

Approches Modernes pour la Super-Résolution d'Images Satellites

SE205 - Initiation à la recherche - Parcours Ingénieur-Docteur

Shakty ROBERT

Département Télécommunications

ENSEIRB-MATMECA, Bordeaux-INP

Talence, France

shakty.robert@bordeaux-inp.fr

Abstract—La super-résolution d'images satellites est un enjeu crucial pour de nombreuses applications en télédétection comme la cartographie, la surveillance environnementale ou la gestion des risques. Les méthodes traditionnelles basées sur l'interpolation ou des modèles statistiques (filtrage de Wiener, estimation MAP) sont très limitées lorsqu'il s'agit de reconstruire des détails fins et/ou réalistes. L'essor du deep learning a permis d'importants progrès, notamment avec les réseaux de neurones à convolutions et les modèles génératifs comme les GAN. En effet, ces derniers sont certes capables de produire des images qui sont visuellement acceptables mais avec un apprentissage instable. Plus récemment, les modèles de diffusion ont émergé comme une alternative à ces méthodes, ils offrent notamment une meilleure stabilité et une capacité à modéliser des distributions complexes. Dans ce rapport, nous présentons l'état de l'art des approches de super-résolution, en insistant sur les modèles de diffusion, et analysons en détail EDiffSR, un modèle efficace spécifiquement conçu pour les images satellites. EDiffSR intègre des modules légers et exploite mieux les informations de l'image basse résolution, tout en atteignant une qualité perceptuelle élevée avec un coût computationnel réduit.

Index Terms—imagerie satellite, super-résolution, deep learning, modèle de diffusion

I. INTRODUCTION

L'amélioration de la résolution spatiale des images satellites est un enjeu crucial pour de nombreuses applications dans les domaines de l'observation de la Terre, de la cartographie, de l'agriculture, ou encore de la surveillance environnementale. La résolution spatiale désigne la densité de pixels par unité de surface, ainsi elle indique directement la quantité et le détail des informations extraites d'une image. Toutefois, des contraintes matérielles (coût élevé des capteurs haute résolution, limitations physiques embarquées dans les satellites) et logistiques (volume de données, bande passante, temps de traitement) limitent la fréquence et la qualité des images à haute résolution (HR).

La super-résolution (SR) consiste à reconstituer une image HR à partir d'une image basse résolution (LR). Le problème est donc que pour une certaine image LR, il existe potentiellement une infinité d'interprétations plausibles à l'œil nu en HR. De plus, les approches classiques comme l'interpolation, le

filtrage et l'estimation mathématique sont peu efficaces et très limitées car incapables de reconstruire des textures assez fines ou encore de s'adapter à du bruits ou des artefacts.

L'émergence du deep learning peut solutionner ces limitations. Les réseaux de neurones convolutionnels (CNN) ont permis de modéliser directement la relation entre les images LR et HR. Par la suite, les modèles génératifs adversariaux (GAN) ont introduit la possibilité de produire des images réalistes. Toutefois, ces méthodes restent très instable à l'entraînement et peuvent générer des artefacts.

C'est pourquoi les modèles de diffusion probabilistes (DPM) se distinguent par leur stabilité d'entraînement et leur capacité à générer des données de qualité en modélisant le processus de génération inverse. Ce rapport propose d'explorer l'évolution des approches de super-résolution, depuis les méthodes traditionnelles jusqu'aux modèles de diffusion, avec un focus particulier sur *EDiffSR* [1], un modèle de diffusion efficace et léger proposé il y a moins d'un an pour les images satellites. Nous analyserons sa structure, ses innovations (modules CPEM et EANet), ainsi que ses performances expérimentales.

II. FONDEMENTS DE LA SUPER-RÉSOLUTION

La super-résolution d'image (SR) [6] en traitement d'image c'est le fait de générer une image haute résolution (HR) à partir d'une ou plusieurs images basse résolution (LR). On cherche en fait à reconstruire une image avec des détails qui seraient cohérents à la réalité terrain.

La formulation mathématique du problème de super-résolution peut être exprimée ainsi :

$$\mathbf{y} = \mathcal{D}(\mathbf{x}) + \mathbf{n} \quad (1)$$

où \mathbf{y} est l'image LR observée, \mathbf{x} l'image HR à reconstruire, $\mathcal{D}(\cdot)$ une fonction de dégradation (flou, sous-échantillonnage, compression, etc..), et \mathbf{n} du bruit. L'objectif est donc de retrouver \mathbf{x} à partir de \mathbf{y} , en inversant l'effet de \mathcal{D} .

Plusieurs approches sont possibles, allant des méthodes simples comme l'interpolation (bicubique, spline, lanczos) à des modèles statistiques plus complexes. Ces méthodes sont "suffisantes" sur des images qui possèdent des propriétés simples (continuité, lissage), mais n'arrive pas à reconstruire des structures complexes ou du contenu vraiment réaliste.

Avec du deep learning, il est désormais possible d'apprendre à faire de la super-résolution [4]. En effet, les modèles de type CNN apprennent à approximer l'inverse de la fonction de dégradation à partir d'un jeu de données d'images [LR, HR]. Cela permet de capturer des informations plus précises et de s'adapter aux variations qui peuvent exister sur données réelles.

Pour évaluer les performances, il existe différentes unités de mesure:

- PSNR (Peak Signal-to-Noise Ratio) : mesure la fidélité pixel à pixel.
- SSIM (Structural Similarity Index Measure) : mesure la similarité structurelle entre l'image reconstruite et la réalité.
- FID (Fréchet Inception Distance) et LPIPS (Learned Perceptual Image Patch Similarity) : mesurent la qualité perceptuelle et la distance dans l'espace des caractéristiques.
- NIQE et DISTS : mesurent l'aspect réaliste de l'image et la cohérence des textures.

Ces unités de mesure permettent de comparer objectivement les performances des différentes approches.

III. MÉTHODES TRADITIONNELLES VS. DEEP LEARNING

A. Approches traditionnelles

Les premières méthodes de super-résolution reposaient sur des techniques simples à mettre en œuvre mais limitées en performance. Parmi les plus utilisées on a:

- Interpolation : consiste à agrandir une image en estimant des valeurs entre chaque pixel de l'image originale, toutefois cela introduit un fort effet de flou et n'ajoute pas d'information réelle.
- Filtrage : méthode statistique qui vise à estimer une image à partir d'une image bruitée, en minimisant l'erreur quadratique moyenne.
- Modèles mathématiques : repose sur des modèles probabilistes pour reconstruire l'image la plus probable étant donné l'image observée.

Ces approches sont peu coûteuses en calcul mais aussi peu performantes pour restituer des détails fins. Ainsi, ces méthodes traditionnelles sont inadaptées à la complexité et à la diversité des textures des images satellites.

B. Transition vers le Deep Learning

Les réseaux de neurones convolutionnels (CNN) ont permis d'apprendre une transformation directe entre images LR et HR

à partir d'énormes bases de données [4]. Cependant, les CNN ont également plusieurs limitations :

- Sur-lissage : les pertes de type MSE conduisent à des images visuellement floues.
- Coût computationnel : Coût élevé pour des architectures assez profondes et complexes.
- Dépendance aux données: la qualité du modèle dépend logiquement de la base d'entraînement (résolution, type d'image, bruit, etc..).

Pour pallier à ces problèmes, les approches génératives comme les GANs ont été étudiés avec pour objectif de générer des images plus réalistes perceptuellement. Cependant, elles sont plus difficiles à entraîner et sensibles aux artefacts. Plus récemment, ce sont les modèles de diffusion qui intéressent la communauté scientifique puisqu'ils combinent qualité de génération, stabilité d'entraînement et flexibilité.

IV. APPROCHES GÉNÉRATIVES : GANs ET MODÈLES DE DIFFUSION

A. GANs pour la super-résolution

Les Generative Adversarial Networks (GANs) [5] [7] fonctionnent grâce à un mécanisme d'entraînement antagoniste entre deux réseaux : un générateur G et un discriminateur D . Le générateur doit apprendre à produire des images réalistes à partir d'une image LR alors que le discriminateur doit apprendre à distinguer les images générées des images HR réelles. La figure 1 présente visuellement ce processus.

L'objectif d'entraînement des GANs est formulé comme un jeu à somme nulle :

$$\min_G \max_D \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}} [\log(1 - D(G(\mathbf{z})))] \quad (2)$$

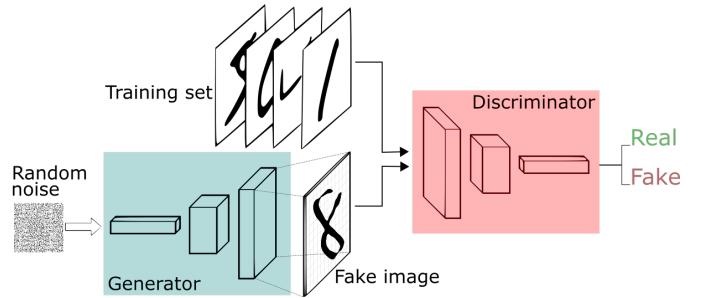


Fig. 1. Architecture typique d'un GAN. Le générateur transforme du bruit en image puis le discriminateur juge si elle est vraie ou fausse.

Dans le cadre de la super-résolution, l'entrée du générateur est une image LR et la sortie est censée correspondre à une version HR. Ensuite, c'est au discriminateur de juger si cette sortie est une image générée ou non en la comparant à de vraies images HR. Le but final étant que le générateur trompe le discriminateur.

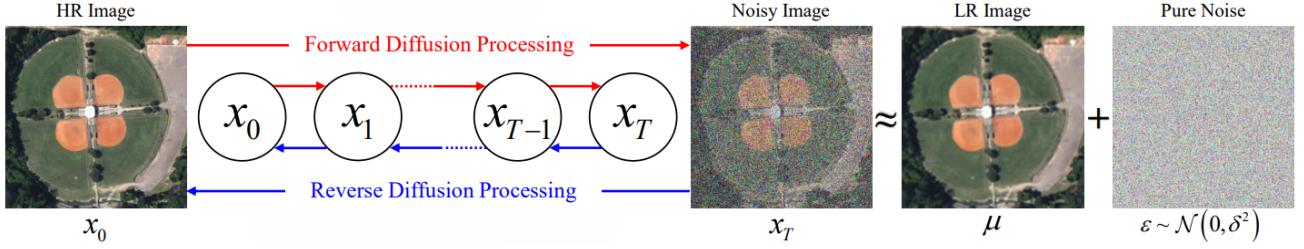


Fig. 2. Illustration du processus de diffusion : dégradation progressive (forward) puis génération inverse (reverse).

Les GANs ont deux principales limites non négligeables :

- L'apprentissage est souvent instable et peut générer des artefacts et/ou du bruit.
- Le réglage des fonctions de perte et des hyperparamètres est difficile.

B. Modèles de diffusion

Les modèles de diffusion [2] [3] (Diffusion Probabilistic Models, DPM) sont une alternative aux GANs qui offrent de meilleurs résultats en SR. Le principe repose sur deux phases :

- une diffusion directe (forward process), qui bruite progressivement une image en y ajoutant du bruit à chaque étape.
- une diffusion inverse (reverse process), qui apprend à inverser ce bruit et à générer une image réaliste à partir d'un bruit blanc.

Dans le cadre des DDPM (Denoising Diffusion Probabilistic Models), la diffusion directe est définie comme une chaîne de Markov discrète :

$$q(\mathbf{x}_t | \mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t} \mathbf{x}_{t-1}, \beta_t \mathbf{I}) \quad (3)$$

où β_t est le taux de bruit à l'étape t .

La diffusion inverse est apprise par un réseau de neurones $\epsilon_\theta(\mathbf{x}_t, t)$ qui prédit le bruit ajouté à chaque étape. L'apprentissage consiste à minimiser :

$$\mathcal{L}_{\text{simple}} = \mathbb{E}_{\mathbf{x}_0, \epsilon, t} \left[\left\| \epsilon - \epsilon_\theta(\sqrt{\alpha_t} \mathbf{x}_0 + \sqrt{1 - \alpha_t} \epsilon, t) \right\|^2 \right] \quad (4)$$

où ϵ est un bruit gaussien, et $\bar{\alpha}_t = \prod_{s=1}^t (1 - \beta_s)$.

Ces modèles possèdent des avantages non négligeables :

- Stabilité d'apprentissage : pas d'interactions adversariales comme dans les GANs.
- Qualité générative : capacité à générer des distributions complexes et des images visuellement cohérentes.
- Flexibilité : adaptables à différents types de données ou scénarios.

Toutefois, ces modèles rencontrent également des limites :

- Temps de génération élevé : processus de diffusion lent, limitant l'utilisation en temps réel.
- Complexité de calcul : Coût élevé en termes de calcul.

V. EDDIFFSR : MODÈLE DE DIFFUSION EFFICACE POUR LA SUPER-RÉSOLUTION SATELLITE

A. Motivations

Les modèles de diffusion sont certes puissants mais ont souvent deux limitations majeures pour la SR d'images satellites : (1) l'utilisation d'architectures lourdes comme le UNet, coûteuses en calcul ; (2) une exploitation sous-optimale de l'information contenue dans les images basse résolution, souvent réduites à une simple interpolation bicubique.

Pour pallier ces problèmes, *EDiffSR* [1]. (Efficient Diffusion Probabilistic Model for Remote Sensing Image Super-Resolution) propose une architecture plus légère et mieux adaptée à la SR, tout en conservant les avantages des DPM. Ce nouveau modèle apporte deux innovations principales : l'Efficient Activation Network (EANet) et le Conditional Prior Enhancement Module (CPEM).

B. Structure générale

La Figure 3 présente la structure globale d'EDiffSR qui est composée de trois blocs :

- un module de conditionnement (CPEM) qui extrait les informations "pertinentes" à partir de l'image LR,
- un débruiteur (EANet) chargé de prédire le bruit injecté à chaque étape de diffusion,
- un bloc d'optimisation basé sur l'apprentissage par maximum de vraisemblance, via des équations différentielles stochastiques (SDE).

C. CPEM : Conditionnement enrichi

Contrairement à d'autres modèles qui se contentent d'une interpolation de l'image LR, EDiffSR introduit le CPEM (Figure 4) pour extraire des informations plus riches. Ce dernier utilise des blocs d'attention résiduelle et de fusion des caractéristiques, c'est aussi lui qui agrandit l'image LR au format de l'image HR grâce à un sur-échantillonage.

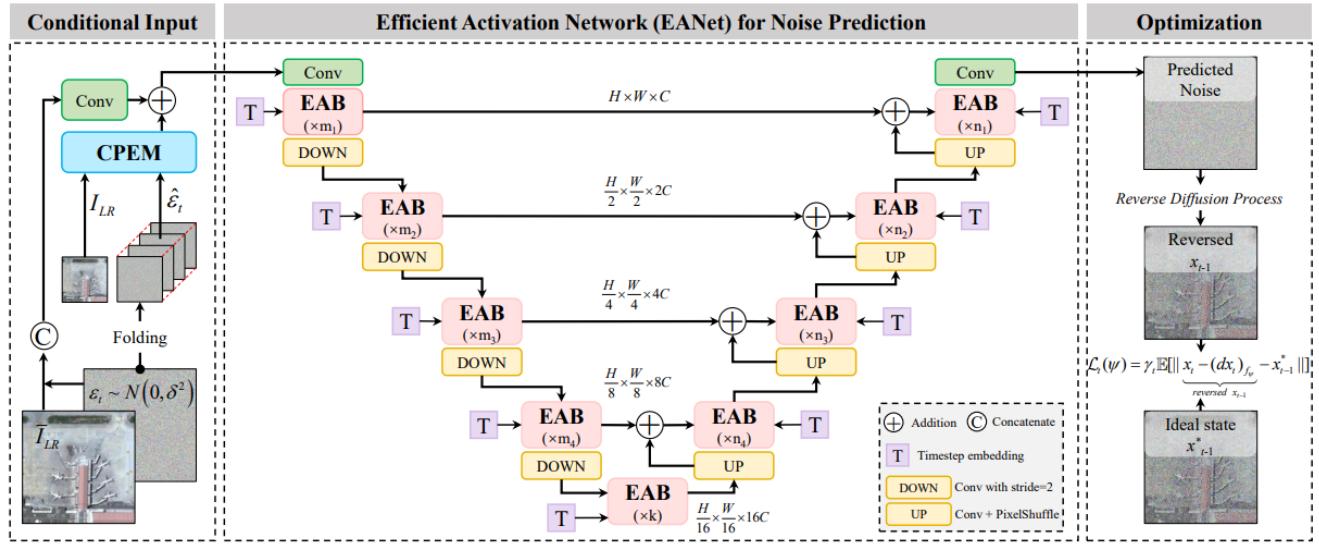


Fig. 3. Architecture générale d'EDiffSR. Le modèle se compose d'un module de conditionnement (CPEM), d'un débruiteur efficace (EANet), et d'un bloc d'optimisation.

Le CPEM produit à chaque étape temporelle t un vecteur de conditionnement et l'injecte dans le processus de génération, ce qui améliore la précision des détails reconstruits dans l'image HR.

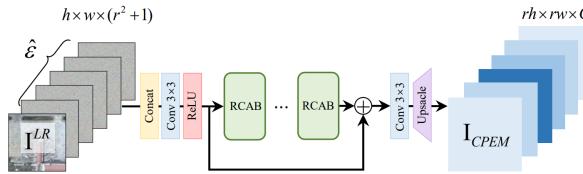


Fig. 4. Structure interne du module CPEM. Il exploite des blocs d'attention pour enrichir le conditionnement à partir de l'image LR.

D. EANet : Réseau efficace pour la prédiction du bruit

L'architecture "classique" des modèles de diffusion utilise un UNet profond et coûteux. Ici, EDiffSR propose une alternative : l'Efficient Activation Network (EANet), basé sur le Efficient Activation Block (EAB) (Figure 5).

Chaque EAB combine :

- un mécanisme de canal attention simplifié,
- une opération de passerelle efficace,
- une hiérarchie de convolutions avec réduction et expansion des canaux.

Cette structure permet d'atteindre une performance de prédiction comparable à celle du UNet tout en réduisant considérablement le nombre de paramètres (voir Tableau I).

E. Bloc d'optimisation : Processus de diffusion basé sur SDE

EDiffSR se base sur un principe de diffusion dit continu, en modélisant les transitions avec des équations différentielles

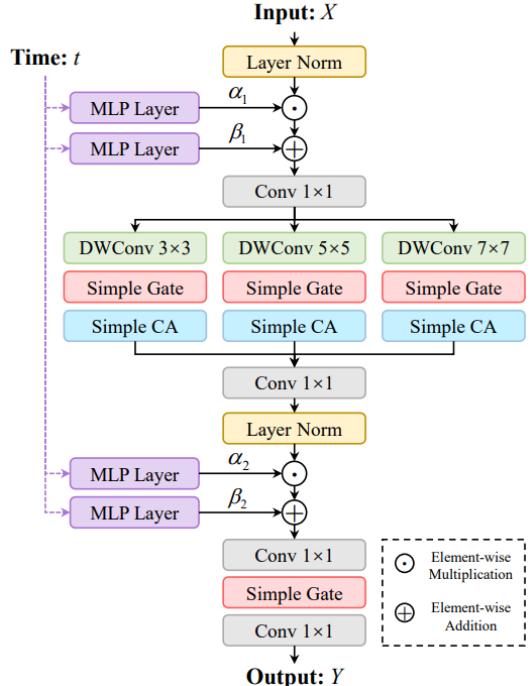


Fig. 5. Architecture des blocs EAB pour une efficacité maximale.

stochastiques (SDE) :

$$dx_t = \lambda_t(u - x_t)dt + \phi_t d\omega \quad (5)$$

où λ_t est le taux de retour à la moyenne, ϕ_t le terme de diffusion, et ω un processus de Wiener. Ce cadre permet une génération plus stable et efficace. La Figure 2 illustre les étapes de diffusion directe et inverse, du bruit blanc jusqu'à l'image HR reconstruite.

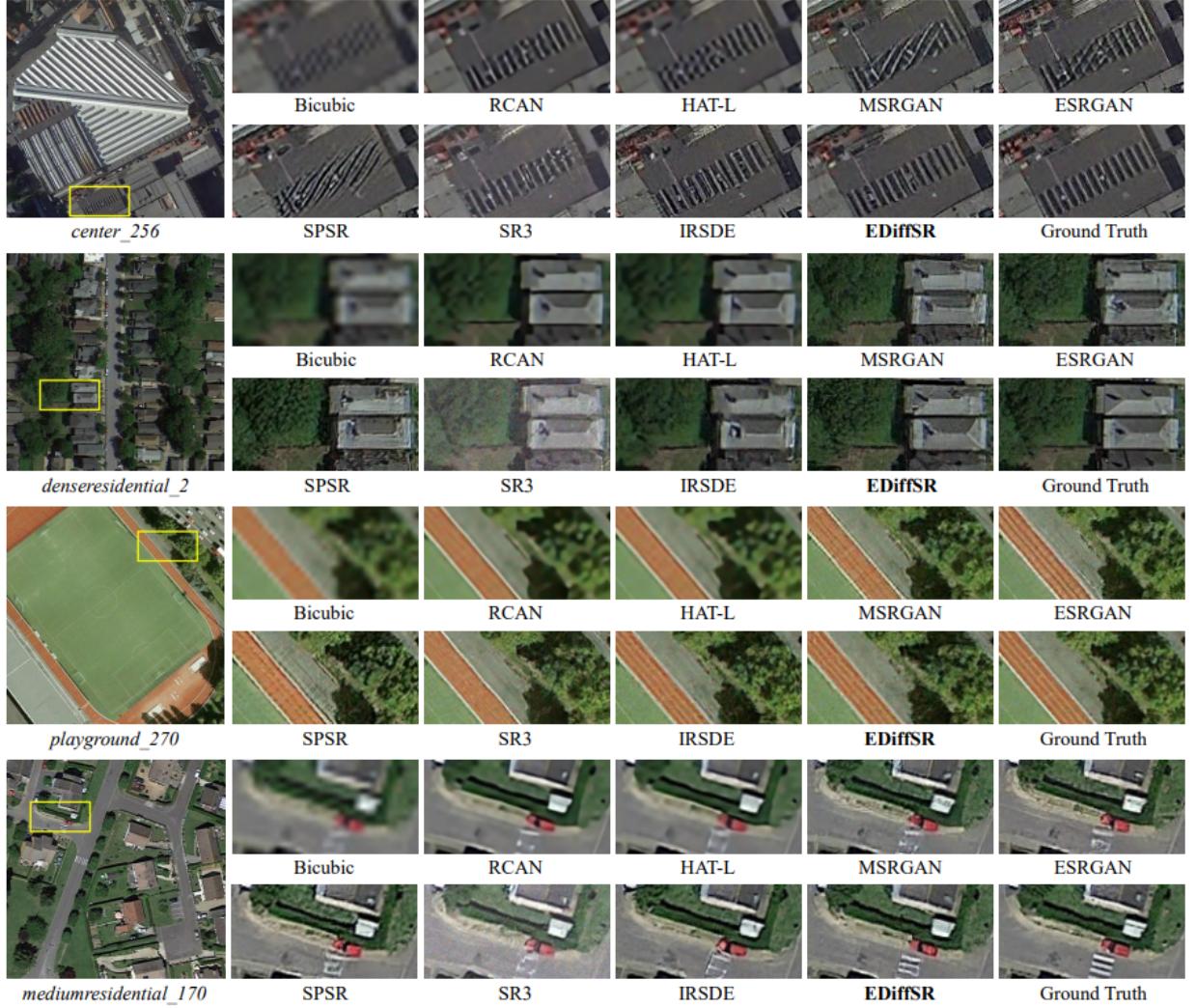


Fig. 6. Comparaison visuelle avec les méthodes de super-résolution sur le jeu de test AID. Les résultats montrent que EDiffSR surclasse les autres approches en termes de restitution des détails haute fréquence, tout en produisant des images visuellement cohérentes et plus naturelles.

TABLE I

ANALYSE DE L'EFFICACITÉ DES MODÈLES COMPARÉE AUX MÉTHODES DE SUPER-RÉSOLUTION. LA MEILLEURE PERFORMANCE EST EN GRAS.

Méthodes	Param. (M)	Temps (s)	FID ↓	PSNR (dB) ↑
<i>Méthodes CNN</i>				
EDSR	43.09	0.93	62.50	31.64
RCAN	15.59	0.29	59.99	31.84
HAT-L	40.32	0.76	58.47	31.89
<i>Méthodes GAN</i>				
MSRGAN	1.52	0.19	34.88	29.02
ESRGAN	16.70	0.22	33.16	28.49
SPSR	24.79	0.60	35.61	27.72
<i>Méthodes de diffusion</i>				
SR3	92.56	137.61	43.61	26.68
IRSDE	137.15	27.90	32.42	27.46
EDiffSR (Ours)	26.79	19.26	30.83	27.75

F. Résultats expérimentaux

EDiffSR a été évalué sur quatre jeux de données :

- AID : Jeu de données pour la classification de scènes aériennes avec 30 catégories différentes (villes, ports, forêts, etc.).
- DOTA : Jeu de données destiné à la détection d'objets dans des images aériennes HR.
- DIOR : Jeu de données pour la reconnaissance d'objets dans des images satellites (véhicules, infrastructures,...).

Comme illustré dans le Tableau II, il surpasse les modèles basés sur CNN, GAN et les autres DPM sur plusieurs métriques : FID, DISTs, NIQE, PSNR.

Une comparaison visuelle avec un ensemble de modèles a été réalisé. Comme le montre la Figure 6, EDiffSR produit systématiquement des résultats réalistes surpassant les approches existantes. Par exemple pour l'image *center_256* du

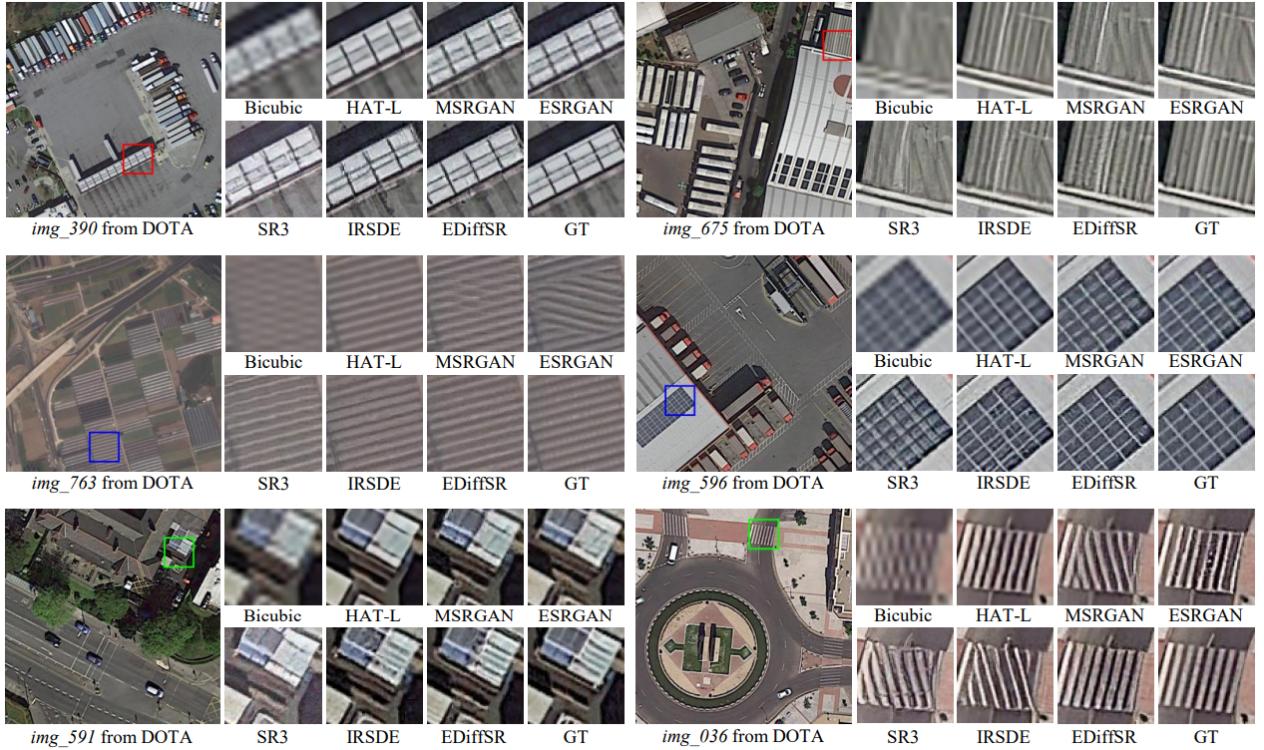


Fig. 7. Comparaisons visuelles avec les méthodes de super-résolution sur le jeu de test DOTA.

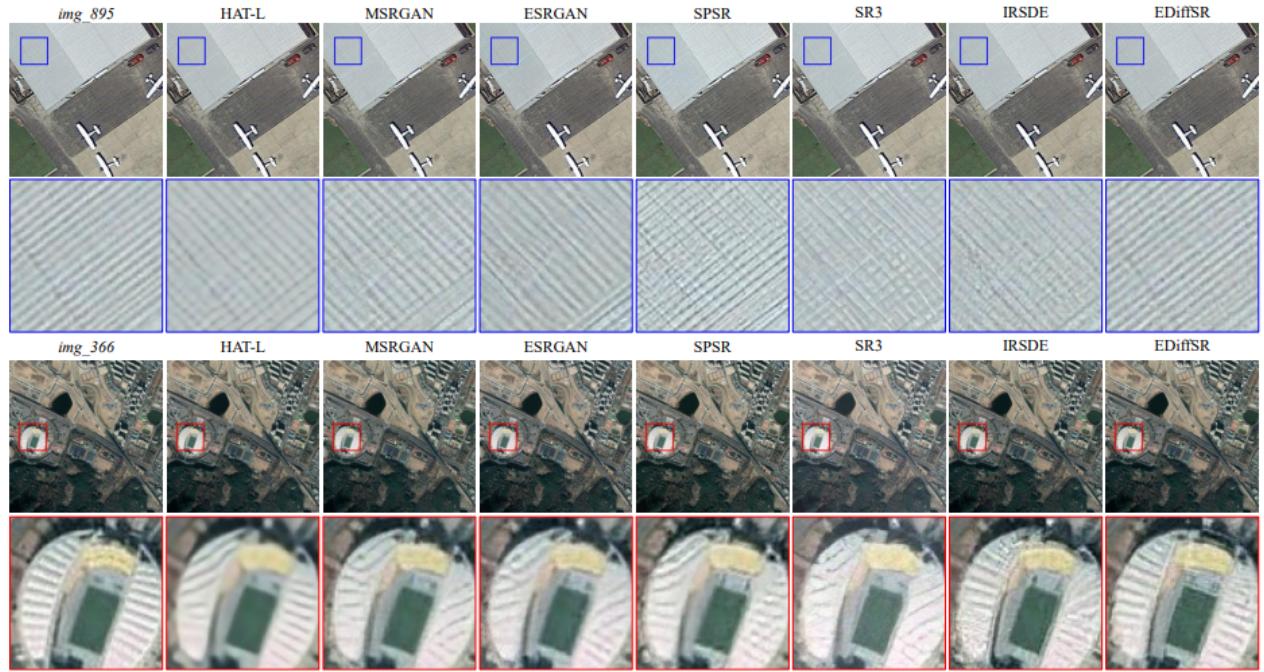


Fig. 8. Comparaisons visuelles avec les méthodes de super-résolution sur le jeu de test DIOR.

jeu de données AID, les modèles RCAN et HAT-L génèrent des sorties floues, révélant la capacité limitée des modèles orientés PSNR à reconstruire des détails riches. À l'inverse, les modèles GAN parviennent à restituer des structures

comme les bords, mais introduisent souvent des artefacts incohérents avec l'image de référence. EDiffSR fournit quant à lui des résultats plus naturels et réalistes, illustrant sa capacité à générer des images visuellement convaincantes.

TABLE II
COMPARAISON QUANTITATIVE AVEC D'AUTRES MÉTHODES SUR LES JEUX DE DONNÉES AID, DOTA ET DIOR, SELON FID, LPIPS, DISTS ET NIQE.
LA MEILLEURE PERFORMANCE EST INDIQUÉE EN ROUGE, LA DEUXIÈME MEILLEURE EN BLEU.

Dataset	Metrics	Bicubic	EDSR	RCAN	HAT-L	MSRGAN	ESRGAN	SPSR	SR3	IRSDE	EDiffSR
AID	FID ↓	126.53	90.40	93.39	90.86	56.36	52.55	56.40	65.51	51.08	49.45
	LPIPS ↓	0.4801	0.3068	0.3112	0.3078	0.1694	0.1695	0.1751	0.2534	0.2247	0.1887
	DISTS ↓	0.1512	0.0880	0.0900	0.0882	0.0590	0.0601	0.0632	0.0903	0.0579	0.0561
	NIQE ↓	21.24	19.79	19.88	20.15	17.46	15.51	17.90	15.16	14.80	14.22
DOTA	FID ↓	65.99	55.22	43.37	42.61	25.12	24.42	26.42	35.74	23.91	21.26
	LPIPS ↓	0.4416	0.2616	0.2629	0.2641	0.1649	0.1506	0.1605	0.2790	0.1919	0.1689
	DISTS ↓	0.1481	0.0831	0.0846	0.0832	0.0589	0.0582	0.0617	0.0933	0.0550	0.0556
	NIQE ↓	19.48	18.37	18.58	18.98	16.71	15.48	18.01	15.49	14.47	14.20
DIOR	FID ↓	57.42	41.87	43.20	41.94	23.16	22.51	24.02	29.60	22.28	21.79
	LPIPS ↓	0.4678	0.3020	0.3048	0.3062	0.1722	0.1863	0.1772	0.2836	0.2419	0.1898
	DISTS ↓	0.1497	0.0886	0.0893	0.0893	0.0605	0.0654	0.0652	0.0924	0.0622	0.0590
	NIQE ↓	20.24	19.16	19.30	19.56	17.61	15.13	18.64	15.55	15.19	15.16

La Figure 7 montre des résultats issus du jeu de test DOTA. Pour l'image *img 591*, seul EDiffSR parvient à restaurer des détails nets du bâtiment. Dans *img 036*, MSRGAN, SR3 et IRSDE présentent des distorsions importantes, éloignées de la vérité terrain. Bien que HAT-L et ESRGAN restituent correctement l'orientation des lignes de route, ESRGAN souffre d'une apparence artificielle due à une accentuation excessive. EDiffSR, au contraire, génère ces détails avec une perception bien plus naturelle. Cela confirme l'intérêt de CPEM dans la récupération d'informations a priori additionnelles, permettant une reconstruction plus réaliste.

Enfin, la Figure 8 présente des résultats issus du jeu DIOR. Sur l'image *img 895*, EDiffSR surpasse les autres modèles en restituant correctement les lignes du toit d'un bâtiment — une tâche difficile car elle implique des détails haute fréquence. Tous les autres modèles échouent ici, contrairement à EDiffSR qui, grâce au module CPEM, extrait davantage d'informations fréquentielles utiles dans le processus de diffusion. Sur l'image *img 366*, centrée sur une zone de stade, EDiffSR fournit une reconstruction nettement plus réaliste, là où les autres approches produisent du flou ou des distorsions notables.

VI. ANALYSE ET PERSPECTIVES

A. Avantages d'EDiffSR

Le modèle EDiffSR est innovant par rapport aux approches précédentes car il mélange qualité de reconstruction, stabilité d'entraînement et efficacité computationnelle. Le module CPEM permet de mettre en avant des détails structurels importants de l'image basse résolution. Par ailleurs, le réseau EANet réduit considérablement le nombre de paramètres et la consommation de mémoire tout en ayant des performances qui sont comparables à UNet. L'utilisation de la diffusion améliore également la stabilité de l'apprentissage en évitant les problèmes observés avec les GANs. Enfin, les résultats

expérimentaux sur des images de télédétection réelles confirment qu'EDiffSR est supérieur avec des reconstructions visuellement plus fidèles et sans artefacts.

B. Limites actuelles

Malgré de très bonnes performances, EDiffSR présente plusieurs limites qui peuvent freiner son adoption à grande échelle :

- Temps de génération élevé : comme pour la plupart des modèles de diffusion, le processus inverse repose sur un grand nombre d'étapes, ce qui rend l'inférence lente et coûteuse.
- Robustesse face aux dégradations réelles : même si le modèle a été testé sur des images simulées et quelques images réelles, il reste sensible au bruit, à la compression, ou à des conditions atmosphériques complexes.
- Complexité d'implémentation : la combinaison des blocs EANet, CPEM, SDE implique un certain degré de connaissance et d'expertise.

C. Perspectives de recherche

Plusieurs pistes d'amélioration peuvent être envisagées pour faire évoluer EDiffSR ou concevoir de nouveaux modèles encore plus efficaces :

- Combinaison de techniques : intégration des approches GANs et modèles de diffusion pour tirer parti des avantages des deux.
- Amélioration des algorithmes existants : optimisation en termes d'efficacité et de réduction du coût computationnel, notamment par des architectures plus légères.
- Nouveau modèle génératif : proposition d'un modèle innovant adapté aux contraintes de la super-résolution d'images satellites.

Ainsi, EDiffSR constitue une base solide pour la prochaine génération de modèles de super-résolution, tout en ouvrant la

voie à des systèmes plus légers, rapides et robustes dans des conditions réelles.

VII. CONCLUSION

L'amélioration de la résolution spatiale des images satellites est un défi technique et économique majeur. Afin de faire face aux limitations des méthodes traditionnelles, les méthodes basées sur le deep learning ont permis des avancées significatives, surtout grâce aux modèles génératifs comme les GANs et plus récemment les modèles de diffusion.

Dans ce rapport, nous avons présenté un état de l'art des différentes approches de super-résolution tout en donnant leurs avantages et leurs limites. Nous avons ensuite analysé en détail le modèle EDiffSR, qui propose une alternative efficace aux architectures classiques de diffusion. Grâce au module CPEM pour un conditionnement enrichi et à EANet pour la prédiction du bruit, EDiffSR représente un parfait compromis entre qualité, stabilité et efficacité.

Les résultats expérimentaux montrent que ce modèle dépasse les performances des approches CNN, GAN et DPM existantes tout en réduisant la complexité computationnelle. Toutefois, il subsiste encore certaines limites comme la lenteur du processus de génération ou la robustesse aux dégradations complexes.

Enfin, ce modèle ouvre de nombreuses perspectives pour le futur et EDiffSR marque ainsi une étape importante vers des solutions de super-résolution plus performantes, déployables à grande échelle dans des contextes opérationnels.

REFERENCES

- [1] Y. Xiao, Q. Yuan, K. Jiang, J. He, X. Jin, and L. Zhang, “EDiffSR: An Efficient Diffusion Probabilistic Model for Remote Sensing Image Super-Resolution,” *arXiv preprint arXiv:2310.19288*, 2024.
- [2] B. Ghojogh and A. Ghodsi, “Diffusion Models: Tutorial and Survey,” 2024. [Online]. Available: <https://hal.science/hal-04642649>
- [3] H. Cao, C. Tan, Z. Gao, Y. Xu, G. Chen, P.-A. Heng, and S. Z. Li, “A Survey on Generative Diffusion Models,” *arXiv preprint arXiv:2209.02646*, 2023.
- [4] Z. Wang, J. Chen, and S. C. H. Hoi, “Deep Learning for Image Super-resolution: A Survey,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 10, pp. 3365–3387, Oct. 2021.
- [5] C. Wang, Y. Zhang, Y. Qiao, C. Zhang, and X. Tang, “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks,” in *Proc. ECCV Workshops*, 2018.
- [6] Wikipédia, “Super-résolution,” Wikipédia, l’encyclopédie libre. [En ligne]. Disponible sur : <https://fr.wikipedia.org/wiki/Super-r%C3%A9solution> [Consulté le 16 avril 2025].
- [7] Wikipédia, “Réseaux antagonistes génératifs,” Wikipédia, l’encyclopédie libre. [En ligne]. Disponible sur : https://fr.wikipedia.org/wiki/R%C3%A9seaux_antagonistes_g%C3%A9n%C3%A9ratifs [Consulté le 16 avril 2025].