



HAL
open science

Super-résolution par apprentissage automatique guidé par la physique : évaluation pour la convection turbulente

Kevin Doria, Anne Sergent, Didier Lucor

► **To cite this version:**

Kevin Doria, Anne Sergent, Didier Lucor. Super-résolution par apprentissage automatique guidé par la physique : évaluation pour la convection turbulente. 25ème Congrès Français de Mécanique (CFM 2022), AFM, Aug 2022, Nantes, France. pp.159-168. hal-04400566

HAL Id: hal-04400566

<https://hal.science/hal-04400566>

Submitted on 17 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Super-résolution par apprentissage automatique guidé par la physique : évaluation pour la convection turbulente

K. DORIA^a, A. SERGENT^{a,b}, D. LUCOR^a

a. Univ. Paris-Saclay, CNRS, Lab. Interdisciplinaire des Sciences du Numérique (LISN), 91405 Orsay

b. Sorbonne Université, Faculté des Sciences et Ingénierie, UFR Ingénierie, F-75005 Paris

kevin.doria@lisn.fr; anne.sergent@lisn.fr; didier.lucor@lisn.fr

Résumé :

On s'intéresse aux méthodes de prédiction d'un champ de température dans un écoulement turbulent de Rayleigh-Bénard ($Ra = 2 \cdot 10^9$) à partir d'images d'ombroscopie. Pour cela, on a recours à des neurones convolutifs assemblés sous la forme d'un auto-encodeur ou d'un réseau en U (U-Net). Une architecture dissymétrique est proposée pour obtenir des images plus détaillées que les données et à moindre coût en paramètres en comparaison avec des auto-encodeurs. De premiers résultats prometteurs sont obtenus.

Abstract :

We are interested in methods for predicting a temperature 2D-field in a turbulent Rayleigh-Benard flow ($Ra = 2 \cdot 10^9$) from shadowgraphs. For this purpose, convolutional neurons forming auto-encoder or U-Net are used. An asymmetric architecture is proposed to obtain with more spatial detail than the data and with lower parameter costs compared to auto-encoders. First promising results are obtained.

Mots clefs : Turbulence, convection, apprentissage automatique, réseaux de neurones

1 Introduction

Résoudre par simulation haute fidélité toutes les échelles spatio-temporelles d'un écoulement turbulent peut se révéler une approche extrêmement coûteuse. D'autres approches, telles que des mesures expérimentales (PIV par exemple) permettent d'acquérir des champs de données, mais le plus souvent celles-ci sont bruitées et partielles en espace (des plans bidimensionnels complets du système, ou des zones plus restreintes avec une meilleure résolution). Nous cherchons dans ce travail à développer une méthode d'apprentissage automatique permettant de reconstruire des échelles plus fines que les données, voire d'inférer des quantités non-mesurées, grâce à la connaissance de la physique du problème considéré.

Notre cas d'étude est la convection turbulente de Rayleigh-Bénard, où un fluide (ici de l'eau) contenu dans une cavité, est chauffé par la surface inférieure de celle-ci. Il s'agit d'un écoulement canonique

largement étudié dans la littérature aussi bien d'un point de vue fondamental, que pour les larges domaines d'application possibles (nucléaire, refroidissement passif, écoulements naturels, etc.). Lorsque l'écoulement est turbulent, il est caractérisé par un large spectre d'échelles spatiales et temporelles, du fait de la variété des structures cohérentes qui le constitue : la circulation à grande échelle, les panaches et les structures turbulentes (voir [2] pour plus de détails).

Nous retenons ici des architectures employant des neurones à convolution (CNN) assemblés sous la forme d'un auto-encodeur (voir par ex. [5]) ou d'un réseau de type U-Net (par ex. [3]), dont l'objectif est de prédire un champ de sortie en fonction d'un champ d'entrée de même dimension. Cependant, dans notre cas, le champ prédit est d'une toute autre nature que le champ d'entrée. Le champ d'entrée correspond à une image d'ombroscopie, telle que celles qui pourraient être enregistrées expérimentalement mais sans bruit car calculées par simulation numérique directe. Il s'agit d'une image compressant l'information tridimensionnelle du champ de température en une image bidimensionnelle, du fait de la modification par le champ de température de l'indice optique de l'eau. Ainsi, les rayons lumineux parallèles illuminant le dispositif sont déviés suivant la structure de l'écoulement (figure 1). Il s'agit de données très riches que nous avons déjà exploitées précédemment pour estimer la vitesse des panaches remplissant la cavité [1]. Le champ de température dans le plan vertical à mi-profondeur est le champ 2D prédit cible. L'ombroscopie serait alors un moyen de mesure non-intrusif du champ de température, difficilement mesurable, en particulier lorsque l'écoulement est turbulent.

Le type d'architecture que nous visons a déjà été employé avec succès pour des écoulements turbulents. Les auto-encodeurs (AE) peuvent être employés pour faire de la super-résolution, i.e. reconstruire des champs prédits sur une grille plus fine que les champs d'entrée, ou encore pour construire des modèles réduits à partir de l'espace latent (couche centrale du AE de dimension minimale) qui peut être vu comme une base de dimension réduite ([5]). D'autre part, les architectures en forme de U (aussi appelées U-Net) sont des variantes des AE, dans lesquels un lien de concaténation est établi entre les couches de convolution et de déconvolution, transférant ainsi directement les détails physiques de l'écoulement capturés par chaque couche de convolution à la couche de déconvolution de niveau équivalent. C'est une façon de transférer les informations physiques de l'encodeur vers le décodeur. Les U-Net sont très utilisés pour des tâches de segmentation, et ont été employés avec succès pour identifier les super-structures à grande échelle en convection turbulente de Rayleigh-Bénard ([3]). Une autre façon de re-créeer du détail serait de contraindre l'apprentissage en imposant une information complémentaire au réseau, comme des informations physiques, à différents niveaux : espace latent, fonction de perte, etc. (voir [4] pour un exemple appliqué à la convection turbulente).

Dans cet article, nous proposons de construire des réseaux de neurones en vue de reconstruire une coupe du champ de température à partir de partir d'une image d'ombroscopie. Pour cela, nous partons d'une architecture simple d'auto-encodeur pour aller vers des architectures plus complexes, dans le but final de construire un réseau capable de réaliser de la super-résolution. Ici, l'information physique est passée dans le réseau via les étapes de concaténation présente dans le U-Net. Différentes architectures et profondeurs de réseau sont testées. Il s'agit d'une étape préalable avant l'ajout de contraintes physiques.

2 Configuration physique et code numérique

On s'intéresse à la convection naturelle de Rayleigh-Bénard se développant dans une cellule parallélépipédique de géométrie $L \times W \times H$ avec $L/H = 1$ et $W/H = 0.5$ et remplie d'eau. Les parois horizontales sont isothermes telles que $T_{bas} - T_{haut} = \Delta T = 1$. Les parois verticales sont adiabata-

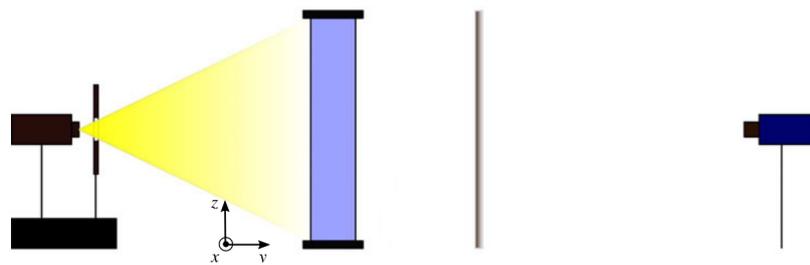


FIGURE 1 – Principe de l’ombroscopie. D’après [1].

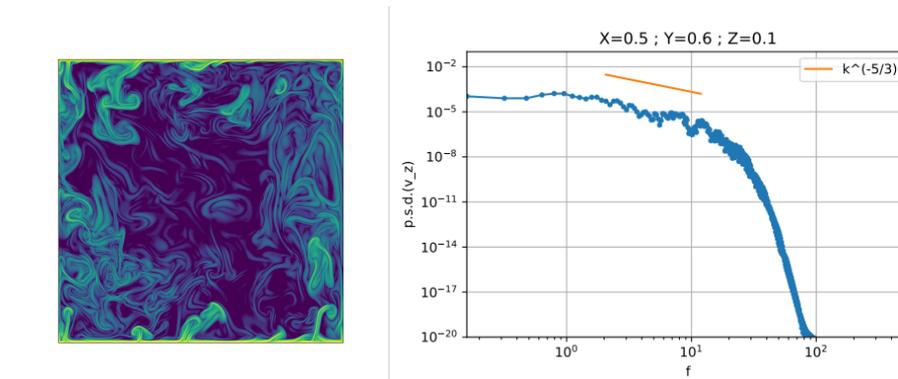


FIGURE 2 – Gauche : exemple de champ DNS à $Ra = 2 \cdot 10^9$. Coupe verticale à mi-profondeur. Le domaine physique 3D d’apprentissage coupant le plan de coupe est matérialisé par le rectangle jaune. Droite : Energie spectrale des fluctuations de vitesse verticale en un point.

tiques. Le problème dépend de deux nombres sans dimension : le nombre de Prandtl ($Pr = \nu/\alpha$) et le nombre de Rayleigh ($Ra = (g\beta\Delta TH^3)/(\alpha\nu)$), où ν est la viscosité cinématique, α la diffusivité thermique, β le coefficient de compressibilité et g la gravité ($\vec{g} = -g\vec{z}$).

L’objectif de ce travail étant de chercher l’architecture de réseaux de neurones la plus à même de reconstruire des coupes bidimensionnelles d’un écoulement tridimensionnel turbulent à partir de champ d’ombroscopie, nous ne retenons ici qu’une configuration particulière : $Ra = 2 \cdot 10^9$ et $Pr = 4.4$ correspondant à un écoulement très turbulent (voir figure 2).

L’écoulement est gouverné par les équations de Navier-Stokes sous l’approximation de Boussinesq, adimensionnées par la vitesse de couche limite de convection naturelle, soit $\alpha\sqrt{Ra}/H$. Les résultats numériques sont obtenus par simulation numérique directe en utilisant SUNFLUIDH, le code développé au laboratoire. Il s’agit d’un code volumes finis d’ordre deux en espace. L’avancée en temps est basée sur un schéma d’Euler retardé. On utilise une méthode de prédiction-projection pour imposer l’incompressibilité. Les systèmes d’Helmholtz sont résolus par des méthodes de Thomas et l’équation de pression par une méthode multigrille. Le code de calcul est parallélisé par décomposition de domaine. Le maillage regroupe $640 \times 320 \times 832$ points répartis sur 512 sous-domaines et le calcul est mené avec un pas de temps sans dimension égal à $4.5 \cdot 10^{-4}$.

3 Méthodes

3.1 Réseaux de neurones

L'architecture de réseau de neurones constituant la base de ce travail est celle de l'auto-encodeur [5], pouvant être vu comme l'ensemble d'un encodeur et d'un décodeur. L'encodeur cherche à réduire les dimensions des données d'entrée afin de les représenter dans un nouvel espace de dimension inférieur appelé espace latent tandis que le décodeur cherche à reconstruire les données à partir de l'encodage (voir Figure 3). Dans l'optique d'aider le processus de reconstruction des données, des cartes d'activations de l'encodage sont ajoutées dans les couches de convolution du décodeur afin de former une architecture de type U-Net (Figure 4). De cette manière, les sorties ne sont pas seulement reconstruites à partir d'une donnée compressée dans l'espace latent mais également à partir de détails progressivement isolés pendant l'encodage. Enfin, la dernière architecture testée concerne directement l'objectif de réaliser un réseau capable d'avoir des propriétés de super-résolution avec l'utilisation d'un U-Net dissymétrique renvoyant des images qui auront une meilleure résolution que celles en entrée (Figure 5).

Ces trois architectures ont été construites hiérarchiquement, l'une à partir de l'autre. De ce fait, elles partagent certains de leurs hyperparamètres et procédés de traitement des données. L'encodage se fait par convolution avec des filtres de taille 3x3 sauf pour la première couche où un filtre 5x5 est utilisé. Dans tous les cas, la taille des images est conservée par du zero-padding. La réduction de la taille des images est assuré par des couches de pooling moyenné. Pour les auto-encodeurs, la dimension de l'espace latent est 128. De l'autre côté des réseaux, la reconstruction des données se fait par convolution transposée avec des filtres 2D de taille 3x3. On retrouve dans chaque couche (de l'encodeur et du décodeur), une normalisation des batchs avant le passage dans la fonction d'activation qui est la fonction *Exponential Linear Unit* (ELU) dans notre étude. Un algorithme de type Adam est utilisé pour l'optimisation d'une erreur quadratique moyenne (ici appelée ε_2 et définie à la section 3.2) sur 2000 epochs avec des lots de taille 32. Le taux d'apprentissage est évolutif en commençant à $1e - 3$, puis valant $5e - 4$ à partir de la 500^e epoch et $1e - 4$ au delà de la 1000^e epoch. Enfin, du Dropout à hauteur de 10% est appliqué à chaque couche de l'encodeur pour limiter les risque de sur-apprentissage. On notera que chaque réseau est entraîné deux fois indépendamment afin de limiter les effets de l'initialisation aléatoire des poids (réalisé par méthode de Glorot uniforme) notamment.

Les réseaux sont entraînés sur des cartes GPU de type NVIDIA Tesla V100. La méthode proposée est mise en œuvre à l'aide de la bibliothèque TensorFlow.

3.2 Données

La base de données consiste en une accumulation de champs instantanés d'ombroscopie I_{2D} pour l'entrée telle que $I_{2D} = \int_y (\nabla_x^2 + \nabla_z^2) \theta dy$ et de la coupe verticale dans le plan médian ($y = 0.25$) du champ de température sans dimension ($\theta \in [0 : 1]$). Il s'agit de champs interpolés des données DNS sur un maillage isotrope de taille 256×256 . La série temporelle contient $n_t = 4714$ échantillons enregistrés tous les 0.02025 unités de temps.

Les données d'entrée sont standardisées, i.e correspondent à la fluctuation réduite du champ d'ombroscopie, soit $I_{input}(t) = \frac{I_{2D}(t) - \bar{I}_{2D}}{\sigma(I_{2D})}$, avec \bar{v} la moyenne temporelle et $\sigma(v)$ l'écart-type d'un champ quelconque v . Dans le processus d'optimisation, ces données seront regroupées par mini-batch regroupant 32 instantanés.

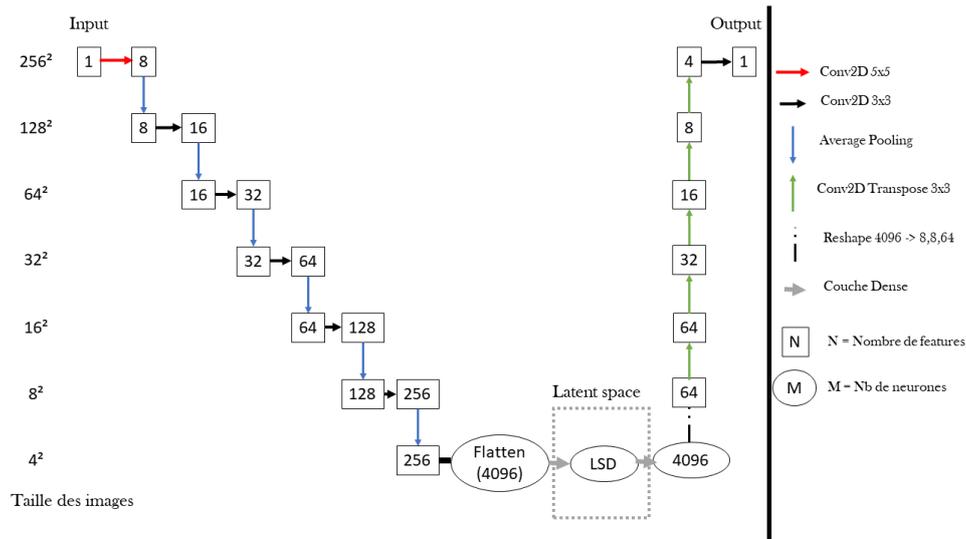


FIGURE 3 – Architecture de type Auto-encodeur. Ici avec 6 couches pour l’encodeur et 5 couches pour le décodeur

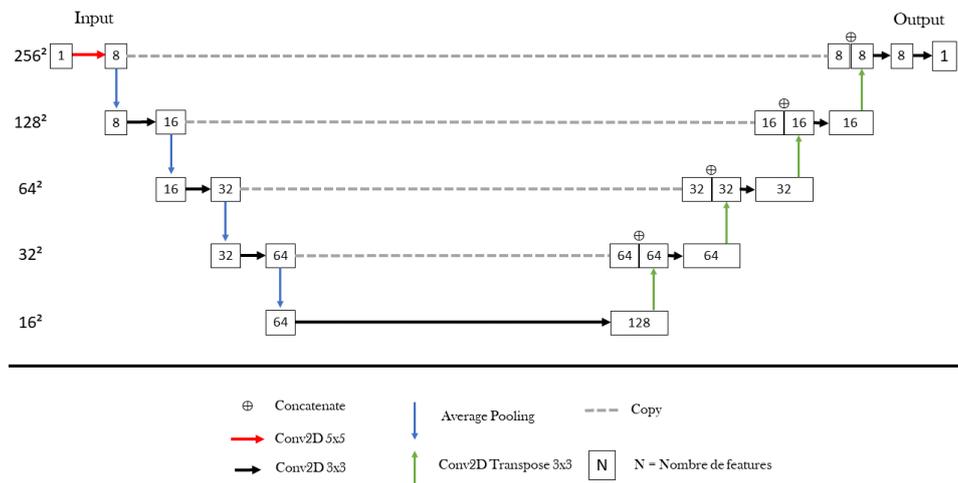


FIGURE 4 – Architecture de type U-Net. Ici avec 5 couches.

Le jeu de données (appelé DB), après un mélange aléatoire appelé *shuffle*, est découpé en trois parties distinctes : la base de données d’entraînement sur laquelle le réseau est entraîné (représentant 70% de DB), la base de données de validation (pour 15% de DB) et la base de données de test (15% de DB) pour l’évaluation des erreurs sur le réseau entraîné.

La comparaison des erreurs de reconstruction obtenues se fait sur quatre estimateurs : l’erreur quadratique moyenne ε_2 , l’erreur absolue moyenne ε_1 , l’erreur moyenne sur la température moyenne de la cavité : ε_m et l’erreur moyenne sur l’écart type spatial de la température de la cavité : ε_σ :

$$\varepsilon_2 = \frac{1}{n_{ech}} \sum_{ech} (\theta_{DB} - \tilde{\theta})^2$$

$$\varepsilon_1 = \frac{1}{n_{ech}} \sum_{ech} |\theta_{DB} - \tilde{\theta}|$$

$$\varepsilon_m = \frac{1}{n_t} \sum_t \frac{|\langle \theta_{DB} \rangle_S - \langle \tilde{\theta} \rangle_S|}{\langle \theta_{DB} \rangle_S}$$

$$\varepsilon_\sigma = \frac{1}{n_t} \sum_t \frac{|\sigma_S(\theta_{DB}) - \sigma_S(\tilde{\theta})|}{\sigma_S(\theta_{DB})}$$

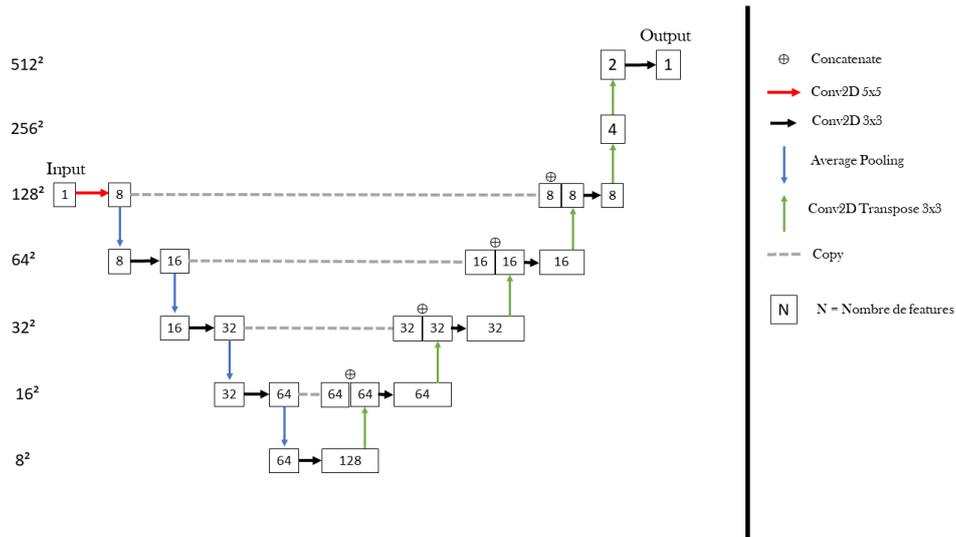


FIGURE 5 – Architecture de type U-Net modifiée pour faire de la super-résolution. Ici avec 5 couches pour l’encodeur et 7 couches pour le décodeur.

où n_{ech} est le nombre de points de l’échantillon temps et espace confondus (soit ici 61,5 millions de points) de DB test, $\langle v \rangle_S$ la moyenne spatiale appliquée au champ 2D instantané v sur les 256² points du maillage interpolé, $\sigma_S(v)$ l’écart type spatial correspondant, θ_{DB} le champ de température de la base de données de test, $\tilde{\theta}$ est le champ de température prédit.

Les visualisations présentent le champ de température prédit sous une forme réduite par rapport aux champs DNS, soit

$$\hat{\theta} = \frac{\tilde{\theta} - \langle \tilde{\theta}_{DNS} \rangle_S}{\sigma_S(\theta_{DNS})}$$

Deux instants particuliers de la base de données de test sont retenus pour mettre en évidence les différences spatiales de la prédiction. Ils correspondent respectivement aux instants t_{max} (et t_{min}) où le transfert de chaleur instantané traversant la cavité et caractérisé par le nombre de Nusselt est maximal $Nu_{max} = 82.0$ (minimal $Nu_{max} = 72.9$).

4 Résultats

Modèles	n_e	n_d	N_p	Input	Sortie	ε_2	ε_1	ε_m	ε_σ
AE-6	6	6	1 509 099	256 ²	256 ²	6.825e-5	4.681e-3	0.0261 %	4.01 %
AE-7	7	7	2 348 651	256 ²	256 ²	7.36e-5	4.99e-3	0.009 %	3.95 %
UNet-5	5	5	295 499	256 ²	256 ²	11.56e-5	6.578e-3	0.0152 %	5.37 %
UNet-6	6	6	1 181 771	256 ²	256 ²	4.63e-5	4.301e-3	0.015 %	1.97 %
UNet-7	7	7	4 723 787	256 ²	256 ²	5.885e-5	4.780e-3	0.0417 %	3.01 %
UNet-6	6	6	1 181 771	128 ²	128 ²	7.43e-5	5.178e-3	0.0348 %	5.33 %
UNet-56	5	6	295 763	128 ²	256 ²	11.03e-5	6.23e-3	0.0359 %	5.73 %
UNet-57	5	7	295 823	128 ²	512 ²	11.28e-5	6.299e-3	0.0269 %	3.83 %

TABLE 1 – Comparaison des différentes architectures testées. n_e, n_d : nombre de couches de convolution pour l’encodeur et le décodeur respectivement, N_p : nombre de paramètres entraînaibles et dimension des champs d’entrée et de sortie.

L'ensemble des indicateurs de précision sont rassemblés dans le tableau 1 pour toutes les architectures testées. De manière générale, on voit qu'augmenter le nombre de couches des réseaux améliore la précision des résultats aussi bien pour les AE que pour les U-Net, même si UNet-7 ne semble pas plus efficace que UNet-6, bien qu'il demande d'optimiser 3 fois plus de paramètres.

En revanche, on voit clairement l'intérêt de transférer des informations physiques par le biais des concaténations dans les U-Net en comparaison avec les auto-encodeurs simples. Une amélioration notable de la distribution spatiale du champ de température est visible sur l'erreur moyenne quadratique ε_2 et ε_σ .

Ceci est confirmé par les champs prédits présentés aux figures 6 et 7. On remarque une meilleure représentation des détails avec les AE et U-Net présentant le plus de couches, et en particulier avec U-Net7. Au niveau du cœur du domaine, on distingue peu de différences mais les différences se situent près des bords. Les contrastes sont ainsi légèrement plus marqués pour les prédictions des U-Nets par rapport à celles des Auto-encodeurs, et surtout plus fins, par exemple lorsque l'on regarde les panaches descendants (en bleu) en haut de la paroi droite de la figure 6 ou encore le panache ascendant (rouge) en bas à droite de la figure 7.

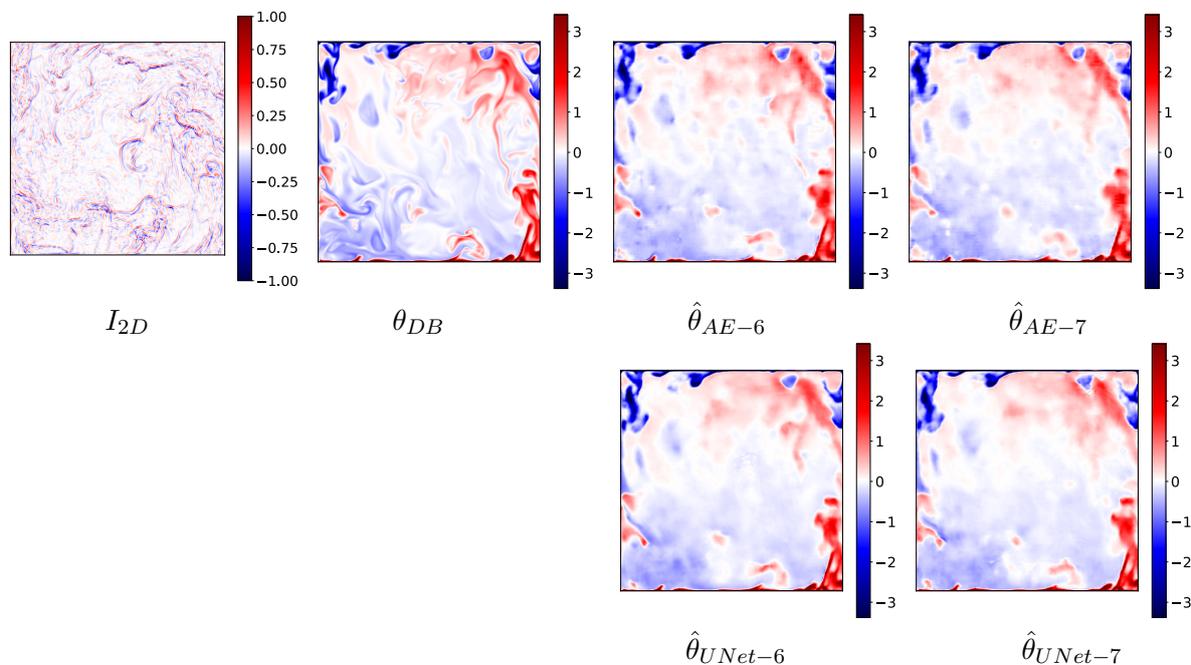


FIGURE 6 – Prédiction du champ de température instantané correspondant au nombre de Nusselt instantané maximal de la base de test suivant l'architecture retenue. Gauche : champ d'ombroscopie en entrée et champ de température cible. Droite : champs prédits réduits par auto-encodeur (haut) et UNet (bas).

La seconde partie de l'analyse concerne les capacités du réseau de prédire les échelles spatiales non vues (et donc plus fines) par les données d'entrée, ie les propriétés de super-résolution du réseau. Pour cela, nous avons proposé des réseaux U-Net dissymétriques, où la branche du décodeur est augmentée d'une ou deux couches convolutives sans concaténation avec l'encodeur, les cartes d'activation étant de dimensions supérieures aux données d'entrée.

Cette fois-ci, on part de données d'entrée de taille 128^2 obtenues à partir des résultats DNS interpolés sur un maillage isotrope. La prédiction porte sur une coupe du champ de température à une résolution supérieure (256^2 pour UNet-56 ou 512^2 pour UNet-57). À des fins de comparaison, on entraîne également un U-Net à 6 couches à prédire la température instantanée sur un maillage 128×128 à partir de données

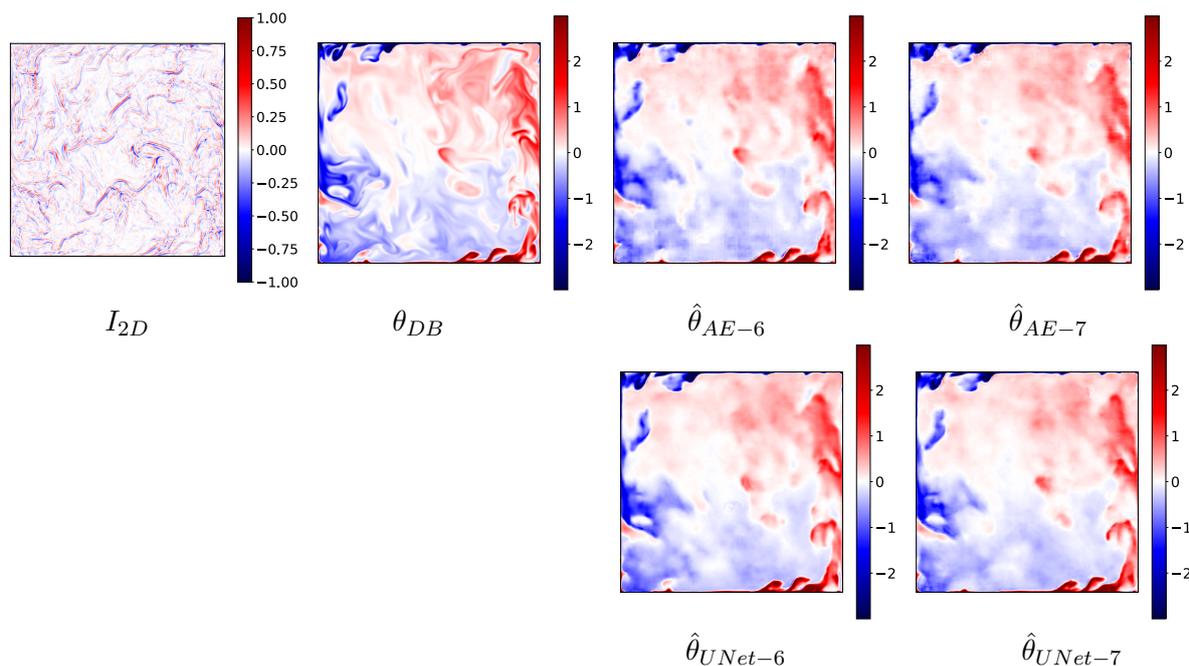


FIGURE 7 – Prédiction du champ de température instantané correspondant au nombre de Nusselt instantané minimal de la base de test suivant l’architecture retenue. Voir figure 6 pour le détail de la légende.

de même taille. Les résultats sont présentés dans le tableau 1 et sur les figures 8 et 9.

On remarque tout d’abord que l’ordre de grandeur des erreurs est conservé pour les réseaux dissymétriques (UNet-56 et UNet-57) malgré une difficulté de prédiction supérieure et une quantité de paramètres inférieure. Avec des prédictions sur des maillages 256^2 ou 512^2 , plus d’extrema locaux de température peuvent apparaître ce qui explique l’augmentation des erreurs point à point $\varepsilon_{1,2}$ et de l’erreur sur la moyenne ε_m . Néanmoins, la présence de ceux-ci permet aussi aux réseaux de mieux capturer les structures cohérentes de l’écoulement ce qui se traduit par une diminution de ε_σ entre U-Net6, U-Net56 et puis entre U-Net56 et U-Net57.

L’interprétation des figures 8 et 9 n’est pas simple. Suivant les instantanés considérés et les zones de l’écoulement, certains détails sont mieux représentés que d’autres. On peut cependant remarquer que les architectures dissymétriques demandent presque quatre fois moins de paramètres à entraîner pour un résultat tout à fait comparable. Cela illustre la nécessité de produire des indicateurs plus fins pour qualifier la qualité de la prédiction.

5 Conclusions

Cette étude est basée sur plusieurs architectures de réseaux à convolution dans le but de prédire une coupe 2D du champ de température à partir d’images d’ombroscopie de l’écoulement turbulent de Rayleigh-Bénard. Pour cela, nous nous sommes concentrés sur des auto-encodeurs et des réseaux de type U-Net. Une évolution vers des U-Net dissymétriques (le décodeur contient plus de couches que l’encodeur) a été proposé pour tester d’obtenir des champs prédits super-résolus par rapport aux données d’entrée.

Comme attendu, nous observons que les réseaux les plus profonds sont les plus efficaces, mais ils demandent plus de paramètres à optimiser. Le transfert d’information physique de l’encodeur vers le décodeur contenu dans les U-Net par construction permet une meilleure prédiction des champs de sortie.

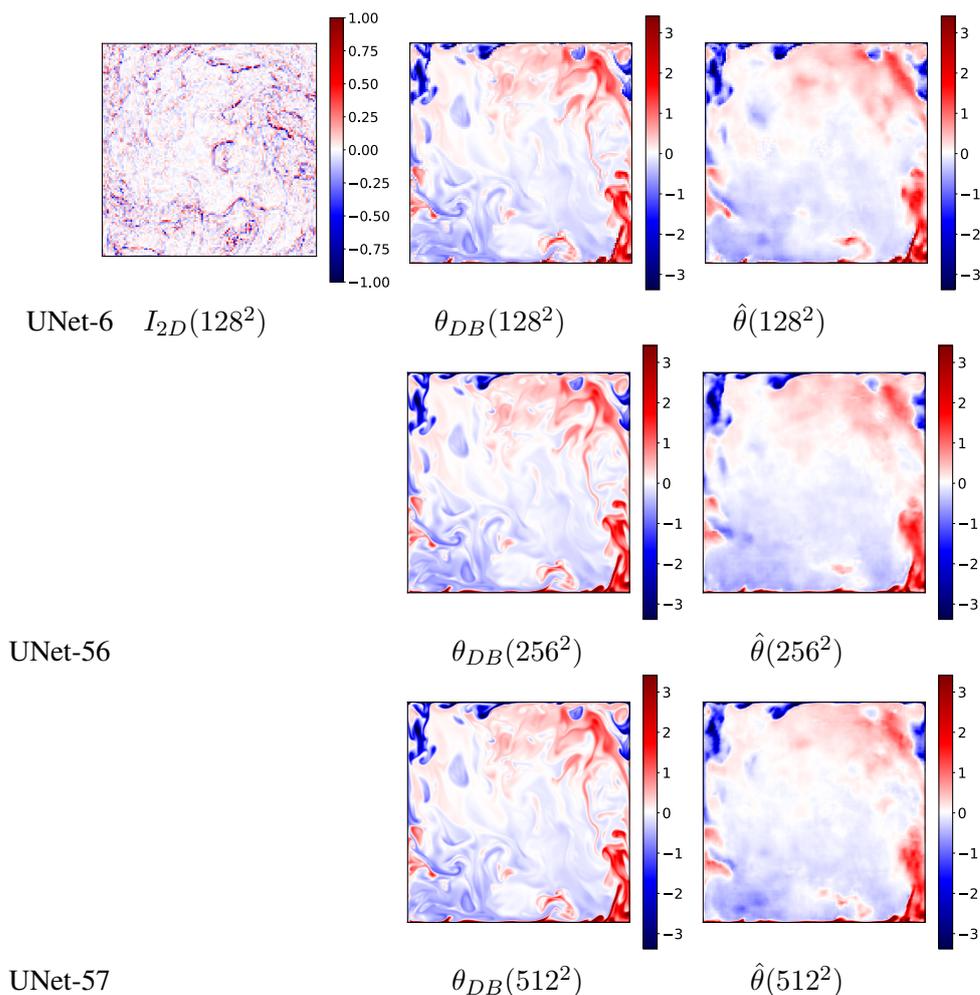


FIGURE 8 – Tests de super-résolution pour différentes architectures à l’instant de nombre de Nusselt maximal. Gauche : champ d’ombroscopie sur un maillage 128^2 . Milieu : champ de température cible. Droite : champ prédit réduit. Le maillage est indiqué entre parenthèse.

Les réseaux dissymétriques proposés offrent une meilleure prédiction générale de l’écoulement avec une erreur sur les déviations standards décroissante avec l’augmentation de la super-résolution, et cela pour un coût en paramètres réduits d’un facteur 4. Cependant, ces conclusions demandent à être affinées avec des métriques qualifiant la distribution spatiale de la température de manière plus fine.

Cette approche est donc prometteuse mais visiblement insuffisante. Il est nécessaire de complexifier davantage la structure des réseaux considérés afin d’améliorer la représentation des détails sans augmenter le nombre de paramètres de façon considérable. Une des pistes privilégiées est l’intégration de connaissances physiques supplémentaires dans l’apprentissage statistique, par exemple en forçant le champ de température à respecter les équations d’un modèle physique via la fonction coût. Nous avons en effet pu constater dans une étude précédente le fort potentiel d’un apprentissage guidé la physique [4].

6 Remerciements

La base de données DNS a été construite en utilisant l’accès accordé par GENCI aux ressources HPC de l’IDRIS sous l’allocation 2a0326. Les réseaux ont été entraînés sur le cluster GPU Lab-IA de l’Université Paris-Saclay dédié à l’Intelligence Artificielle.

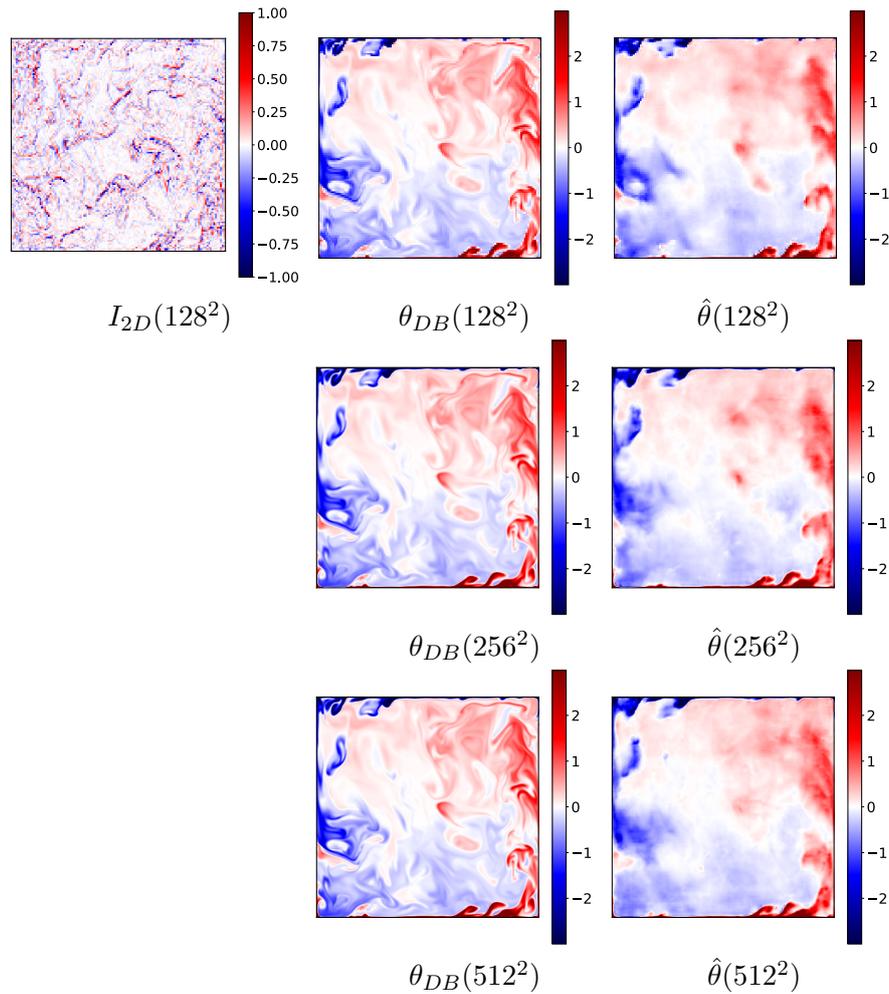


FIGURE 9 – Idem figure 8 pour l’instant de nombre de Nusselt minimal.

Références

- [1] M. Belkadi, L. Guislain, A. Sergent, B. Podvin, F. Chillà and J. Salort, Experimental and numerical shadowgraph in turbulent Rayleigh–Bénard convection with a rough boundary : investigation of plumes, *J. Fluid Mech.*, vol. 895 (2020), A7.
- [2] F. Chillà J. Schumacher, New perspectives in turbulent Rayleigh–Bénard convection. *Eur. Phys. J. E*, 35 (2012), 58.
- [3] E. Fonda, A. Pandey, J. Shumacher and K.R. Sreenivasan, Deep learning in turbulent convection networks, *Proceedings of the National Academy of Sciences*, 116 (2019), 8667-8672.
- [4] D. Lucor, A. Agrawal and A. Sergent, Simple computational strategies for more effective physics-informed neural networks modeling of turbulent natural convection, *J. Computational Physics*, 456 (2022), 111022.
- [5] M. Morimoto, K. Fukami, K. Zhang, A.G. Nair and K. Fukagata, Convolutional neural networks for fluid flow analysis : toward effective metamodeling and low dimensionalization. *Theor. Comput. Fluid Dyn.* 35 (2021), 633–658.