fitted to the measured data (see Subsection 4.B). Three tolerance criteria are chosen: one fast but less accurate $\epsilon_1^{\text{PD}} = 10^{-4}$, one finer but slower $\epsilon_2^{\text{PD}} = 10^{-7}$, and a very slow and precise $\epsilon_3^{\text{PD}} = 10^{-15}$. The principle is also iterative for IS: the metric is maximized until a tolerance criterion is reached; here, a tolerance criterion of $\epsilon^{\text{IS}} = 2 \cdot 10^{-4}$ is chosen. In our case, the metric employed is the maximum pixel intensity of a 2×2 window centered on the image. At each step, PTT corrections are applied (i.e., mirrors are moved), another image is acquired, and the maximum intensity is subsequently calculated.

Data used for the comparison is from the same dataset so that the same data is compared between the three methods.

We aim to compare the three methods in terms of wavefront estimation performance under the presence of noise and higher-order aberrations. NN3 and NN5 are the two neural networks used to compare our three methods, trained on at the corresponding SNR with HO aberrated data. NN coefficients estimation yields better phase estimations for both the NN5 and NN3 across all SNRs levels. Plus these models demonstrate a significant improvement in terms of computational time compared to PD and IS. NN models with their relatively lightweight architecture are much faster, with the majority of computational time occurring during training (which is obviously not taken into account for the inference). Figures 10 and 11 illustrate the performances of both methods relative to their execution time. The $\epsilon_3^{\text{PD}} = 10^{-15}$ criterion and $\epsilon_2^{\text{PD}} = 10^{-7}$ achieve nearly the same performance at SNR = 100 but with a different computing time. In Fig. 11 only ϵ_2^{PD} is used to provide a fairer comparison between the methods. Compared to the NN inference, which is stable in terms of computing time, PD shows limitations in execution time. First, PD takes longer than NN inferences, and second, the standard deviation of the computing time is larger. In addition, two distinct clusters are identified computing phase diversity with a tolerance criterion

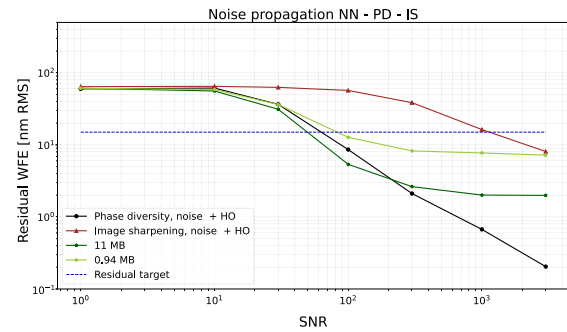


Fig. 10. PTT estimation of the NN (green line), phase diversity method at $\epsilon_3^{\text{PD}} = 10^{-15}$ (dark line), and image sharpening (brown line) facing the same dataset of PSF generated with 30 nm RMS of higher-order aberrations at several SNRs regardless of the computing time.

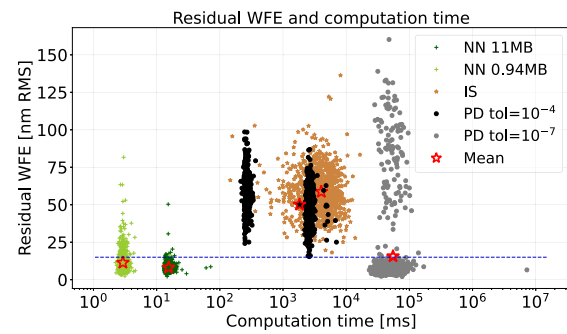


Fig. 11. Estimation time and precision of estimation for NN, phase diversity (PD), and image sharpening (IS). Two types of PD are studied, one fast but less accurate, one slow and more accurate. Test made over the same 1000 samples from the NN testset with HO aberrations for SNR = 100. Algorithm computed on a 2,3 GHz Quad-Core Intel Core i7.

of 10^{-4} , indicating that the number of iterations to reach convergence depends on the criterion shape. Better estimations can be achieved with a convergence criterion of 10^{-7} . However, it requires a longer computing time. Additionally, while PD reaches the diffraction limit requirement on average (red star), it does exhibit approximately 100 outliers out of 1000 samples. While IS is widely spread in terms of computing time and is not reliable, only PD follows NN performance in terms of WFE. However, the estimation time is at least three orders of magnitude slower compared to NN estimation. Therefore, the NN method appears to be more reliable to accurately estimate the wavefront compared to a mono-image phase diversity and image sharpening.

In this section, our primary focus has been examining NN methods and comparing them to two alternative approaches. Both phase diversity and image sharpening can naturally be improved upon. PD methods, for instance, could potentially benefit from linearized analytic phase diversity [31] to improve computation time and load. For IS, one could potentially investigate different image sharpness metrics that provide better

performance, but also look at faster optimization algorithms correcting one mode at a time [32].

6. TOWARDS A COMPLETE PHASING SEQUENCE

Once launched, the telescope will deploy the four petals with a mechanical accuracy of a few micrometers [5]. The initial stage of the phasing strategy aims to mitigate the significant piston and tip-tilt errors introduced during deployment. After initial segment identification and phasing procedure recurring, for instance, to the ELASTIC method [2,23], the expected WFE reaches a sub- λ amplitude over the entire pupil. We have not managed to train a suitable single NN able to retrieve high performance on large aberrations. However, we propose a solution with two NNs in cascade (a first NN for the initial large tip-tilt sensing ignoring the piston, and a second NN for the piston and the remaining PTT). The AZIMOV phasing strategy relies on two main steps. First, after deployment as the TT errors are large (as well as the piston errors), the spots do not superimpose and the piston errors do not affect (or only marginally) the resulting image. The primary mirror needs to be first corrected in tip-tilt in order to recenter each of the four PSFs generated by the individual segments. This allows the four PSFs to be superimposed on top of each other, resulting in a single PSF in the focal plane image. After this superposition, the piston error can be estimated. Finally, the remaining piston and residual tip-tilt errors can be estimated using the NN presented in Section 5. Our correction can be applied iteratively in closed-loop at SNR = 100, to process real-time wavefront estimation. The proposed phasing strategy can be found Fig. 12.

The first iteration corrects for the coarse tip-tilt errors. Figure 13(b) shows this while the piston errors remain unchanged. Then after the second iteration, the NN responsible for fine PTT phasing corrects for the remaining aberrations eventually converging towards about a nanometer of piston and residual tip-tilt resulting in <10 nm WFE over the full pupil. The diffraction limit is reached after two or three iterations for the 11 MB model, and slightly after four or five iterations for the 0.94 MB model, Fig. 13(a). The presence of noise at SNR = 100 increases the standard deviation of the estimation error. However, the estimation remains well under the requirement and shows a great stability during 100 iterations for both models. In addition, it is important to notice that we limited the tip-tilt amplitude due to the amount expected in the deployment, yet larger tip-tilt could be easily added and learned by our method as long as the PSFs remain in the field of view of the detector.

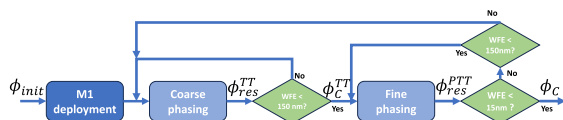


Fig. 12. Diagram of the AZIMOV phasing strategy, where ϕ_{init} is the initial aberrated phase, $\phi_{\text{res}}^{\text{TT}}$ the residual tip-tilt phase with uncorrected piston, $\phi_{\text{res}}^{\text{PTT}}$ the phase with pre-corrected tip-tilt, $\phi_{\text{res}}^{\text{PTT}}$ the residual phase after PTT correction, and ϕ_{C} the phase corresponding to the diffraction limit.

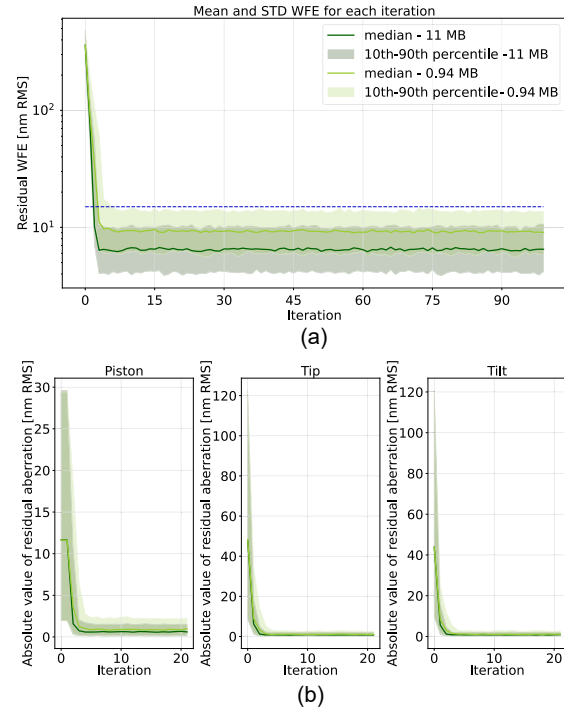


Fig. 13. Closed-loop residuals for AZIMOV deployment regime; mean and STD results over the same 1000 PSFs of the testset at SNR = 100. Both figures contain the same legend. (a) Average residual WFE for 100 iterations in a closed-loop over 1000 PSFs of the testset. The green shaded area represents the bounds of the 10th and 90th percentiles. (b) Average PTT residuals for 15 iterations in a closed-loop. The green shaded area represents the bounds of the 10th and 90th percentiles. Iteration 0 corresponds to the initial WFE.

7. CONCLUSION

This study explored the performance of neural networks for phasing a four-segment deployable space telescope with piston, tip, and tilt aberrations in the range of (20 – 120 nm RMS). The existing state-of-the-art algorithms are deemed too complex because of their iterative nature and large computational burden for space applications. Alternatively, simplified NN architectures are explored, which exhibit promising results for accurately estimating the piston-tip-tilt phasing errors. A five-convolutional-layer-based Resnet has been implemented for demonstrating robustness even in the presence of additional high-order aberrations and noise, by reaching the diffraction limit in different cases at SNR = 50 and for 30 nm RMS and 60 nm RMS of HO aberrations. Additionally, it outperforms concurrent methods such as mono-image phase diversity and image sharpening in terms of computation time by at least a factor of 100 with comparable noise propagation performance. Our implementation shows great performance and stability once applied in a closed loop, for both coarse and fine phasing, using a relatively simple architecture allowing for fast inference times of about a few milliseconds per step.

However, using a single in-focus image for FPF sensing appears to limit the piston capture range between $-\frac{\lambda}{4}$ and $\frac{\lambda}{4}$. Nevertheless, this limitation might be overcome by using several

input images. For instance, capturing two images at different wavelengths or alternatively using an in-focus and an out-of-focus image could overcome the lambda ambiguity and extend the capture range. Ultimately, the NN approach provides a more generic solution, both in terms of aberration capture range and potentially in terms of scene (point source versus extended).

In our case of simulated data, NNs are driven by the image formation model. Once on-sky data will be available, fine-tuning our method can enhance the performance to fit better to realistic cases.

Finally, our work presents results employing a point source. Finding such a source may be possible during the initial calibration stage. However, in the context of Earth observation, obtaining such a point source may be difficult in practice as objects of interest are typically more complex and extended. Further work will be directed towards the study of extended objects in regions of adequate contrast to convey sufficient phase information.

Funding. European Structural and Investment Funds (POCI-01-0247-FEDER-047264); Centre National d'Etudes Spatiales.

Acknowledgment. This work is supported by European Structural and Investment Funds in the FEDER component, through the Operational Competitiveness and Internationalization Programme (COMPETE 2020) [Project no. 047264]. The authors acknowledge the support by the Action Spécifique Haute Résolution Angulaire (ASHRA) of CNRS/INSU co-funded by CNES. This project is supported by ANR WOLF. This research has made use of computing facilities operated by CeSAM data center at LAM, Marseille, France.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

1. P. A. Sabelhaus and J. E. Decker, "An overview of the James Webb space telescope (JWST) project," *Proc. SPIE* **5487**, 550–563 (2004).
2. N. Schwartz, W. Brzozowski, M. Milanova, *et al.*, "High-resolution deployable CubeSat prototype," *Proc. SPIE* **11443**, 1144331 (2020).
3. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Elsevier, 2013).
4. N. Schwartz, W. Brzozowski, Z. Ali, *et al.*, "6U CubeSat deployable telescope for optical earth observation and astronomical optical imaging," *Proc. SPIE* **12180**, 1073–1084 (2022).
5. J.-F. Sauvage, N. Schwartz, S. Vievard, *et al.*, "First error budget of a deployable CubeSat telescope," *Proc. SPIE* **11443**, 1144330 (2020).
6. G. A. Chanan, J. E. Nelson, and T. S. Mast, "Segment alignment for the Keck telescope primary mirror," *Proc. SPIE* **628**, 466–471 (1986).
7. S. Y. Haffert, L. M. Close, A. D. Hedglen, *et al.*, "Phasing the Giant Magellan Telescope with the holographic dispersed fringe sensor," *J. Astron. Telesc. Instrum. Syst.* **8**, 021513 (2022).
8. D. S. Acton, S. Knight, M. Carrasquilla, *et al.*, "Phasing the Webb telescope," *Proc. SPIE* **12180**, 121800U (2022).
9. M. D. Perrin, D. S. Acton, C.-P. Lajoie, *et al.*, "Preparing for JWST wavefront sensing and control operations," *Proc. SPIE* **9904**, 142–160 (2016).
10. M. P. Lamb, C. Correia, J.-F. Sauvage, *et al.*, "Quantifying telescope phase discontinuities external to adaptive optics systems by use of phase diversity and focal plane sharpening," *J. Astron. Telesc. Instrum. Syst.* **3**, 039001 (2017).
11. L. M. Mugnier, A. Blanc, and J. Idier, *Advances in Imaging & Electron Physics* (2006), Vol. **141**, p. 1.
12. F. Martinache, "The asymmetric pupil Fourier wavefront sensor," *Publ. Astron. Soc. Pac.* **125**, 422–430 (2013).
13. F. Rossi, C. Plantet, A.-L. Cheffot, *et al.*, "Machine learning techniques for piston sensing," *Proc. SPIE* **12185**, 1669–1674 (2022).
14. J. Herbel, T. Kacprzak, A. Amara, *et al.*, "Fast point spread function modeling with deep learning," *J. Cosmol. Astropart. Phys.* **2018**, 054 (2018).
15. S. W. Paine and J. R. Fienup, "Machine learning for improved image-based wavefront sensing," *Opt. Lett.* **43**, 1235–1238 (2018).
16. Y. Wang, F. Jiang, G. Ju, *et al.*, "Deep learning wavefront sensing for fine phasing of segmented mirrors," *Opt. Express* **29**, 25960–25978 (2021).
17. H. F. Rajaoberison, J. S. Tang, and J. R. Fienup, "Machine learning wavefront sensing for the James Webb space telescope," *Proc. SPIE* **12180**, 2210–2216 (2022).
18. T. Andersen, M. Owner-Petersen, and A. Enmark, "Image-based wavefront sensing for astronomy using neural networks," *J. Astron. Telesc. Instrum. Syst.* **6**, 034002 (2020).
19. G. Orban De Xivry, M. Quesnel, P. Vanberg, *et al.*, "Focal plane wavefront sensing using machine learning: performance of convolutional neural networks compared to fundamental limits," *Mon. Not. R. Astron. Soc.* **505**, 5702–5713 (2021).
20. Y. Nishizaki, M. Valdivia, R. Horisaki, *et al.*, "Deep learning wavefront sensing," *Opt. Express* **27**, 240–251 (2019).
21. M. Quesnel, G. Orban de Xivry, G. Louppe, *et al.*, "A deep learning approach for focal-plane wavefront sensing using vortex phase diversity," *arXiv*, arXiv:2210.00632 (2022).
22. B. Pope, N. Cvetojevic, A. Cheetham, *et al.*, "A demonstration of wavefront sensing and mirror phasing from the image domain," *Mon. Not. R. Astron. Soc.* **440**, 125–133 (2014).
23. S. Vievard, F. Cassaing, and L. Mugnier, "Large amplitude tip/tilt estimation by geometric diversity for multiple-aperture telescopes," *J. Opt. Soc. Am. A* **34**, 1272–1284 (2017).
24. K. Dohlen, F. Wildi, J. Beuzit, *et al.*, "SPHERE: confronting in-lab performance with system analysis predictions," in *Second International Conference on Adaptive Optics for Extremely Large Telescopes* (LESIA, 2011).
25. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT, 2016).
26. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv*, arXiv:1409.1556 (2014).
27. K. He, X. Zhang, S. Ren, *et al.*, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778.
28. A. Blanc, "Identification de réponse impulsionnelle et restauration d'images: apport de la diversité de phase," Ph.D. thesis (Paris 11, 2002).
29. B. T. Polyak, "The conjugate gradient method in extremal problems," *USSR Comput. Math. Math. Phys.* **9**, 94–112 (1969).
30. M. J. Powell, *A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation* (Springer, 1994).
31. I. Mocoeur, "Analyse de front d'onde en plan focal: développement d'algorithmes temps-réel et application au cophasage de télescopes multipupilles imageurs," Ph.D. thesis (Université Paris Sud-Paris XI, 2008).
32. D. Débarre, E. J. Botcherby, M. J. Booth, *et al.*, "Adaptive optics for structured illumination microscopy," *Opt. Express* **16**, 9290–9305 (2008).

