



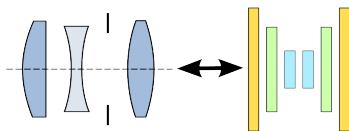
RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*

ONERA

THE FRENCH AEROSPACE LAB

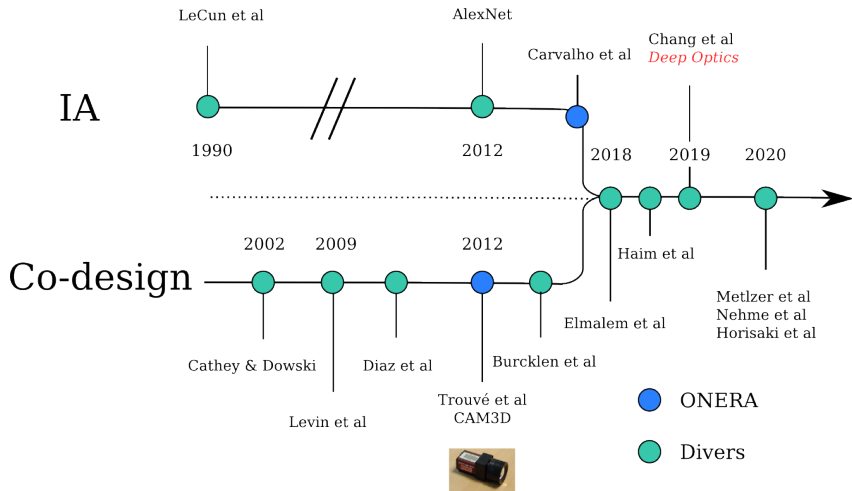
Conception conjointe optique/traitement: du Codesign au DeepCodesign



Pauline Trouvé-Peloux
ONERA DTIS/IVA
pauline.trouve@onera.fr

- 1 Introduction
- 2 Quelques notions utiles sur les réseaux de neurones
- 3 Quelques exemples de l'état de l'art
- 4 DeepCodesign à l'ONERA
- 5 Conclusion et perspectives

- 1 Introduction
- 2 Quelques notions utiles sur les réseaux de neurones
- 3 Quelques exemples de l'état de l'art
- 4 DeepCodesign à l'ONERA
- 5 Conclusion et perspectives



Plusieurs applications développées :

- Extension de profondeur de champ [1, 2, 3]
- Extension du champ de vue latéral [4]
- 3D [5, 6, 7, 8, 9]
- Classification d'objets [10, 6]
- HDR [11]
- Imagerie sans lentille [12]

Avec des optiques non conventionnelles :

- Masque de phase [1, 2, 3, 5, 7]
- Pupille codée en amplitude [12]
- Lentille *Freeform* [6]

⇒ L'enjeu est de modéliser l'optique dans le réseau de neurones

- 1 Introduction
- 2 Quelques notions utiles sur les réseaux de neurones
- 3 Quelques exemples de l'état de l'art
- 4 DeepCodesign à l'ONERA
- 5 Conclusion et perspectives

Réseaux de neurones - convolutifs

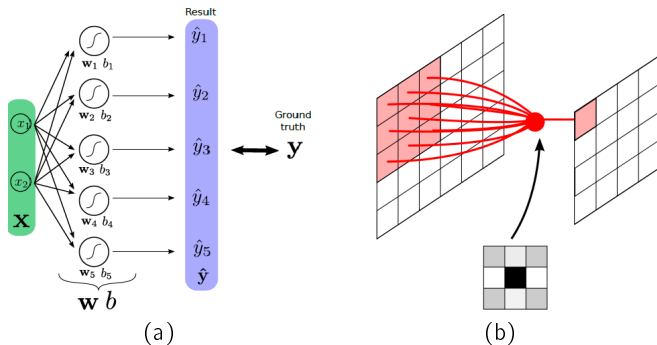
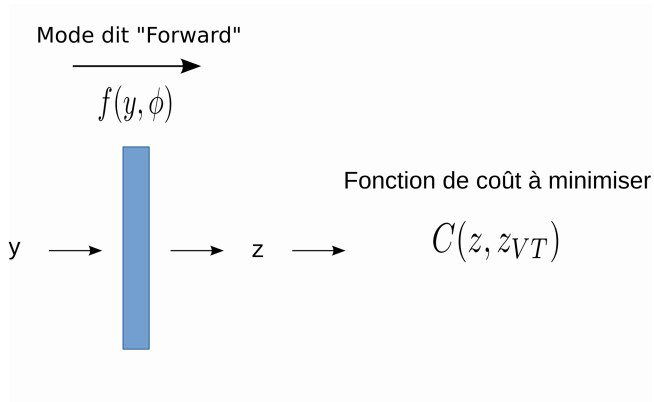


Figure : Extraits du cours d'A. Boulch pour l'IOGS. (a) Couche Réseau de neurones générique. (b) Réseau de neurones convolutif.

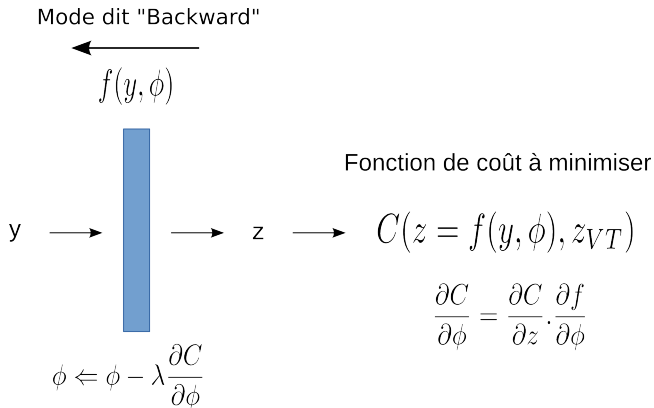
- ⇒ Réseaux de neurones *convolutifs* majoritairement utilisés pour des images
- ⇒ Enjeux : apprendre les coefficients des noyaux de convolutions

- 2 Quelques notions utiles sur les réseaux de neurones
 - Principe de l'optimisation d'un réseau de neurone
 - Comment introduire l'optique dans un réseau de neurones convolutif?

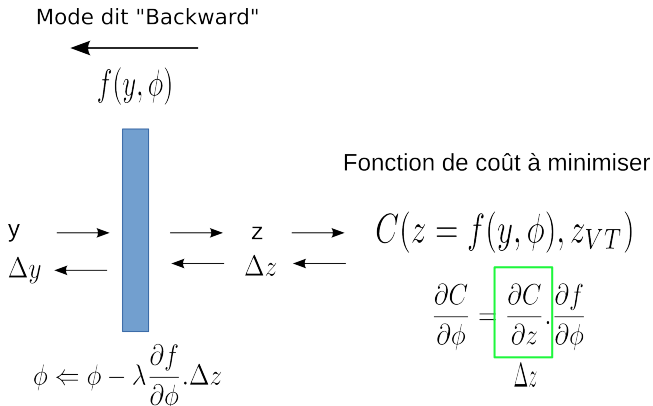
Comment est optimisé un réseau de neurones ?



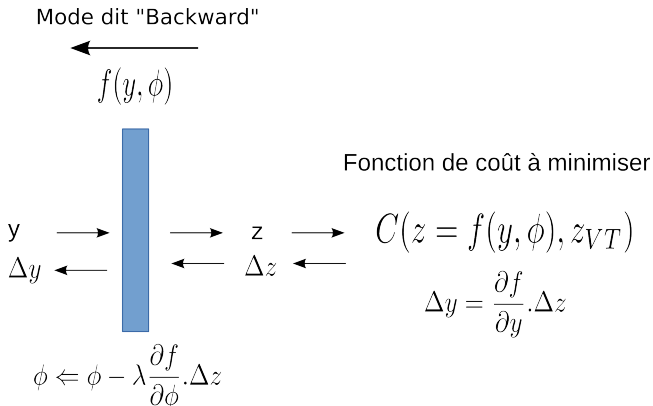
Comment est optimisé un réseau de neurones ?



Comment est optimisé un réseau de neurones ?



Comment est optimisé un réseau de neurones ?



- 2 Quelques notions utiles sur les réseaux de neurones
 - Principe de l'optimisation d'un réseau de neurone
 - Comment introduire l'optique dans un réseau de neurones convolutif?

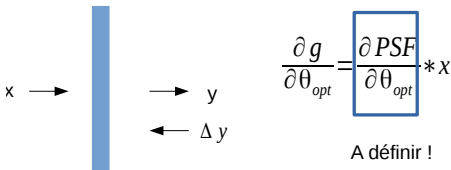
Comment introduire l'optique dans un réseau de neurones convolutif?

Modèle de formation d'images : $y = g(x, \theta_{opt}, b) = PSF(\theta_{opt}) * x + b$

⇒ Une optique peut être modélisée par une couche du réseau convolutif

Couche "optique"

$$g(x, \theta_{opt}, b) = PSF(\theta_{opt}) * x + b$$

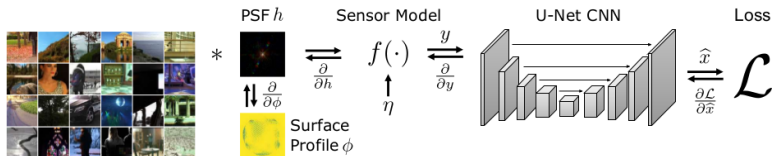


$$\theta_{opt} \leftarrow \theta_{opt} - \lambda \frac{\partial g}{\partial \theta_{opt}} \Delta y$$

⇒ En ajoutant un réseau de neurones pour traiter y on peut faire de la co-conception !

Comment introduire l'optique dans un réseau de neurones convolutif?

Schéma complet issu de la littérature Metzler et al 2020 [11]



- 1 Introduction
- 2 Quelques notions utiles sur les réseaux de neurones
- 3 Quelques exemples de l'état de l'art
- 4 DeepCodesign à l'ONERA
- 5 Conclusion et perspectives

- 3 Quelques exemples de l'état de l'art
 - Optimisation d'un masque de phase pour l'EDOF
 - Lentille "freeform" pour l'estimation de profondeur

Travaux proposés par Elmalem et al, (2018) [1]

Principe : Un masque de phase pour créer du chromatisme et faire de l'EDOF

$$\phi_{Masque}(r) = \phi \text{ si } r_1 < r < r_2$$

$$\phi_{Masque}(r) = 0 \text{ sinon}$$

Modèle optique : Optique de Fourier

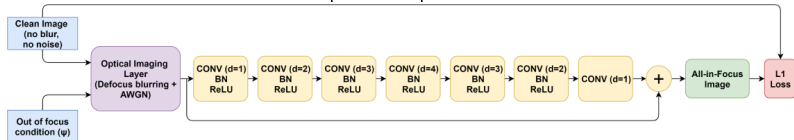
$$h_{\psi}^{\theta}(x_P, y_P) \propto \left| TF \left(A(x, y) \exp^{i\psi \frac{(x^2+y^2)}{R^2}} \exp^{i\phi_{Masque}^{\theta}(x,y)} \right) \right|_{\left(\frac{x'_P}{\lambda z_i}, \frac{y'_P}{\lambda z_i} \right)}^2$$

Ce modèle est dérivable !

Optimisation des paramètres r_1 , r_2 et ϕ à l'aide d'un réseau de neurones

Base de données : Fenêtres de texture variable de taille 64x64
(hypothèse de profondeur localement constante)

Principe de l'optimisation



Des idées à retenir sur ces travaux :

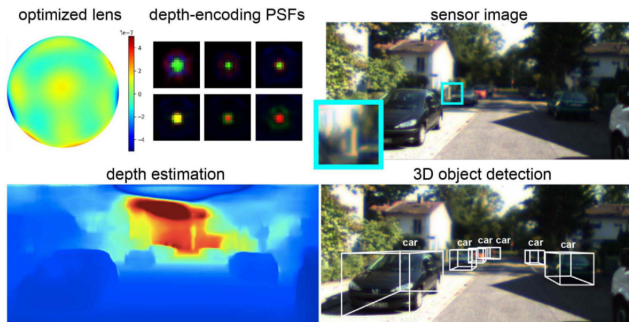
- Sensibilité du résultat au point de départ
- Le composant optique non conventionnel permet d'utiliser un réseau peu profond
- Etude de robustesse du réseau de déconvolution à des perturbations de fabrication du masque
- Validation expérimentale après fabrication du masque

- 3 Quelques exemples de l'état de l'art
 - Optimisation d'un masque de phase pour l'EDOF
 - Lentille "freeform" pour l'estimation de profondeur

Optimisation d'une lentille feeform pour la 3D

Equipe très active sur la thématique **Deep Optics** localisée à Stanford dirigée par G. Wetzstein

Travaux de J. Chang et al [6]



Autres publications sur la thématique *DeepOptics* : [2, 6, 11, 13, 14, 9]

Les idées principales :

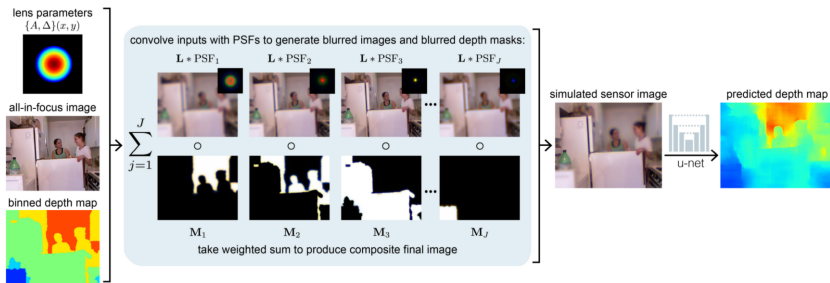
- Optimiser un élément optique non conventionnel qui *encode* la profondeur dans les images (comparaison de différentes optiques)
- Modèle optique basé sur de l'optique de Fourier
- Simulation d'images 3D à partir de base de données "RGB-D" Kitti, NYUV-2, "Rectangles"
- Réseau d'estimation de profondeur de type U-Net

Différentes aberrations optiques sont comparées :

- Defocalisation seule
- Astigmatisme
- Chromatisme (intrinsèque à une lentille)
- Optique freeform (surface de la lentille laissée libre)

Comment simuler les images ?

A partir du modèle optique $h(\phi_{opt})$ et d'une base d'images RGB-D



→ Peut-être vu comme un encodeur des données

→ On peut se poser la question de la VT "cachée" dans l'image

Optimisation de l'optique

Optical model	Rectangles		NYU Depth v2		KITTI*	
	RMSE _{lin}	RMSE _{log}	RMSE _{lin}	RMSE _{log10}	RMSE _{lin}	RMSE _{log}
All-in-focus	0.4626	0.3588	0.9556	0.1452	2.9100	0.1083
Defocus, achromatic	0.2268	0.1805	0.4814	0.0620	2.5400	0.0776
Astigmatism, achromatic	0.1348	0.0771	0.4561	0.0559	2.3634	0.0752
Chromatic aberration	<i>0.0984</i>	<i>0.0563</i>	<i>0.4496</i>	<i>0.0556</i>	<i>2.2566</i>	<i>0.0702</i>
Optimized, annular	0.1687	0.1260	0.4817	0.0623	2.7998	0.0892
Optimized, freeform	0.0902	0.0523	0.4325	0.0520	1.9288	0.0621

Table 1. Depth estimation error with different optical models for various datasets. RMSEs are reported for linear and log (base e or 10) scaling of depth (m or log(m)). Lowest errors are bolded, and second-lowest are italicized. The KITTI* dataset is our KITTI dataset subset.

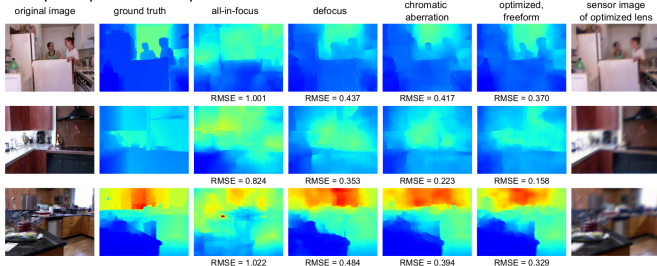
→ L'optique free-form obtient les meilleures performances

Optimisation d'une lentille feeform pour la 3D

Résultats qualitatifs sur données simulées

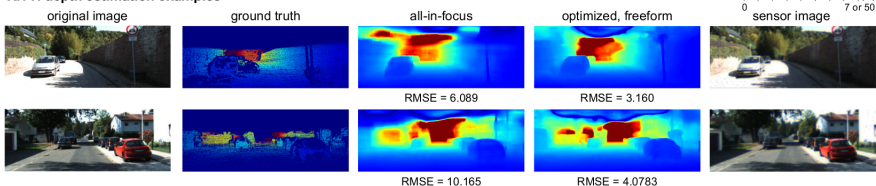
Résultats sur NYU-V2

NYU Depth v2 depth estimation examples



Résultats sur Kitti

KITTI depth estimation examples



Comparaison avec l'état de l'art sur NYU-V2

Method	rel	log10	rms	δ_1	δ_2	δ_3
Laina <i>et al.</i> [19]	0.127	0.055	0.573	0.811	0.953	0.988
MS-CRF [47]	0.121	0.052	0.586	0.811	0.954	0.987
DORN [6]	0.115	0.051	0.509	0.828	0.965	0.992
All-in-focus	0.293	0.145	0.956	0.493	0.803	0.936
Defocus	0.108	0.062	0.481	0.893	0.981	0.996
Astigmatism	0.095	0.056	0.456	0.916	0.986	0.998
Chromatic	0.095	0.056	0.450	0.916	0.987	0.998
Freeform	0.087	0.052	0.433	0.930	0.990	0.999

Table 2. Comparative performance on NYU Depth v2 test set, as calculated in [5]. Units are in meters or log10(m). Thresholds are denoted δ_i : $\delta_i > 1.25^i$. Lowest errors and highest δ s are bolded.

→ Amélioration des performances avec l'optique optimisée

Entraînement de deux réseaux de l'état de l'art de détection d'objet :

- Fast RCNN (2D)
- FPointNet (3D)

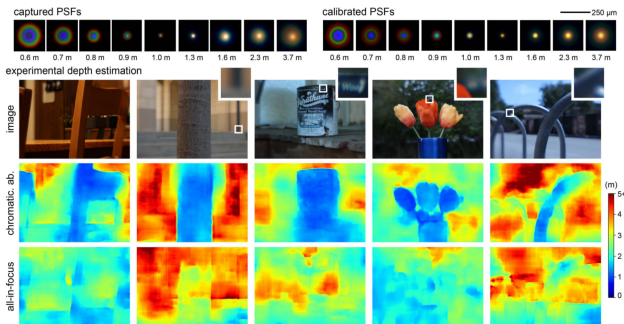
Object detection metric	All-in-focus	Optimized
2D mAP	78.01	78.96
2D AP, Car	95.50	95.15
2D AP, Pedestrian	80.06	80.22
2D AP, Cyclist	89.77	88.11
3D AP, Ped., Easy	9.74	13.86
3D AP, Ped., Moderate	7.10	11.74
3D AP, Ped., Hard	6.21	11.90
3D AP, Cyc., Easy	2.27	7.18
3D AP, Cyc., Moderate	2.36	4.89
3D AP, Cyc., Hard	1.98	4.95

Table 3. Object detection performance measured by 2D AP % (IoU = 0.5) and 3D AP % (IoU = 0.5) on our validation split of the KITTI object detection dataset using the all-in-focus and optimized mask models. Higher values are bolded.

- L'optique optimisée ne détériore pas la tâche de détection 2D
- L'optique optimisée améliore les tâches de détection d'objet 3D

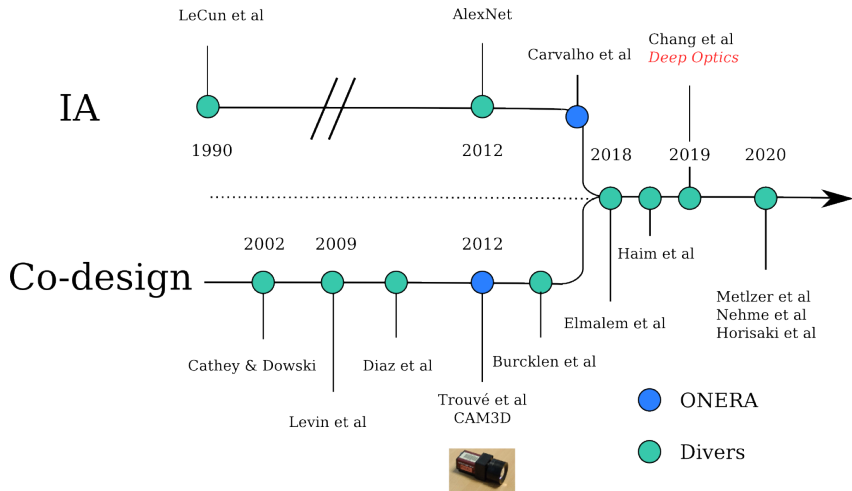
En pratique :

- Réalisation d'un système avec la solution d'une lentille simple (chromatisme intrinsèque)
- Etalonnage des PSF
- "Finetuning" du réseau d'estimation de profondeur sur images simulées
- Application à de vraies acquisitions



- 1 Introduction
- 2 Quelques notions utiles sur les réseaux de neurones
- 3 Quelques exemples de l'état de l'art
- 4 DeepCodesign à l'ONERA**
- 5 Conclusion et perspectives

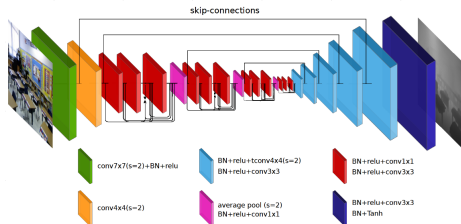
- 4 DeepCodesign à l'ONERA
 - Historique
 - Deep depth from defocus
 - Deep codesign



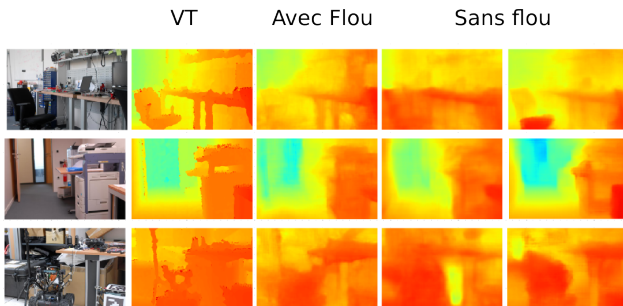
- 4 DeepCodesign à l'ONERA
 - Historique
 - Deep depth from defocus
 - Deep codesign

Thèse de M. Carvalho [15]

- Réseau de neurones D3-Net pour l'estimation de profondeur monoculaire
- Ajout du flou de défocalisation dans les images (bases de données simulées et réelles)
- Illustration de l'apport du flou de défocalisation sur les performances



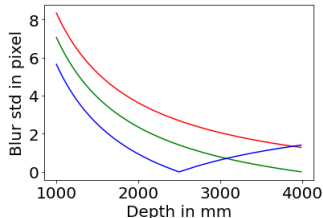
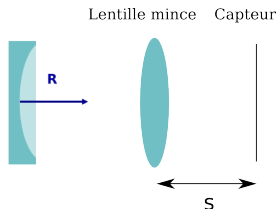
Comparaison de performance sur des images réelles



- Meilleure performance sur des critères quantitatifs
- Meilleure généralisation du réseau

- 4 DeepCodesign à l'ONERA
 - Historique
 - Deep depth from defocus
 - Deep codesign

Add-on chromatique développé en collaboration avec l'IOGS [16]

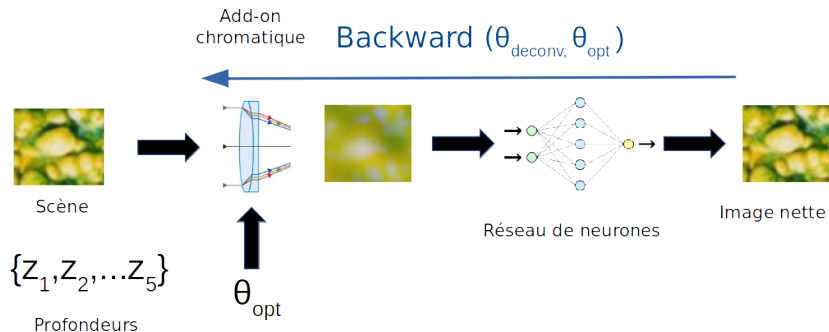


Permet d'ajouter du chromatisme à un objectif existant
⇒ Applications en EDOF et/ou 3D

Degrés de liberté :

- rayon de courbure (R)
- **position du capteur** (s)

PSF supposée gaussienne d'écart type : $\sigma(\lambda, R, s, z) = Ds \left(\frac{1}{f'(\lambda, R)} - \frac{1}{z} - \frac{1}{s} \right)$



- Implémentation d'une couche "optique" devant un réseau de neurones
- Tests d'architectures de déconvolution [17, 18]
- Optimisation conjointe de la mise au point de l'optique et du réseau

Plusieurs architectures de l'état de l'art : [17, 18]

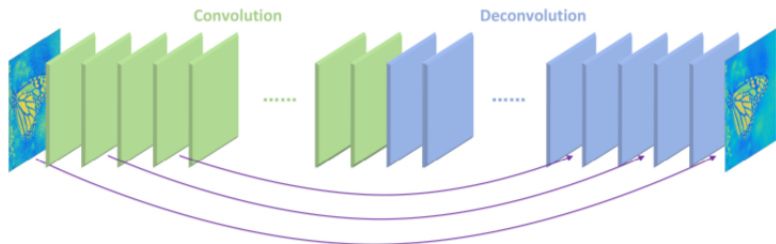
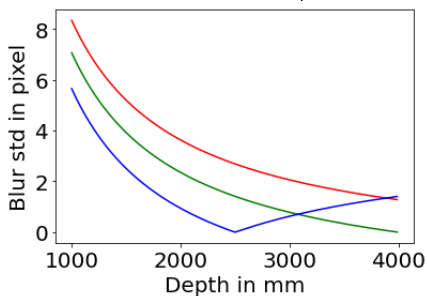


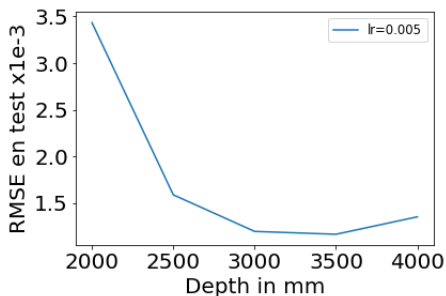
Figure : Architecture de [18]

Point de départ - optique fixée

Variation du flou avec la profondeur



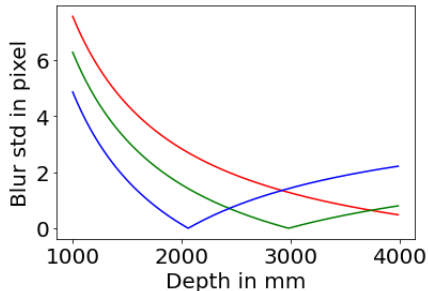
RMSE après déconvolution



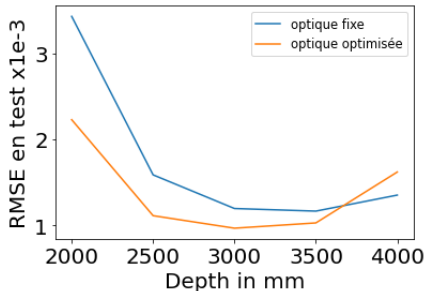
⇒ Le réseau de déconvolution fonctionne bien

Optimisation de la position du détecteur (même lr)

Variation du flou avec la profondeur



RMSE après déconvolution



⇒ L'optimisation conjointe diminue le RMSE par rapport au cas d'une optique fixée

Conclusion

- Preuve de concept de co-conception optique/réseau avec un modèle optique simple
- Premières analyses de co-conception pour un seul paramètre à optimiser

Perspectives

- Augmentation du nombre de paramètres à optimiser
- Etude de la sensibilité au point de départ
- Analyse de l'optimisation des deux composants optique vs traitement

Mais tous les travaux de l'état de l'art reposent sur des modèles optiques "simples" (optique de Fourier)

Principe : Développer un modèle optique complet et différentiable

Comment ? Travaux de Volatier et al (2017) [19]

- Par tracé de rayon en tirant parti de la différentiabilité du principe de Fermat
- Avec un algorithme qui calcule la dérivée de la propagation d'un rayon en fonction des paramètres "système "
- En utilisant les outils du deep learning : la différentiation algorithmique

Modèle développé à l'ONERA (DRT)

Entrées :

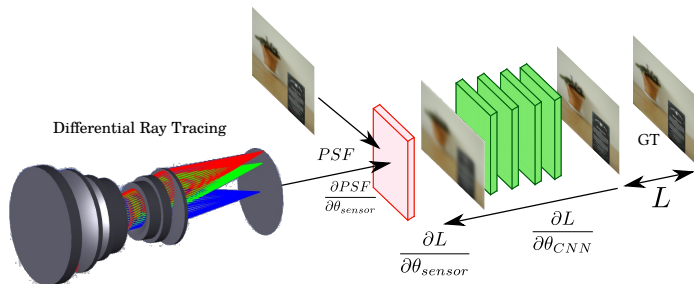
- Paramètres des lentilles (épaisseurs, positions,...)
- Paramètres du point source (**champ**, profondeur, longueur d'onde)

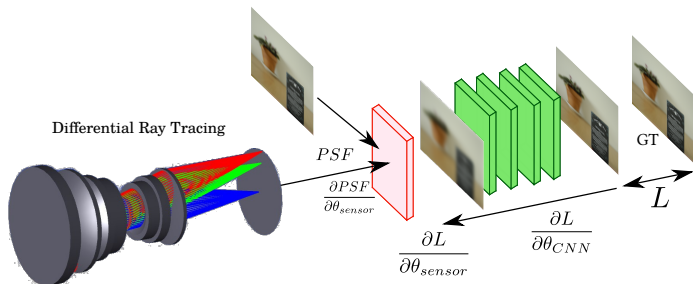
Sorties :

- La PSF en n'importe quel point du champ
- Le jacobienne de la PSF pour chaque paramètre optique

Modèle optique complet et différentiable

Schéma de principe

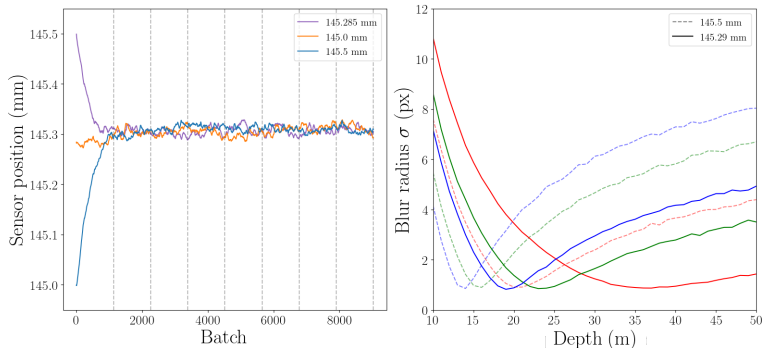




Travaux en cours de réalisation (stage d'A. Halé)

- Ajout d'un add-on chromatique devant un double Gauss
- Réseau de déconvolution de type [18] appliqué sur **des fenêtres**
- Optimisation **uniquement de la mise au point** pour augmenter la profondeur de champ sur une plage de distance fixée

Résultats d'optimisation pour plusieurs points de départ différents



⇒ Validation de la compatibilité des outils de DRT et Pytorch

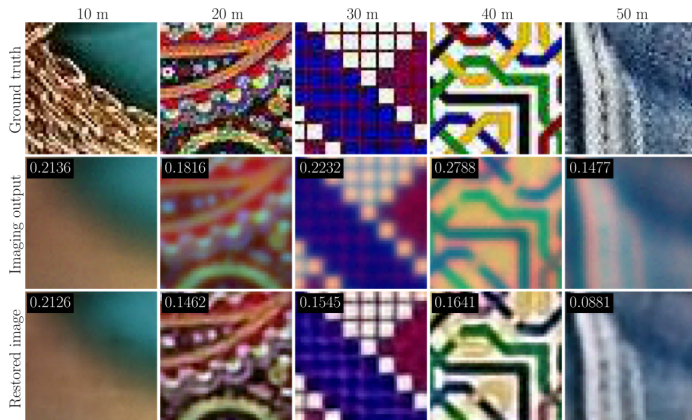
Qualité image avec et sans optimisation de la mise au point

Object position	Without optical optimization				With optical optimization			
	RMSE	R	G	B	RMSE	R	G	B
5 m	0.1364	0.1326	0.1381	0.1382	0.1275	0.1274	0.1283	0.1266
10 m	0.1067	0.1094	0.1089	0.1017	0.1151	0.1169	0.1160	0.1122
15 m	0.0464	0.0514	0.0417	0.0453	0.0801	0.0838	0.0798	0.0765
20 m	0.0453	0.0385	0.0459	0.0506	0.0483	0.0547	0.0475	0.0416
25 m	0.0710	0.0668	0.0734	0.0724	0.0482	0.0504	0.0434	0.0504
30 m	0.0990	0.0914	0.1041	0.1010	0.0509	0.0482	0.0509	0.0534
35 m	0.1110	0.0983	0.1151	0.1184	0.0442	0.0380	0.0452	0.0486
40 m	0.1157	0.1021	0.1191	0.1246	0.0466	0.0401	0.0479	0.0511
45 m	0.1180	0.1044	0.1211	0.1271	0.0532	0.0475	0.0544	0.0571
50 m	0.1195	0.1062	0.1227	0.1284	0.0618	0.0567	0.0632	0.0650
Average	0.0969	0.0901	0.0990	0.1007	0.0676	0.0664	0.0677	0.0682

⇒ L'optimisation conjointe permet d'adapter la position du plan de mise au point

Modèle optique complet et différentiable

Performance qualitative de restauration d'images



Conclusion

- Développements d'outils dédiés de DRT et de traitement par réseaux de neurones
- Validation de leur compatibilité sur un exemple jouet
- Permet de considérer l'ensemble des paramètres du système optique
- Sans approximation de lentille mince, approximation paraxiale, permet l'étude de champ importants

Perspectives (thèse de R. Leroy)

- Optimisation d'un plus grand nombre de paramètres
- Variation du champ (quid de la couche optique ?)
- Application à l'estimation de la 3D par DFD
- Analyse de cette nouvelle approche de co-conception

- 1 Introduction
- 2 Quelques notions utiles sur les réseaux de neurones
- 3 Quelques exemples de l'état de l'art
- 4 DeepCodesign à l'ONERA
- 5 Conclusion et perspectives

Conclusion

- L'utilisation de réseaux de neurones amène à de nouvelles méthodes de co-conception
- Pose la question de la différentiabilité du modèle de formation d'image
- Grande variété d'applications possibles (cf intervention de G. Wetzstein à la journée co-conception)

Enjeux

- "Disparition" du modèle de performance analytique
- Comment interpréter cette co-conception ? Comprendre l'influence de composant non conventionnels ?
- Comment "maîtriser" l'optimisation ?
- Comment inclure des contraintes physiques dans l'optimisation ?

"De nos jours, les images ne sont plus vues par des humains mais par des machines"

Devrons-nous toujours justifier l'intérêt de l'approche de co-conception ?



**RÉPUBLIQUE
FRANÇAISE**

*Liberté
Égalité
Fraternité*

ONERA

THE FRENCH AEROSPACE LAB

Merci de votre attention !

Des questions ?

www.onera.fr

Bibliographie

- [1] S. Elmalem, R. Giryes et E. Marom,
Learned phase coded aperture for the benefit of depth of field extension,
Opt. Express, 26 (12), pp. 15316–15331 (Jun 2018).
- [2] V. Sitzmann, S. Diamond, Y. Peng, X. Dun, S. Boyd, W. Heidrich,
F. Heide et G. Wetzstein,
*End-to-end Optimization of Optics and Image Processing for Achromatic
Extended Depth of Field and Super-resolution Imaging*,
ACM Trans. Graph. (SIGGRAPH) (2018).
- [3] U. Akpınar, E. Sahin et A. Gotchev,
Learning Optimal Phase-Coded Aperture for Depth of Field Extension,
Dans *2019 IEEE International Conference on Image Processing (ICIP)*,
pp. 4315–4319 (2019).
- [4] Y. Peng, Q. Sun, X. Dun, G. Wetzstein, W. Heidrich et F. Heide,
Learned Large Field-of-View Imaging With Thin-Plate Optics,
ACM Trans. Graph. (SIGGRAPH Asia), (6) (2019).

- [5] H. Haim, S. Elmalem, R. Giryas, A. M. Bronstein et E. Marom, *Depth Estimation From a Single Image Using Deep Learned Phase Coded Mask*, IEEE Transactions on Computational Imaging, 4 (3), pp. 298–310 (2018).
- [6] J. Chang et G. Wetzstein, *Deep optics for monocular depth estimation and 3d object detection*, Dans *Proceedings of the IEEE International Conference on Computer Vision*, pp. 10193–10202 (2019).
- [7] Y. Wu, V. Boominathan, H. Chen, A. Sankaranarayanan et A. Veeraraghavan, *PhaseCam3D — Learning Phase Masks for Passive Single View Depth Estimation*, Dans *2019 IEEE International Conference on Computational Photography (ICCP)*, pp. 1–12 (2019).
- [8] E. Nehme, D. Freedman, R. Gordon, B. Ferdman, L. E. Weiss, O. Alalouf, T. Naor, R. Orange, T. Michaeli et Y. Shechtman, *DeepSTORM3D : dense 3D localization microscopy and PSF design by deep learning*, Nature Methods, 17 (7), p. 734–740 (Jun 2020).

- [9] H. Ikoma, C. M. Nguyen, C. A. Metzler, Y. Peng et G. Wetzstein, *Depth from Defocus with Learned Optics for Imaging and Occlusion-aware Depth Estimation*, IEEE International Conference on Computational Photography (ICCP) (2021).
- [10] J. Chang, V. Sitzmann, X. Dun, W. Heidrich et G. Wetzstein, *Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification*, Scientific Reports (2018).
- [11] C. Metzler, H. Ikoma, Y. Peng et G. Wetzstein, *Deep Optics for Single-shot High-dynamic-range Imaging*, Dans *Proc. CVPR* (2020).
- [12] R. Horisaki, Y. Okamoto et J. Tanida, *Deeply coded aperture for lensless imaging*, Optics Letters, 45 (11), pp. 3131–3134 (2020).
- [13] G. Wetzstein, A. Ozcan, S. Gigan, S. Fan, D. Englund, M. Soljačić, C. Denz, D. A. Miller et D. Psaltis, *Inference in artificial intelligence with deep optics and photonics*, Nature, 588 (7836), pp. 39–47 (2020).

- [14] Y. Peng, S. Choi, N. Padmanaban et G. Wetzstein,
Neural Holography with Camera-in-the-loop Training,
ACM Trans. Graph. (SIGGRAPH Asia) (2020).
- [15] M. Carvalho, B. Le Saux, P. Trouve-Peloux, A. Almansa et
F. Champagnat,
*Deep Depth from Defocus : how can defocus blur improve 3D estimation
using dense neural networks ?*,
Dans *Proceedings of the European Conference on Computer Vision
(ECCV) Workshops* (September 2018).
- [16] P. Trouvé-Peloux, J. Sabater, A. Bernard-Brunel, F. Champagnat, G. L.
Besnerais et T. Avignon,
Turning a conventional camera into a 3D camera with an add-on,
Appl. Opt., 57 (10), pp. 2553–2563 (Apr 2018),
<http://ao.osa.org/abstract.cfm?URI=ao-57-10-2553>.
- [17] K. Zhang, W. Zuo, S. Gu et L. Zhang,
Learning deep CNN denoiser prior for image restoration,
Dans *Proceedings of the IEEE conference on computer vision and
pattern recognition*, pp. 3929–3938 (2017).

- [18] X. Mao, C. Shen et Y.-B. Yang,
Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections,
Dans *Advances in neural information processing systems*, pp. 2802–2810
(2016).
- [19] J.-B. Volatier, Álvaro Menduiña Fernández et M. Erhard,
Generalization of differential ray tracing by automatic differentiation of computational graphs,
J. Opt. Soc. Am. A, 34 (7), pp. 1146–1151 (Jul 2017).