

Évaluation des CNN pour la super-résolution de l'IRM de flux 4D

Simon PERRIN¹ Sébastien LEVILLY¹ Harold MOUCHÈRE¹ Jean-Michel SERFATY²

¹Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France

²Nantes Université, CHU Nantes, CNRS, INSERM, Institut du thorax, Nantes, France

Résumé – Dans le contexte de l'IRM de flux 4D, la fiabilité des biomarqueurs hémodynamiques est influencée par la résolution spatiale. La super-résolution résout ce problème en améliorant la résolution et en réduisant le bruit du champ de vitesse. Le succès récent de l'apprentissage profond dans le domaine du traitement d'image, et plus spécifiquement pour la super-résolution, a conduit au développement de méthodes pour le débruitage et la super-résolution d'IRM de flux 4D. Cependant, la plupart des solutions de l'état de l'art évaluent leurs performances sur des jeux de données privés, ce qui réduit la possibilité de les comparer objectivement entre elles. Dans ce travail, nous proposons une méthode de comparaison pour l'entraînement et l'évaluation de trois solutions CNN. Des simulations de mécanique des fluides numérique sont utilisées pour générer des données synthétiques d'IRM de flux 4D pour l'entraînement et l'évaluation des méthodes. Dans cette évaluation, nous analysons l'influence des fonctions de coût sur les réseaux, l'architecture des réseaux et le mécanisme de *channel attention*.

Abstract – Hemodynamics biomarkers reliability is known to be impacted by the spatial resolution in the context of 4D Flow MRI. Super-resolution tackles this issue by enhancing the resolution and reducing the velocity field noise. The recent success of deep learning in the field of image processing, and more specifically for super-resolution, has led to the development of super-resolution and denoising methods for 4D Flow MRI. However, most state-of-the-art solutions assess their performance on private/project datasets which reduce the possibility to objectively compare one from an other. In this work, we propose a benchmark for training and evaluating three CNN methods. Computational fluid dynamic simulations are used to generate synthetic 4D Flow MRI data for training and testing. In this benchmark, we analyze the influence of channel attention, loss functions on networks and network architectures.

1 Introduction

L'IRM de flux 4D [7] est un moyen non invasif de mesurer le flux sanguin dans le système cardiovasculaire humain. Cette acquisition fournit des images en 3D de l'anatomie et du flux sanguin tout au long du cycle cardiaque, et est par conséquent construite sur plusieurs battements de cœur. Ce type d'IRM souffre de bruit, d'artefacts de repliement et d'une résolution spatiale très faible, *eg.* 2.5mm isotropique (ISO) pour une aorte thoracique [7]. Ces inconvénients ont un impact sur l'évaluation des biomarqueurs hémodynamiques tels que le cisaillement à la paroi [6]. De plus, le temps d'acquisition, pour une image du thorax, est compris entre 10 et 15 minutes ce qui est contraignant en routine clinique.

Dans le domaine du traitement d'images, la super-résolution (SR) consiste à améliorer la résolution d'une image basse résolution (LR) à l'aide de méthodes basées sur l'interpolation ou l'apprentissage profond [1]. Dans le contexte de l'IRM de flux 4D, la super-résolution peut être utilisée pour deux stratégies : soit améliorer la résolution clinique afin d'obtenir une meilleure évaluation des biomarqueurs hémodynamiques, ou bien améliorer la résolution d'une acquisition accélérée, ce qui conduit à une résolution plus faible. Dans la littérature (Section 2), les performances de la plupart des solutions sont évaluées sur des ensembles de données privés ou dédiés à un projet, ce qui affecte toute analyse comparative. Par conséquent, nous proposons de définir une base commune pour comparer les méthodes de SR, *i.e.* de définir des ensembles de données d'entraînement et de test. Étant donné que l'acquisition d'IRM de flux 4D à haute résolution (HR) serait trop longue et potentiellement très impactée par du bruit, des artefacts et du

flou respiratoire, nous avons décidé de nous concentrer sur des simulations de la dynamique des fluides numérique (CFD) comme champ de vitesse de référence [3]. Dans la section 3, le processus de génération des données synthétiques d'IRM de flux 4D imitant la méthode d'acquisition est expliquée, ainsi que les métriques et les solutions évaluées. Dans la section 4, les résultats des différentes méthodes sont comparés sur une base de test commune. Enfin, nous proposons une analyse de l'influence de la fonction de coût, le mécanisme de *channel attention* (CA) et des architecture des réseaux sur la solution prédite.

2 État de l'art

Les méthodes d'apprentissage profond pour la super-résolution de l'IRM de flux 4D sont divisées en deux parties : les réseaux de neurones convolutifs (CNN) et les réseaux de neurones guidés par la physique (PINN) [8].

Les méthodes d'apprentissage profond pour la super-résolution utilisent principalement des CNN [5]. Basiquement, le CNN va prédire l'image super-résolue (SR) à partir d'une basse résolution (LR). Plusieurs couches de convolution sont utilisées pour extraire les caractéristiques de l'image d'entrée, qui sont ré-échantillonnées grâce à une méthode d'interpolation ou une couche dédiée. Pour finir, les caractéristiques sur-échantillonnées sont utilisées pour reconstruire l'image SR. Ferdian *et al.* (2020) [3] est le premier travail utilisant un CNN pour cette tâche. Un réseau résiduel est utilisé pour tirer avantages des connexions raccourci qui permettent d'avoir un entraînement plus stable et de meilleures performances

grâce à l'extraction de caractéristiques de bas niveaux. Shit *et al.* (2022) [9] introduit l'utilisation du mécanisme de *channel attention* pour cette tâche, ce type de mécanisme d'attention permet au réseau de porter son attention sur les caractéristiques extraites les plus intéressantes pour la tâche. Par ailleurs, le modèle utilise la *efficient wider activation* qui permet de réduire la complexité du modèle et gagner en efficacité. Dans les solutions précédentes, la tâche ne portait que sur une partie du système cardiovasculaire, généralement l'aorte thoracique. Ericsson *et al.* (2023) [2] propose d'utiliser des stratégies d'apprentissage ensembliste pour traiter l'IRM cérébrovasculaire, cardiaque et aortique. Pour la super-résolution d'IRM de flux 4D à l'aide de CNN, les méthodes de la littérature [3, 9, 2] utilisent des simulations CFD comme données d'entraînement car il est trop complexe, et prendrait beaucoup de temps, d'acquérir des IRMs de flux 4D HR fiables. Les images obtenues par les simulations CFD sont modifiées pour correspondre aux caractéristiques de cette acquisition.

Les PINN [8] sont des perceptrons multicouches qui intègrent des connaissances physiques par le biais d'équations aux dérivées partielles (EDP) utilisées dans la fonction de coût. En raison de son architecture, un PINN n'est pas en mesure de généraliser sur un ensemble de données, il doit être entraîné sur chaque nouvelle image, ce qui dans l'état actuel ne permet pas une utilisation clinique.

3 Méthodologie

3.1 Images d'apprentissage et de test

Pour former et tester les approches de manière supervisée, nous avons besoin de paires de données HR et LR. L'acquisition d'IRM de flux 4D HR est cliniquement trop complexe en raison du long temps d'acquisition. Nous avons donc choisi d'utiliser un ensemble de données publiques de simulations CFD pour générer des images synthétiques en utilisant la même méthodologie que [3]. Cet ensemble de données est composé de 3 simulations CFD, chacune d'entre elles utilisant une géométrie différente. La première géométrie, *aorta01*, a été acquise à partir d'un volontaire sain et *aorta02* et *aorta03* présentent une coarctation de l'aorte.

Les images obtenues lors de la CFD ne sont pas du même type que l'IRM de flux 4D, elles ne présentent aucun bruit et aucune limite de valeur *VENC* n'est définie. Pour générer les images synthétiques, les images complexes C sont d'abord calculées en utilisant les images de magnitude M et celle de vitesse V : $C = M e^{\frac{i\pi}{VENC} V}$, avec *VENC* le paramètre d'acquisition de l'IRM (*Velocity Encoding*). C est ensuite projetée dans l'espace k (domaine d'acquisition) en utilisant la transformée de Fourier. Puis, les hautes fréquences sont coupées (facteur de 2) et un bruit Gaussien est ajouté pour imiter la distribution de bruit présente dans une *vraie* IRM de flux 4D [4]. Enfin, une transformée de Fourier inverse permet de revenir dans le domaine de l'image. Les parties réelles et imaginaires sont séparées pour obtenir une image de magnitude et de phase sous-échantillonnée. Le paramètre d'acquisition *VENC* est choisi dans l'ensemble $[0.3, 0.6, 1, 1.5, 2, 3]$ m.s⁻¹ et est toujours supérieur à la vitesse maximale dans l'image pour éviter le repliement de la vitesse. Un *VENC* est choisi pour chaque image temporelle contrairement aux données réelles où qu'une seule valeur est choisie pour toute l'acquisition.

aorta01 et *aorta02* sont utilisées pour l'apprentissage. Le jeu de validation est composé de patch provenant d'*aorta03*. Pour l'apprentissage, 10 patchs de taille $16 \times 16 \times 16$ sont découpées aléatoirement pour chaque image temporelle. Pour 9 de ces 10 patchs, au moins 20% du patch doit se trouver dans le domaine fluide, le dernier n'ayant aucune contrainte. Pour l'augmentation des données, 9 rotations sont calculées pour chaque patch (3 rotations sur les 3 axes). Comme les simulations CFD comportent 71 images temporelles, cela résulte en un ensemble d'apprentissage constitué de 14 200 patchs. Pour le jeu de validation, un seul patch, avec au moins 20% dans le domaine fluide, est extrait d'*aorta03* pour chaque image temporelle et est augmenté avec les 9 rotations.

Comme dans [3], l'évaluation des performances est effectuée sur l'image complète de *aorta03*. L'image entière est découpée en patchs de $16 \times 16 \times 16$, avec un pas de 12, pour éviter les effets de bord, et est reconstruite à l'aide de patchs super-résolus. Bien que l'ensemble de validation soit sélectionné dans la même image, il est important de noter qu'un seul patch par période de temps est partagé avec l'ensemble de test (ce qui représente environ 1% de l'image totale).

3.2 Métriques d'évaluation

Nous avons choisi quatre métriques utilisées dans la littérature [3, 9] pour évaluer les SR par rapport à une référence. Chaque métrique est calculée pour chaque période de temps, afin d'évaluer les performances tout au long d'un cycle cardiaque, et uniquement dans le domaine fluide, afin de n'évaluer que la super-résolution du fluide.

Pour les métriques suivantes, \mathbf{V} est le champ de vitesse de référence, \mathbf{V}' est celui prédit par le réseau et i est l'indice du voxel. La métrique \mathcal{E}_{rel} correspond à l'erreur vectorielle relative, ϵ est fixé à $1e^{-6}$ pour éviter la division par zéro.

$$\mathcal{E}_{\text{rel}} = \frac{1}{N} \sum_{i=1}^N \frac{\|V_{i,t} - V'_{i,t}\|}{\|V_{i,t}\| + \epsilon} \quad (1)$$

Le *Peak Velocity-to-Noise Ratio* (PVNR) permet d'évaluer l'erreur vectorielle relative, mais avec une autre normalisation.

$$\text{PVNR} = 20 \log_{10} \frac{1}{\text{nRMS}_{\text{vel}}} \text{dB} \quad (2)$$

$$\text{nRMS}_{\text{vel}} = \frac{1}{\max_{i,t} \|V_{i,t}\|} \sqrt{\frac{1}{N} \sum_{i=0}^N \|V_{i,t} - V'_{i,t}\|^2} \quad (3)$$

L'erreur quadratique moyenne normalisée de la vitesse (nRMS_{sp}) évalue la quantité d'erreur de vitesse du flux sanguin dans l'image SR.

$$\text{nRMS}_{\text{sp}} = \frac{1}{\max_{i,t} \|V_{i,t}\|} \sqrt{\frac{1}{N} \sum_{i=1}^N (\|V_{i,t}\| - \|V'_{i,t}\|)^2} \quad (4)$$

\mathcal{E}_{dir} évalue l'erreur de direction du vecteur par rapport à la référence.

$$\mathcal{E}_{\text{dir}} = \frac{1}{N} \sum_{i=0}^N \left(1 - \frac{\langle V_{i,t}, V'_{i,t} \rangle}{\|V_{i,t}\| \|V'_{i,t}\| + \epsilon} \right) \quad (5)$$

3.3 Méthodes évaluées

Trois méthodes de SR de l'état de l'art utilisant un CNN [3, 9, 11] sont évaluées. **4DFlowNet** [3] est un CNN utilisant une architecture résiduelle. Elle dispose de deux entrées, une dédiée aux images de magnitude et de norme de la vitesse, et la seconde pour les images de phase. Il produit trois images SR à partir de trois chemins différents, chacune correspondant à une composante de vitesse. **SRFlow** [9] est un CNN résiduel pour la SR utilisant le mécanisme de CA et l'*efficient wider activation* [10]. De plus, nous proposons **RCANFlow** combinant RCAN [11] et 4DFlowNet. Il utilise la même architecture que 4DFlowNet avec plusieurs chemins, mais l'extracteur de caractéristiques utilisé est celui de RCAN.

3.4 Implementation

Toutes les méthodes d'apprentissage profond ont été implémentées en PyTorch. Le GPU utilisé est un NVIDIA GeForce RTX 2080 Ti 12 GB avec CUDA 12.1. Le pré-traitement des données est fait avec le script de [3].

4DFlowNet est implémenté avec la même configuration que l'article original, il utilise 8 blocs LR et 4 blocs HR, ce qui correspond à 3456004 paramètres. **SRFlow**, est entraîné avec 4 blocs résiduels et 4 groupes résiduels. Cette configuration a été choisie parce qu'elle a 3697731 paramètres, ce qui est très proche de 4DFlowNet. **RCANFlow** utilise 3 blocs et groupes résiduels. Nous avons choisi d'avoir moins de groupes et de blocs que SRFlow pour obtenir une comparaison équitable entre les deux architectures car RCANFlow a naturellement plus de paramètres en raison des branches multiples. Dans cette configuration, RCANFlow a 4129804 paramètres.

Pour l'apprentissage, le taux d'apprentissage est de 10^{-4} , la taille du *mini-batch* est de 32. L'entraînement dure 300 époques et le modèle sauvegardé est celui qui présente le meilleur résultat sur \mathcal{E}_{rel} sur la validation. Chaque réseau est entraîné avec sa fonction de coût originale : \mathcal{L}_2 pour 4DFlowNet et $\mathcal{L}_1 + mp\mathcal{L}_1$ pour SRFlow. Nous avons aussi entraîné les différents modèles avec \mathcal{L}_1 pour avoir une comparaison entre les architectures et non la contribution de la fonction de coût. Ici, \mathcal{L}_1 et \mathcal{L}_2 désignent respectivement l'erreur absolue moyenne (MAE) et l'erreur quadratique moyenne (MSE).

La fonction de coût $mp\mathcal{L}_1$ utilisée dans [9] prend en compte la direction des vecteurs prédits dans la pénalité. RCANFlow a été entraîné avec et sans *channel attention* (*w/o CA*) pour comprendre la contribution de ce mécanisme.

4 Résultats

La Table 1 présente les résultats pour chaque approche à l'image du pic systolique (10°), où la vitesse est la plus haute et à une image arbitrairement choisie de la phase diastolique (32°). Les erreurs sont plus élevées pendant la phase diastolique en raison d'un faible rapport vitesse sur bruit. En outre, les schémas de vitesse sont plus complexes en raison de la fin de l'impulsion sanguine. En revanche, pendant la phase systolique, la vitesse est plus élevée par rapport au VENC. Le flux sanguin ressemble plus à un écoulement laminaire et est donc plus simple à prédire.

Premièrement, nous comparons les différents réseaux en utilisant la fonction de coût \mathcal{L}_1 . RCANFlow est plus performant

que les autres réseaux en ce qui concerne les erreurs relatives et de direction. On observe que l'erreur relative pénalise particulièrement les petits vecteurs de vitesse. Comme 4DFlowNet présente une erreur plus importante sur cette mesure, le réseau semble avoir plus de mal à gérer les plus petites vitesses, comme celles au bord de la paroi aortique. SRFlow et RCANFlow utilisent tous les deux de longues connexions raccourci et la *channel attention*. En comparaison à 4DFlowNet, les deux réseaux démontrent une meilleure capacité à reconstruire les écoulements turbulents (phase diastolique). L'utilisation des raccourcis permet la préservation de caractéristiques de bas niveau à la fin du réseau. La *channel attention* permet la sélection des caractéristiques les plus efficaces pour la reconstruction, assurant ainsi l'acquisition d'informations cruciales pour les écoulements turbulents.

\mathcal{L}_1 produit systématiquement de meilleurs résultats par rapport à \mathcal{L}_2 , quelle que soit le réseau. Pour 4DFlowNet, des performances supérieures sont observées pour la majorité des métriques, tant dans la phase systolique que dans la phase diastolique. Le même résultat est observé pour RCANFlow. En effet, on observe un gain de 1.39% en erreur relative ainsi qu'une réduction de 1.45% de son écart-type dans la phase systolique. De plus, la diminution de $nRMS_{sp}$ avec \mathcal{L}_1 démontre la capacité de tous les réseaux à mieux quantifier la quantité de vitesse (malgré des résultats proches dans la phase diastolique de 4DFlowNet). \mathcal{L}_1 permet à ces réseaux d'obtenir de meilleures solutions qu'avec \mathcal{L}_2 , ce qui est conforme aux travaux de SR de ces dernières années. De plus, \mathcal{L}_1 permet aux réseaux de prendre des décisions plus précises, en particulier sur les bords comme le montre la Figure 1. RCANFlow entraîné avec \mathcal{L}_1 montre une erreur plus petite au bord de l'aorte sur ce patch que celui entraîné avec \mathcal{L}_2 .

L'ajout de la fonction de coût $mp\mathcal{L}_1$ semble faciliter la convergence de SRFlow vers la solution correcte. En fait, $mp\mathcal{L}_1$ ajoute une pénalité de direction à la perte classique \mathcal{L}_1 qui aide le réseau à prédire les vecteurs de vitesse qui vont dans la bonne direction, comme le montre la métrique \mathcal{E}_{dir} .

Les résultats démontrent une différence statistiquement significative (valeur $p < 0,05$) entre les méthodes SRFlow ($\mathcal{L}_1 + mp\mathcal{L}_1$) et RCANFlow (\mathcal{L}_1), pour la métrique \mathcal{E}_{rel} , à la fois pour le pic systolique et l'image de la phase diastolique.

De plus, RCANFlow a été entraîné sans la CA pour tester son importance sur notre tâche. Les métriques montrent que la CA améliore les performances sur la plus part des métriques sauf \mathcal{E}_{dir} au pic systolique qui présente des valeurs presque identiques. CA est conçue pour permettre au réseau de choisir les caractéristiques les plus importantes pour la tâche.

5 Discussion et Conclusion

Ce papier fournit une première comparaison entre les différentes approches de l'état de l'art. Nous avons utilisé la méthode d'apprentissage et les données de 4DFlowNet [3] pour normaliser le processus d'apprentissage et d'évaluation à travers plusieurs architecture de réseaux de neurones, fonctions de coût ou le mécanisme de *channel attention*. Une analyse a été menée sur l'apport des différentes fonctions de coût dans la phase d'entraînement, indiquant la supériorité de la fonction de coût \mathcal{L}_1 par rapport à la \mathcal{L}_2 . L'incorporation de la fonction de coût complémentaire $mp\mathcal{L}_1$ dans le processus d'apprentissage

TABLE 1 : Résultats des différentes approches au pic systolique et un temps à la phase diastolique. Les **meilleurs** et **seconds meilleurs** résultats sont respectivement en gras et soulignés. L’astérisque (*) indique la version originale de l’algorithme publié.

Réseau	Pic systolique				Diastolique			
	\mathcal{E}_{rel} ↓	PVNR ↑	nRMS _{sp} ↓ (10^{-3})	\mathcal{E}_{dir} ↓ (10^{-3})	\mathcal{E}_{rel} ↓	PVNR ↑	nRMS _{sp} ↓ (10^{-3})	\mathcal{E}_{dir} ↓ (10^{-3})
4DFlowNet (\mathcal{L}_2)*	9.39±15.36	26.67	39.5	6.3±61.9	23.17±21.99	24.32	49.5	34.7±148.2
4DFlowNet (\mathcal{L}_1)	9.39±16.32	<u>26.75</u>	<u>38.7</u>	6.0±61.5	22.52±21.82	24.33	49.7	32.3±142.1
SRFlow (\mathcal{L}_1)	9.02±16.10	26.58	41.2	5.6±61.0	20.26±21.20	25.94	41.1	28.2±134.3
SRFlow ($\mathcal{L}_1 + mp\mathcal{L}_1$)*	7.31±14.25	27.34	36.7	5.2±58.1	19.59±20.31	<u>25.89</u>	<u>41.8</u>	25.5±123.5
RCANFlow (\mathcal{L}_2)	9.55±16.58	26.33	42.0	5.9±60.7	20.60±20.88	25.20	45.9	26.4±125.4
RCANFlow (\mathcal{L}_1)	<u>8.16±15.13</u>	26.55	40.5	5.6±58.0	<u>20.16±21.46</u>	25.63	43.0	27.5±130.8
RCANFlow (w/o CA)	8.67±15.79	26.18	43.0	5.4±57.6	20.85±22.12	24.44	49.6	31.0±142.1

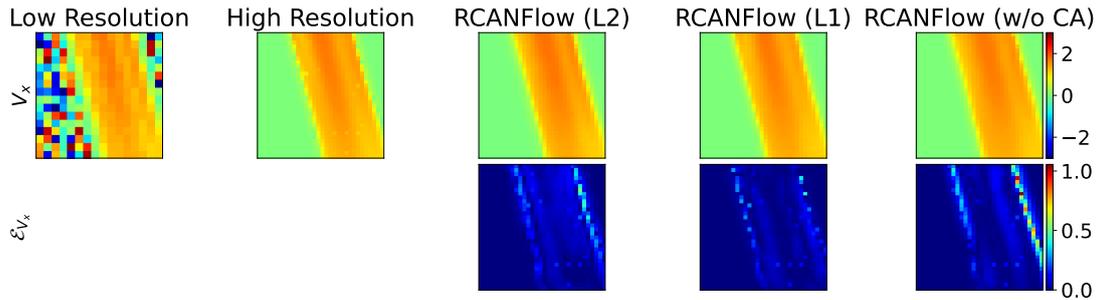


FIGURE 1 : Comparaison de différentes configurations de RCAN sur un patch $16 \times 16 \times 16$ au début de la crosse aortique au pic systolique pour V_x (haut); erreur absolue des prédictions (bas). La sortie haute résolution est de tailles $32 \times 32 \times 32$.

a permis de minimiser l’erreur et d’améliorer la reconstruction. Toutefois, une analyse approfondie de l’utilisation de la perte $mp\mathcal{L}_1$ pour la super-résolution IRM de flux 4D sur d’autres architectures devrait être menée à l’avenir afin de comprendre ses avantages.

L’usage de CA a démontré une amélioration des performances pour la tâche de super-résolution pour les réseaux résiduels aux temps évalués. D’autres analyses devraient être menées à l’avenir pour tester l’influence de ce type d’attention sur différentes architectures de réseau de neurones. Actuellement, l’étude ne compare que des méthodes basées sur les CNN, il serait nécessaire d’ajouter des méthodes comme les PINNs [8] ou non basées sur des réseaux de neurones. Bien qu’un patch par temps soit partagé entre les ensembles de données de validation et de test, nous considérons que cette analyse est pertinente compte tenu de la proportion de chevauchement (1 %). Cependant, le jeu de données, et donc l’évaluation, peuvent être améliorés en incorporant davantage de données CFD. De plus, de nouvelles métriques basées sur les biomarqueurs hémodynamiques et des données réelles permettraient d’obtenir une étude plus proche des besoins cliniques.

Références

- [1] C. DONG *et al.* : Learning a Deep Convolutional Network for Image Super-Resolution. In *Computer Vision – ECCV*, volume 8692, pages 184–199. 2014.
- [2] L. ERICSSON *et al.* : Generalized super-resolution 4D flow MRI - using ensemble learning to extend across the cardiovascular system. *IEEE Journal of Biomedical and Health Informatics*, pages 1–12, 2024.
- [3] E. FERDIAN *et al.* : 4DFlowNet : Super-Resolution 4D Flow MRI Using Deep Learning and Computational Fluid Dynamics. *Frontiers in Physics*, 8:138, 2020.
- [4] HáKon GUDBJARTSSON et Samuel PATZ : The rician distribution of noisy mri data. 34(6):910–914, 1995.
- [5] C. LEDIG *et al.* : Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In *IEEE Conference on CVPR*, pages 105–114. IEEE, 2017.
- [6] S. LEVILLY *et al.* : Towards quantitative evaluation of wall shear stress from 4D flow imaging. *Magnetic Resonance Imaging*, 74:232–243.
- [7] M. MARKL *et al.* : 4D flow MRI. *Journal of Magnetic Resonance Imaging*, 36(5):1015–1036, 2012.
- [8] M. RAISSI *et al.* : Physics-informed neural networks : A deep learning framework for solving forward and inverse problems involving narticlear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.
- [9] S. SHIT *et al.* : SRflow : Deep learning based super-resolution of 4D-flow MRI data. *Frontiers in Artificial Intelligence*, 5:928181, 2022.
- [10] Jiahui YU *et al.* : Wide activation for efficient and accurate image super-resolution. *CoRR*, abs/1808.08718, 2018.
- [11] Y. ZHANG *et al.* : Image Super-Resolution Using Very Deep Residual Channel Attention Networks. In *Computer Vision – ECCV*, volume 11211, pages 294–310. 2018.