

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne démocratique et populaire

وزارة التعليم العالي والبحث العلمي

Ministère de l'enseignement supérieur et de la recherche scientifique

جامعة سعد دحلب البليدة

Université SAAD DAHLAB de BLIDA

كلية التكنولوجيا

Faculté de Technologie

قسم الإلكترونيك

Département d'Électronique



Mémoire de Master

Filière : Electronique

Spécialité : Electronique des systèmes embarqués

Présenté par

ADOUR HANENE AHLEM

&

BENAOUDA LOUBNA

Déconvolution d'images réelles par Le GAN

Proposé Par :

Mr NAMANE Abderrahmane.

Mr AOUINANE Bilal.

Année Universitaire 2023-2024.

Remerciement

Nous tenons tout d'abord à remercier Dieu le Tout-Puissant et Miséricordieux, qui nous a donné la force et la patience nécessaires pour accomplir ce travail.

Nous exprimons ensuite notre profonde gratitude à Mr NAMANE Abderrahmane, notre promoteur, pour son encadrement et ses précieux conseils tout au long de ce projet. Son expertise et sa confiance en notre travail ont été des moteurs essentiels pour mener à bien cette étude.

Nous remercions également Mr AOUINANE Bilel, notre co-promoteur, pour son soutien constant et son aide précieuse. Sa disponibilité et son engagement ont grandement contribué à la qualité de ce travail.

Nous sommes également très reconnaissants envers les membres de notre jury, Mr ZAIR et Mr MAAMOUN, pour l'honneur qu'ils nous font en participant à ce jury. Leur expertise et leurs suggestions constructives ont enrichi notre travail de manière significative.

Nous adressons un grand merci à nos familles, dont le soutien et les encouragements ont été indispensables pour mener à bien ce projet.

Enfin, nos remerciements vont à toutes les personnes qui nous ont aidés, de près ou de loin, dans l'élaboration de ce travail. Leur soutien et leurs conseils ont été précieux tout au long de cette aventure.

Loubna et Ahlem.

Dédicace

Je dédie ce mémoire à mes chers parents qui ont été toujours à mes côtés et m'ont toujours soutenu tout au long de ces longues années d'études. En signe de reconnaissance, qu'ils trouvent ici, l'expression de ma profonde gratitude pour tout ce qu'ils ont consenti d'efforts et de moyens pour me voir réussir dans mes études.

A toute ma famille

Et A toutes mes amies,

A tous les gens qui me connaissent et que je connais.

Et à tous ceux qui aiment le bon travail et ne reculent pas devant les obstacles de la vie.

Ahlem.

Dédicace

Je dédie ce travail

À mon cher père,

Tu m'as inculqué le vrai sens du travail dur et de la responsabilité. Mon succès a été fondé sur ton amour, ta générosité et ton soutien indéfectible. Ce travail rend hommage aux efforts que tu as déployés pour mon éducation. Que Dieu te bénisse d'une longue et heureuse vie.

À ma chère mère,

Les mots ne peuvent pleinement exprimer l'amour profond et la gratitude que j'ai pour toi. Ta tendresse et ton soutien constant ont été ma force. Tu as toujours été là pour m'encourager et me réconforter. Que Dieu te protège, te donne santé et bonheur, et te préserve de tout mal.

À mon frère et ma sœur,

Merci pour votre amour et votre soutien. Vous apportez joie et sens à ma vie. Ce travail reflète mon affection profonde et ma reconnaissance pour vous.

À mes meilleurs amis,

Votre amitié signifie tout pour moi. Merci d'avoir toujours été à mes côtés, d'avoir cru en moi et de m'avoir soutenu dans mes moments difficiles. Je vous dédie ce travail avec une gratitude sincère et vous souhaite une vie pleine de bonheur et de santé.

À mes chers amis et à mon binôme,

Merci pour votre soutien, vos encouragements et votre attention. Merci pour tout.

À toute ma famille,

Aucun mot ne peut vraiment exprimer le respect et la gratitude que j'ai pour votre soutien, Merci pour votre amour et vos encouragements.

À tous ceux qui nous ont aidés de près ou de loin,

Je vous adresse ma plus profonde gratitude.

Loubna.

الملخص

هدف هذا المشروع هو تصميم نظام ذكي قادر على إزالة التمويه من الصور الضبابية وتحويلها إلى صور واضحة. تعتمد مقاربتنا على الأساليب الجديدة في التعلم العميق، وبشكل خاص الشبكات التوليدية التنافسية. يتكون المشروع أولاً من تقديم نماذج GAN و MNET، وكذلك تحويل الموجات. بعد ذلك، سنقوم بإنشاء وتدريب هذه النماذج باستخدام أدوات مخصصة لمجالات التعرف المراقب.

الكلمات المفتاحية: شبكات الخصومة التوليدية، والصور الحادة، وإنهاء الثورة، والصور الضبابية، والتعلم العميق

Résumé

L'objectif de ce projet est de concevoir un système intelligent capable de déconvoluer des images floues pour avoir des images nettes. Notre approche est basée sur les nouvelles méthodes de l'apprentissage profond, en particulier les réseaux génératifs antagonistes. Le projet consiste d'abord en une présentation des modèles GAN et MNET, ainsi que de la transformation en ondelettes. Ensuite, nous créerons et entraînerons ces modèles en utilisant des outils dédiés aux domaines de la reconnaissance supervisée.

Mots clés : réseaux génératifs adversaires, modèle MNET, images nettes, déconvoluer, image floues, apprentissage profond.

Abstract

The objective of this project is to design an intelligent system capable of deconvolving blurry images into sharp images. Our approach is based on new methods of deep learning, particularly generative adversarial networks. The project first consists of a presentation of the GAN and MNET models, as well as wavelet transformation. Then, we will create and train these models using tools dedicated to the fields of supervised recognition.

Keywords: generative adversarial networks, sharp images, deconvolution, blurry images, deep learning.

Liste des acronymes et abréviations

| | |
|----------------|--|
| AI | Artificial Intelligence. |
| ML | Machine Learning. |
| DL | Deep Learning. |
| GAN | Generative Adversarial Network. |
| FC | Fully Connected Layer. |
| DWT | Discrete Wavelet Transform |
| IDWT | Inverse Discrete Wavelet Transform |
| CWT | Continuos Wavelet Transform |
| PSNR | Peak Signal-to-Noise Ratio. |
| MSE | Mean Squared Error. |
| BRISQUE | Blind/Referenceless Image Spatial Quality Evaluator. |
| NIQE | Natural Image Quality Evaluator. |
| PIQE | Perception Image Quality Evaluator. |

Tables des matières

| | |
|---|----|
| Remerciement | |
| Dédicace | |
| Résumé | |
| Liste des acronymes et abréviations | |
| Table des matières | |
| Liste des figures | |
| Liste des tableaux | |
| Introduction générale..... | 1 |
| Chapitre I : Généralités sur Traitement d'image et Apprentissage profond | |
| I.1 Introduction..... | 4 |
| I.2 Image numérique | 4 |
| I.3 Flou..... | 5 |
| I.4 Origine du flou..... | 6 |
| I.4.1 Flou généré optiquement..... | 6 |
| I.4.2 Flou généré numériquement | 8 |
| I.5 Floutage | 10 |
| I.6 Types de flou..... | 10 |
| I.6.1 Flou gaussien | 10 |
| I.6.2 Filtre médian | 12 |
| I.6.3 Flou de mouvement..... | 12 |
| I.6.4 Convolution..... | 12 |
| I.7 Technique classique pour déflouter | 13 |
| I.7.1 Déconvolution image | 13 |
| I.7.2 Déconvolution aveugle (Blind déconvolution) | 13 |
| I.7.3 Déconvolution non aveugle | 13 |
| I.8 Transformée en ondelette..... | 14 |
| I.8.1 Transformée en ondelette discrète (DWT) | 14 |
| I.9 Intelligence artificielle | 19 |
| I.10 Apprentissage automatique | 19 |
| I.11 Apprentissage profond..... | 19 |
| I.12 Réseau de neurone | 20 |
| I.13 Perceptron monocouche..... | 20 |

| | | |
|--------|---|----|
| I.14 | Perceptron multicouche | 21 |
| I.15 | Réseau de neurone U-NET | 21 |
| I.15.1 | Architecture de U-NET | 21 |
| I.16 | Réseau antagoniste génératif (GAN) | 23 |
| I.17 | Architecture du réseau antagoniste génératif (GAN)..... | 23 |
| I.17.1 | Modèle générateur | 23 |
| I.17.2 | Modèle discriminateur | 24 |
| I.18 | Conclusion..... | 24 |

Chapitre II : Méthodes utilisées

| | | |
|--------|--|----|
| II.1 | Introduction..... | 26 |
| II.2 | Méthodologie..... | 26 |
| II.3 | Générateur..... | 27 |
| II.3.1 | Définition..... | 27 |
| II.3.2 | Architecture M-Net améliorée..... | 27 |
| II.4 | Discriminateur..... | 33 |
| II.5 | Fonctions de perte (Loss function) | 35 |
| II.5.1 | Perte du générateur et du discriminateur : | 35 |
| II.5.2 | Perte de Charbonnier | 35 |
| II.5.3 | Fonction de Perte Perceptuelle | 35 |
| II.6 | Conclusion..... | 36 |

Chapitre III : Résultats et Implémentations

| | | |
|-------|---|----|
| III.1 | Introduction..... | 38 |
| III.2 | Outils utilisés..... | 38 |
| III.3 | Datasets..... | 38 |
| III.4 | Détails de l'implémentation..... | 39 |
| III.5 | Métriques d'évaluation de la qualité de l'image | 39 |
| III.6 | Résultats et discussion..... | 41 |
| III.7 | Conclusion..... | 47 |
| | Conclusion générale..... | 49 |
| | Références..... | 51 |

Liste des figures

| N° | Figure | Page |
|-----------|---|-----------|
| 01 | Figure I. 1: Image numérique | 5 |
| 02 | Figure I. 2: Image couleur | 5 |
| 03 | Figure I. 3: Diaphragme collé, avant et après la lentille | 6 |
| 04 | Figure I. 4: flou de mouvement | 7 |
| 05 | Figure I. 5: Profondeur de champ | 7 |
| 06 | Figure I. 6: différents types de capteurs | 8 |
| 07 | Figure I. 7: image avec bruit numérique | 9 |
| 08 | Figure I. 8: Compression sans perte ou avec perte | 9 |
| 09 | Figure I. 9: Redimensionnement d'une image | 10 |
| 10 | Figure I. 10: La différence entre un petit et un grand flou gaussien | 11 |
| 11 | Figure I. 11: Calcul de la valeur médian | 12 |
| 12 | Figure I. 12: noyau de convolution (au centre) contient les coefficients de pondération | 13 |
| 13 | Figure I. 13: Application du filtre de Wiener sur une image dégradée par du bruit de sel et de poivre | 14 |
| 14 | Figure I. 14: Implémentation d'une transformée en ondelettes dans un banc de filtres | 18 |
| 15 | Figure I.15: La relation entre l'intelligence artificielle, le ML et le Deep Learning | 19 |
| 16 | Figure I.16 : Réseau de neurone | 20 |
| 17 | Figure I.17: Perceptron Monocouche | 20 |
| 18 | Figure I.18: Perceptron Multicouche | 21 |
| 19 | Figure I.19: L'architecture UNET | 23 |
| 20 | Figure I.20: L'architecture du GAN | 24 |
| 21 | Figure II. 1: Déconvolution d'images basée sur GAN | 27 |
| 22 | Figure II. 2: Architecture M-Net | 27 |
| 23 | Figure II. 3: Architecture proposée | 28 |
| 24 | Figure II. 4: Half Wavelet Attention Block (HWAB) | 29 |
| 25 | Figure II. 5: Dual attention unit (DAU) | 30 |
| 26 | Figure II. 6: Selective kernel feature fusion (SKFF) | 32 |

| | | |
|-----------|--|-----------|
| 27 | Figure II. 7: Exemple de modèle du Discriminateur | 33 |
| 28 | Figure II. 8: Discriminateur Markovien (PatchGAN) | 34 |
| 29 | Figure II. 9: Architecture du discriminateur Markovien (PatchGAN) | 34 |
| 30 | Figure III.1 : Résultats de notre système | 42 |
| 31 | Figure III.2 : Comparaison qualitative des résultats obtenus avec la déconvolution par aveugle | 43 |
| 32 | Figure III.3 : Comparaison qualitative des résultats obtenus avec la déconvolution par le filtre de Gaussien | 44 |
| 33 | Figure III.4 : Comparaison qualitative des résultats obtenus avec la déconvolution par le filtre de wiener | 45 |

Liste des tableaux

| N° | Tableaux | Page |
|----|--|------|
| 01 | Tableau III.1 : Comparaison quantitative des résultats sur des ensembles de données. (+) indique que le score le plus élevé est meilleur, (-) indique le score le plus bas est meilleur. Les deux meilleurs résultats sont en rouge et bleu respectivement. | 46 |

Introduction générale

Introduction générale

Les images capturées par une caméra subissent souvent diverses formes de dégradation lors du processus de formation d'images, ce qui modifie la représentation de l'objet original. Parmi ces altérations, le flou est un effet indésirable souvent observé, dû à différentes causes comme la distorsion, les aberrations optiques et le mouvement de la caméra ou de l'objet, ce qui altère la qualité des images et rend parfois difficile l'interprétation des détails importants. Au fil du temps, diverses méthodes de déconvolution ont été mises au point pour résoudre le problème de flou dans les images, y compris des méthodes classiques telles que le filtre de Wiener, la déconvolution aveugle (Blind Deconvolution) et le filtrage inverse (Inverse Filtering). Bien que ces techniques puissent améliorer la netteté des images, elles présentent des limites en termes de performances, de flexibilité et de capacité à gérer efficacement les différents types de flou.

Les techniques d'apprentissage profond, en particulier les réseaux antagonistes génératifs (GANs), ont ouvert de nouvelles opportunités pour corriger les images floues. Les modèles d'apprentissage non supervisés appelés GANs sont constitués de deux réseaux de neurones en concurrence : un générateur et un discriminateur. Le générateur génère de nouvelles images à partir des données d'entrée, tandis que le discriminateur cherche à différencier les images réelles des images générées. Le générateur est incité par cette compétition à générer des images de plus en plus réalistes, ce qui permet d'acquérir des connaissances sur les caractéristiques des données d'entraînement. En ce qui concerne la déconvolution d'images floues, il est possible d'entraîner les GANs à produire des images nettes à partir d'images floues en prenant connaissance des caractéristiques des images nettes à partir de données d'entraînement. Il est possible de combiner des architectures aux GANs afin d'améliorer les performances et la qualité des résultats.

Ce projet vise à explorer et à mettre en œuvre des techniques avancées de déconvolution d'images en utilisant des réseaux antagonistes génératifs (GANs). Il s'agit de montrer comment les GANs, associés à d'autres architectures de réseaux de neurones, peuvent être exploités afin d'améliorer la qualité des images floues capturées par les caméras. Grâce à ces méthodes, notre objectif est de produire des images plus nettes et de meilleure qualité, ce qui pourrait avoir des utilisations importantes dans différents domaines, tels que la photographie et l'imagerie médicale.

La mémoire contient trois différents chapitres :

- Le 1er chapitre présente le concept des images floues, les techniques de la déconvolution et de donner certaines notions de base de l'architecture de U-NET et le réseau antagoniste génératif (GAN) que nous utilisons dans notre projet.
- Le 2ème chapitre, Introduit la technique de déconvolution basée sur les réseaux antagonistes génératifs (GAN : Generative Adversarial Neural network en anglais) proposé.
- Le 3ème chapitre montre les résultats obtenus et une comparaison avec des méthodes classiques.
- Enfin, une conclusion générale pour résumer les résultats obtenus et les perspectives.

Chapitre I :
Généralités sur Traitement
d'image et Apprentissage profond

I.1 Introduction

Au cours du processus de formation d'images, les images capturées par une caméra souffrent généralement de plusieurs degrés de dégradation dans la représentation de l'objet original. Parmi ces dégradations se trouve le flou, le bruit, la distorsion, la compression etc. Le flou est un artefact indésirable courant associé à la formation d'images. Il est un processus déterministe qui peut se produire pour diverses raisons, telles que la distorsion, les aberrations optiques, le mouvement, etc. On trouve plusieurs techniques pour résoudre le problème : techniques classiques, filtre Wiener, déconvolution aveugle (Blind Deconvolution), filtre inverse (inverse filtering). Actuellement, des techniques avancées d'apprentissage en profondeur telles que; l'architecture U-Net et les réseaux antagonistes génératifs (GAN) sont de plus en plus utilisées et les résultats obtenus sont prometteur et encourageant.

En utilisant ces techniques, les chercheurs et les ingénieurs peuvent obtenir des résultats exceptionnels dans la correction d'images floues, ce qui permet d'obtenir des images plus nettes et de meilleure qualité.

Le but de ce chapitre est de présenter le concept des images floues, les techniques de la déconvolution et de donner certaines notions de base de l'architecture de U-NET et le réseau antagoniste génératif (GAN) que nous utilisons dans notre projet.

I.2 Image numérique

L'image est une représentation visuelle de quelque chose, qu'elle soit naturelle ou artificielle, visuelle, tangible ou conceptuelle. Elle peut créer un lien direct avec son modèle, ou au contraire y être liée par un lien plus symbolique [1].

Une image numérique (voir la figure I.1) représente une image réelle sous la forme d'un ensemble de nombres qui peut être stocké et géré par un ordinateur numérique. Afin de traduire l'image en chiffres, elle est divisée en petites zones appelées pixels (éléments d'image). Chaque pixel est enregistré par le dispositif d'imagerie avec un nombre, ou un petit ensemble de nombres, qui décrivent certaines propriétés de ce pixel, comme sa luminosité (intensité de la lumière) ou sa couleur. Les nombres sont placés dans un tableau de lignes et de colonnes correspondant aux positions verticales et horizontales des pixels de l'image [2].

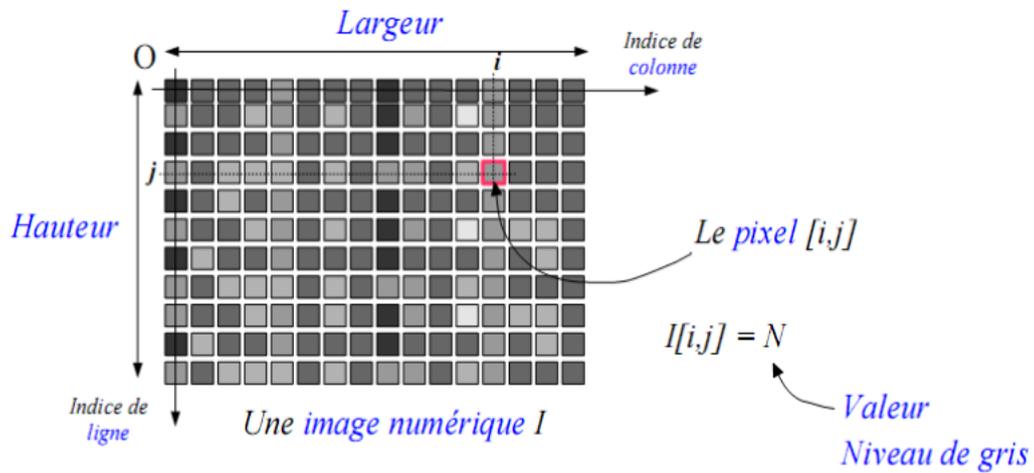


Figure I. 1: Image numérique.

Une image couleur (voir la figure I.2) est la composition de trois (ou plus) images en niveaux de gris sur trois composantes (ou plus). On définit donc trois plans de niveaux de gris, un rouge, un vert et un bleu. La couleur finale est obtenue par synthèse additive de ces trois (ou plus) composantes [3].

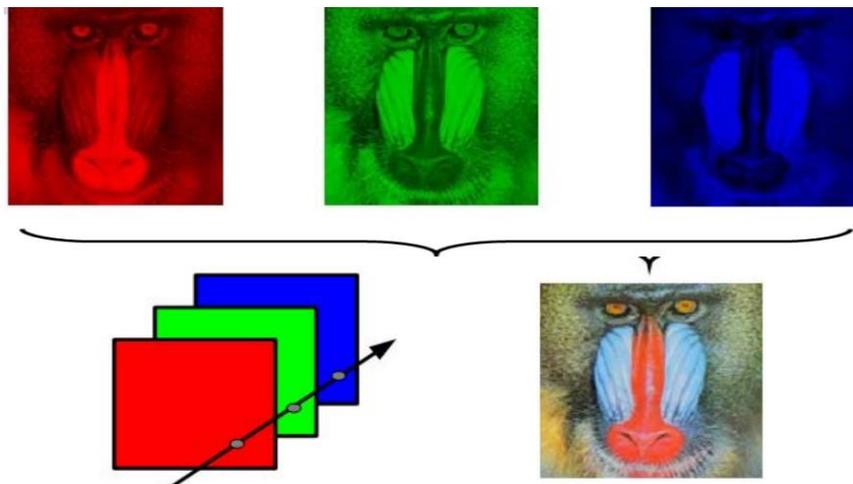


Figure I. 2: Image couleur.

I.3 Flou

Le flou est un manque de netteté de l'image, provenant soit d'une mise au point défectueuse à la projection, soit d'un effet délibéré à la prise de vues [4].

I.4 Origine du flou

Le flou, qu'il soit optique ou numérique, résulte de la combinaison de différents facteurs qui affectent la manière dont la lumière est captée, focalisée et transmise pour former une image.

I.4.1 Flou généré optiquement

I.4.1.1 Distorsion

La distorsion apparaît lorsque l'agrandissement de l'objet n'est pas uniforme. Elle se traduit par une déformation de l'image.

La distorsion résulte du fait que les rayons traversant les bords de la lentille sont plus déviés que ceux passant au voisinage du centre de la lentille.

Pour la mettre en évidence, on forme l'image d'un objet étendu en utilisant un petit diaphragme pour limiter l'influence des autres aberrations (voir la figure I.3) [5].

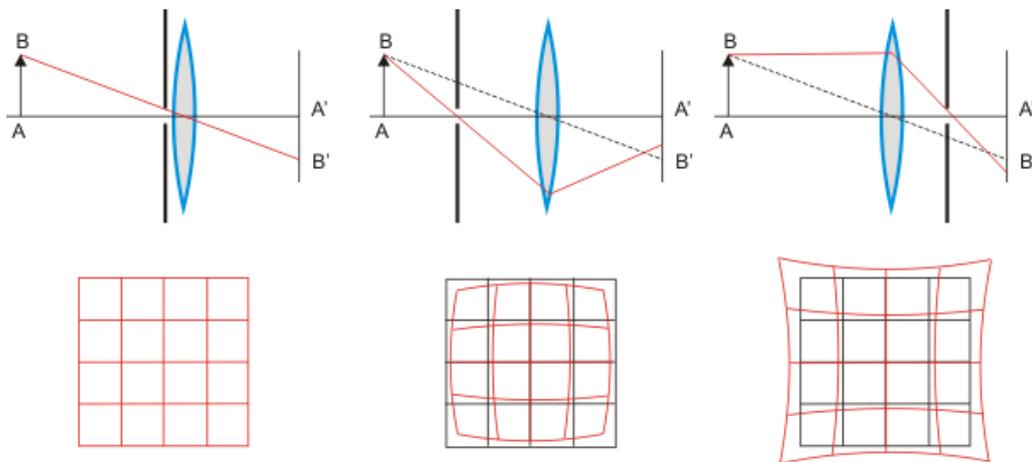


Figure I. 3: Diaphragme collé, avant et après la lentille.

I.4.1.2 Mouvement

Si l'objet ou l'appareil photo bougent pendant l'acquisition de l'image, un flou de mouvement est créé. Ce flou est d'autant plus important que la vitesse de déplacement est élevée ou que le temps d'exposition est long (voir la figure I.4) [6].



Figure I. 4: flou de mouvement [7].

I.4.1.3 Profondeur de champ

La zone nette de l'image est limitée par la profondeur de champ. Seuls les objets situés à une distance précise de l'appareil photo seront nets. Les objets en dehors de cette zone apparaîtront flous [8], illustrée dans la Figure I.5.

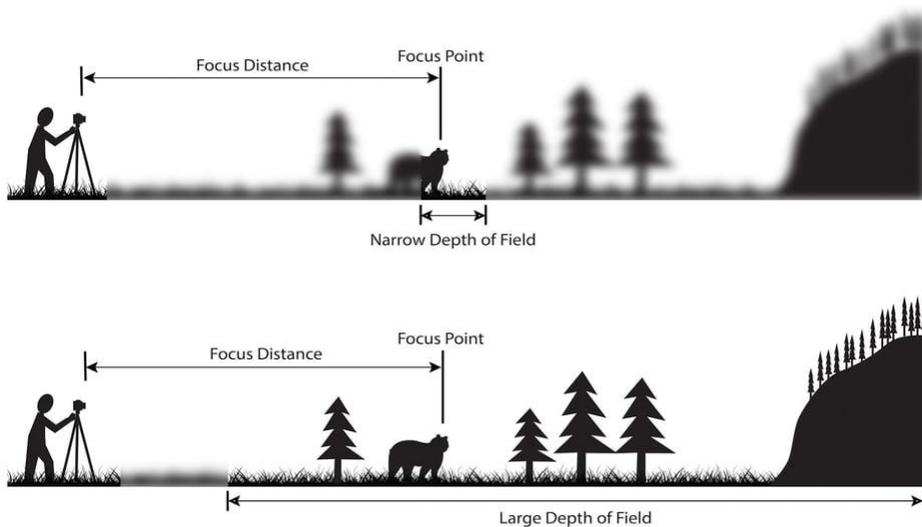


Figure I. 5: Profondeur de champ.

I.4.2 Flou généré numériquement

I.4.2.1 Résolution des capteurs

Les capteurs numériques capturent l'image en la divisant en une matrice de pixels. La résolution du capteur, définie par le nombre de pixels, détermine la limite de la finesse des détails que l'image peut capturer. Un capteur avec une faible résolution produira des images floues, surtout pour les sujets avec des détails fins [9],(voir la figure I.6).

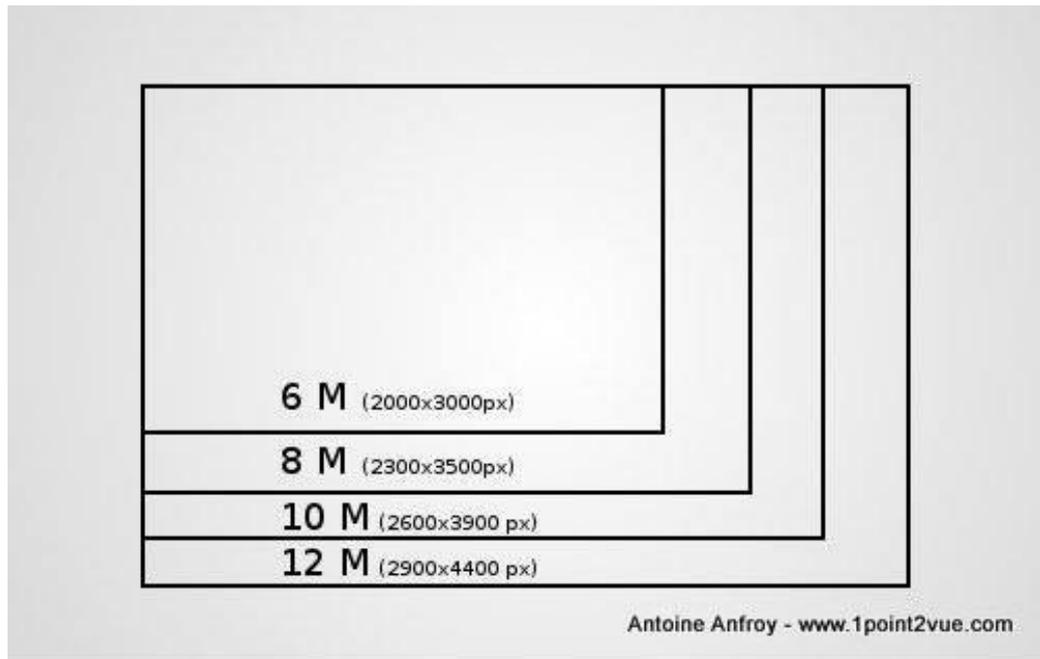


Figure I. 6:diffèrent types de capteurs [10].

I.4.2.2 Flou numérique

Le flou numérique est un phénomène indésirable qui apparaît dans les images numériques. Il peut être causé par des facteurs tels que ; la sensibilité du capteur, la température de fonctionnement ou encore les processus de conversion analogique-numérique. Le bruit numérique peut donner un aspect flou et granuleux à l'image [11],(voir la figure I.7).

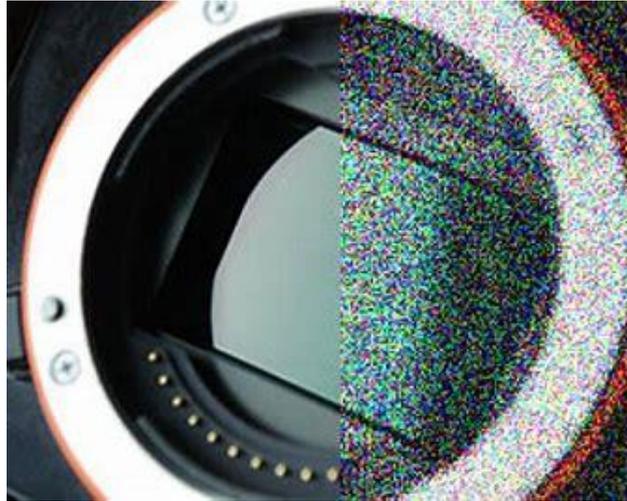


Figure I. 7: image avec bruit numérique [12].

I.4.2.3 Compression d'image

Pour réduire la taille des fichiers numériques, des techniques de compression sont souvent utilisées. La compression d'image peut entraîner une perte de détails, ce qui peut se manifester par un flou, surtout si le taux de compression est élevé [13], comme illustré dans la figure I.8



Figure I. 8: Compression sans perte ou avec perte.

I.4.2.4 Redimensionnement

Le redimensionnement ou la modification de la résolution d'une image numérique peut également entraîner un flou. Lors de l'opération de sous échantillonnage de nouveaux pixels sont créés ou supprimés, ce qui peut altérer la qualité de l'image et la rendre floue [14],(voir la figure I.9).

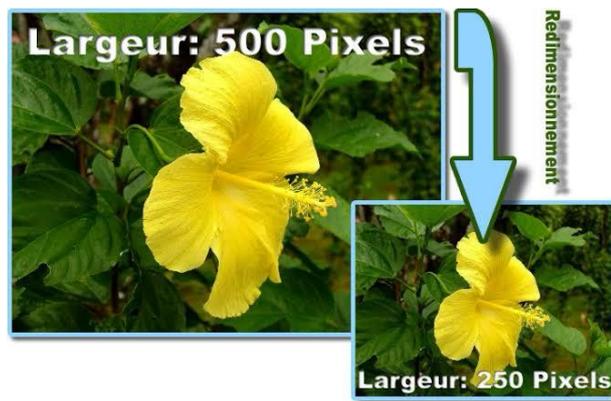


Figure I. 9: Redimensionnement d'une image.

I.5 Floutage

Le floutage d'une image est un processus qui rend l'image moins nette et réduit son niveau de détail. Il déforme les détails d'une image, ce qui la rend moins claire. Nous pouvons flouter une image à l'aide de différents filtres de flou passe-bas [15].

I.6 Types de flou

Les filtres flous sont des outils numériques utilisés pour adoucir, estomper ou modifier la netteté d'une image. Ils sont couramment utilisés dans divers domaines tels que ; la retouche photo, le graphisme, le traitement d'image et la vision par ordinateur. Chaque type de flou produit un effet spécifique sur l'image et sert des objectifs différents. Voici quelques types de flous couramment utilisés :

I.6.1 Flou gaussien

Le flou gaussien est l'un des filtres flous les plus populaires. Il utilise une fonction de distribution gaussienne pour pondérer les pixels voisins, ce qui crée un effet d'atténuation progressive des détails fins vers les bords de l'image. Ce type de flou est souvent utilisé pour

créer un effet de bokeh artistique, comme illustré dans la figure I.10, lisser les imperfections ou réduire le bruit numérique [16].



Figure I. 10: La différence entre un petit et un grand flou gaussien.

L'équation de la courbe est

$$h(x, y) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\left(\frac{x^2+y^2}{2\sigma}\right)} \quad (\text{I. 1})$$

Avec :

- $h(x,y)$: La valeur de la fonction gaussienne en un point (x,y) .
- σ : L'écart-type, une mesure de la dispersion ou de la largeur de la courbe gaussienne.
- π : Une constante mathématique (approximativement 3.14159).
- e : La base du logarithme naturel (approximativement 2.71828).
- x et y : Les variables indépendantes de la fonction, représentant des coordonnées dans un plan bidimensionnel.

I.6.2 Filtre médian

Le filtre médian remplace chaque pixel par la valeur médiane des couleurs dans son voisinage, (voir la figure I.11). Contrairement au flou gaussien, qui préserve les contours généraux de l'image, le flou médian peut estomper les détails fins tout en préservant les structures et les textures importantes. Il est souvent utilisé pour réduire le bruit tout en conservant la netteté des contours et des arêtes [17].

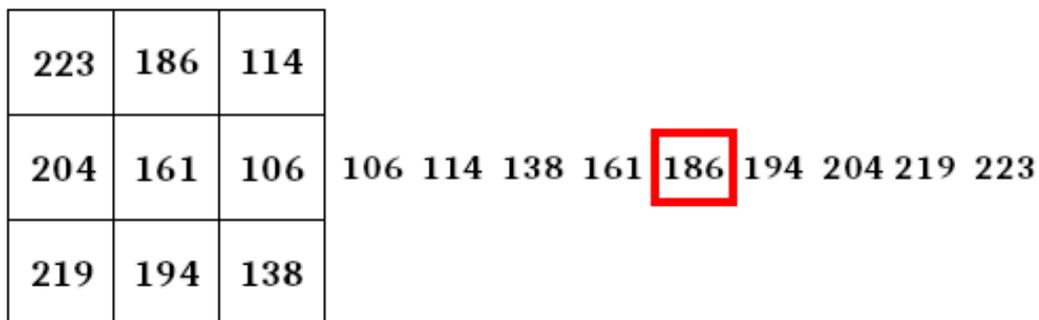


Figure I. 11: Calcul de la valeur médian.

I.6.3 Flou de mouvement

Le flou de mouvement est la trace apparente d'objets en mouvement dans une photographie ou une séquence d'images. Il se produit lorsque l'image enregistrée change pendant l'enregistrement d'une seule exposition, en raison d'un mouvement rapide ou d'une longue exposition [18].

I.6.4 Convolution

La convolution, ou produit de convolution, est une généralisation du filtre moyenneur où l'on considère cette fois une moyenne pondérée. La fenêtre glissante est alors elle-même une image qui contient les coefficients de pondération. On l'appelle généralement noyau déconvolution ou masque de convolution (kernel ou mask en anglais) [19], (voir la figure I.12).

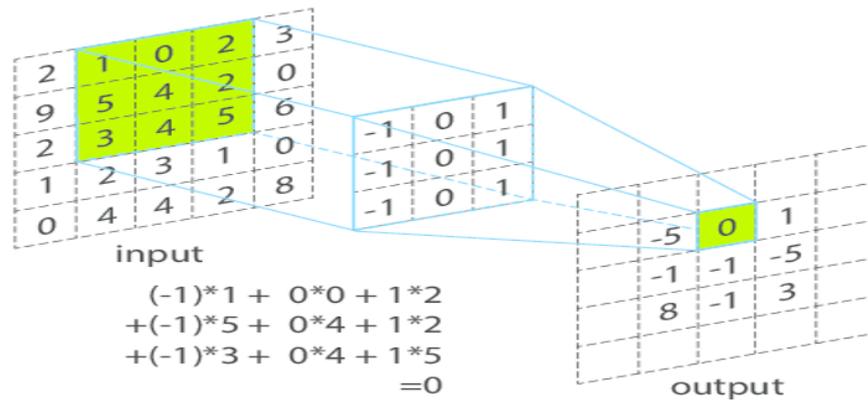


Figure I. 12:noyau de convolution (au centre) contient les coefficients de pondération.

I.7 Technique classique pour déflouter

I.7.1 Déconvolution image

La déconvolution est une méthode de traitement d'image utilisés pour améliorer une image dégradée et souvent floue afin de la rendre nette. Son objectif principal est d'aider à retrouver les détails les plus fins de l'image d'origine qui ont été perdus [19].

I.7.2 Déconvolution aveugle (Blind déconvolution)

Déconvolution aveugle est un algorithme pour estimer la fonction de flou (noyau de convolution) et en même temps effectuer un défloutage de l'image. Le but est de trouver à la fois l'image nette et les détails de la manière dont le flou a été effectué.

I.7.3 Déconvolution non aveugle

➤ Filtre de Wiener (linéaire) :

Le filtrage de Wiener est l'une des approches les plus anciennes et les plus connues de la restauration d'images linéaires. Le filtre de Wiener est le filtre le plus approprié pour gérer les images endommagées par le bruit et le mouvement. Il a été utilisé pour filtrer l'image bruyante observée. Il est utilisé pour le bruit additif, les spectres de bruit et le signal stationnaire de l'image (voire la figure I.13) [20].

L'objectif de cette technique est de trouver une estimation de l'image originale f de manière à minimiser l'erreur quadratique moyenne entre eux. Cette mesure d'erreur est donnée par :

$$e^2 = E \{ (f - \hat{f})^2 \} \quad (\text{I. 1})$$

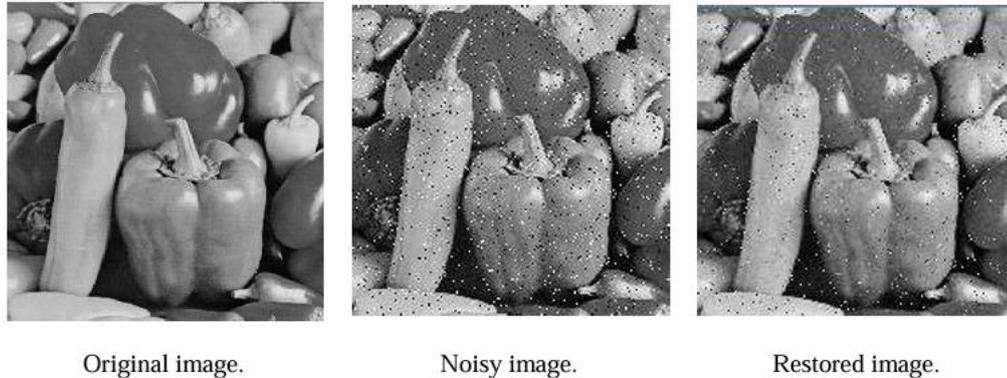


Figure I. 13: Application du filtre de Wiener sur une image dégradée par du bruit de sel et de poivre. [20]

I.8 Transformée en ondelette

La transformée en ondelettes décompose les images d'entrée en bandes de fréquences. Cela permet d'extraire et d'améliorer les détails haute fréquence tout en supprimant le bruit présent dans les bandes basse fréquence. Contrairement à la transformée de Fourier qui décompose un signal en sinus à fréquences spécifiques, la transformation en ondelettes se base sur des versions décalées et dilatées d'une 'ondelette', une oscillation en forme d'onde qui décroît rapidement. Elle permet une analyse à la fois temporelle et fréquentielle, avec une résolution en temps plus fine pour les hautes fréquences et une meilleure résolution en fréquence pour les basses fréquences [21].

En plus de la transformation continue, il existe une version discrète de la transformation en ondelettes, plus adaptée à l'analyse de signaux numériques.

I.8.1 Transformée en ondelette discrète (DWT)

La transformée en ondelette discrète est une technique du traitement de signal et un outil qui découpe les données, les fonctions ou les opérateurs en composantes fréquentielles suivant une résolution adaptée à l'échelle [21].

Si $x(t)$ est une fonction réelle de variable réelle la transformée en ondelettes de f est:

$$g(a, b) = \frac{1}{\sqrt{a}} \int_{t=-\infty}^{t=\infty} x(t) \psi_{a, b}(t) dt \quad (I.2)$$

Avec :

$a \neq 0$ La fonction $\psi_{a,b}(t)$ est obtenue par translation et dilatation d'une fonction particulière appelée ondelette mère:

$$\psi_{a, b}(t) = \psi\left(\frac{t - b}{a}\right) \quad (I.3)$$

- b : détermine la position et a donné l'échelle.
- Cas d'un signal : a est la fréquence et b le temps.

➤ Définition et échantillonnage de la DWT

Une DWT est normalement définie comme un échantillonnage non redondant de la transformée en ondelettes continue (CWT). La DWT vise à représenter une série temporelle discrète, $x(n)$, comme un ensemble de coefficients d'ondelettes. Ces coefficients sont échantillonnés à partir d'une CWT, généralement de manière à produire un ensemble de fonctions de base orthogonales (ou biorthogonales). Il existe de nombreuses formulations de la DWT, avec des caractéristiques différentes. Ici, l'attention se porte sur le cas des ondelettes orthogonales à support fini [21].

➤ Bases orthogonales et algorithmes efficaces

L'utilisation de bases orthogonales garantit que la représentation n'est pas redondante. En général, l'utilisation de représentations orthogonales conduit à des algorithmes simples pour la reconstruction et la décomposition. Des algorithmes DWT efficaces existent, nécessitant des charges de calcul souvent inférieures à celles nécessaires pour implémenter une transformée de Fourier rapide [21].

➤ Concept de bancs de filtres dans la DWT

La DWT peut être considérée sous plusieurs points de vue équivalents. Ici, nous choisissons de discuter de la DWT à travers le concept de bancs de filtres. Une paire de filtres à réponse impulsionnelle finie (FIR), avec L coefficients, est définie. L'un de ces filtres est un filtre passe-haut, tandis que le second est un filtre passe-bas ; les deux filtres coupent à la moitié

de la fréquence d'échantillonnage [21]. La DWT peut être définie en utilisant ces filtres et en les appliquant de manière récursive. Les filtres sont d'abord appliqués à la série temporelle d'entrée pour produire des composantes passe-bas et passe-haut, $x_1(n)$ et $x_2(n)$ respectivement:

$$x_1(n) = \sum_{k=0}^{L-1} c_k x(n-k) \quad (\text{I.4})$$

$$x_2(n) = \sum_{k=0}^{L-1} d_k x(n-k) \quad (\text{I.5})$$

Où c_k et d_k sont les coefficients des filtres passe-bas et passe-haut. Il est courant de construire le filtre passe-haut basé sur le filtre passe-bas, ce qui est généralement réalisé en utilisant le schéma de renversement alterné, de sorte que les deux ensembles de coefficients de filtre sont liés par :

$$d_k = (-1)^k c_{L-k} \quad (\text{I.6})$$

➤ Sous-échantillonnage et resampling

La sortie des deux filtres est la moitié de la bande passante de la séquence d'entrée, de sorte que $x_1(n)$ occupe la gamme de fréquences inférieures et $x_2(n)$ la bande supérieure. Les sorties de chaque filtre ont une bande passante équivalente à la moitié de celle de la séquence d'entrée $x(n)$, de sorte que ces deux séries temporelles contiennent des informations redondantes. On peut sous-échantillonner les deux sorties de filtre à la moitié du taux d'échantillonnage initial. Le sous-échantillonnage est réalisé en éliminant un échantillon sur deux dans chaque séquence. Ce processus de resampling signifie que les deux signaux occupent à nouveau la bande passante complète. En raison de la conception imparfaite des deux filtres FIR, le resampling introduit un aliasing dans les deux composantes [21].

➤ Annulation de l'aliasing et reconstruction du signal

Pour contrer les artefacts introduits par l'aliasing, on peut concevoir les deux filtres de manière à ce que les deux effets d'aliasing se compensent. Si les filtres sont conçus pour garantir l'annulation de l'aliasing, il est possible de reconstruire le signal d'origine à partir des deux séquences sous-échantillonnées. La récupération de la séquence d'entrée originale à partir des deux séquences sous-échantillonnées est réalisée en sur-échantillonnant les deux signaux sous-

échantillonnés, en les filtrant, puis en sommant les deux composantes. Pour une transformée en ondelettes orthogonale, utilisant des filtres FIR, les filtres de synthèse sont les mêmes que ceux utilisés pour la décomposition, mais avec les réponses impulsionnelles inversées dans le temps. Le processus de décomposition d'un signal en deux séquences sous-échantillonnées puis de reconstruction du signal original est illustré dans la Figure I.14. Tout au long de la Figure I.14, les symboles $\uparrow 2$ et $\downarrow 2$ sont utilisés pour désigner le sur-échantillonnage et le sous-échantillonnage, et $H_1(z)$ et $H_2(z)$ sont les fonctions de transfert des filtres de décomposition passe-bas et passe-haut respectivement, avec $H_3(z)$ et $H_4(z)$ étant les fonctions de transfert des filtres de synthèse correspondants [21].

➤ Structure arborescente des bancs de filtres

Le schéma montré dans la Figure I.14(A), représente le bloc de construction de base d'un banc de filtres en ondelettes. Ces éléments sont appliqués dans une structure arborescente. Après avoir appliqué le filtrage et le sous-échantillonnage à la série temporelle d'origine pour générer des composantes basse et haute fréquence, l'opération peut être répétée sur l'une des nouvelles composantes. Le plus souvent, on ne décompose que la composante basse fréquence (grande échelle). Cette forme de schéma est représentée dans la Figure I.14(B) [21].

➤ Décomposition en échelles multiples

Chaque étape de la décomposition de la séquence d'entrée produit deux séquences, chacune de longueur égale à la moitié de la séquence originale. Ainsi, chaque étape de la décomposition préserve le nombre total de points de données. Le résultat final de la décomposition, indiqué sur le côté droit de la Figure I.14 (B), consiste en N coefficients d'ondelettes, qui représentent la séquence d'entrée originale de N points (on suppose ici que N est une puissance de 2). Les N coefficients d'ondelettes comprennent $N/2$ coefficients à la plus petite échelle, $N/4$ coefficients à l'échelle suivante, etc., et à la plus grande échelle, il y a deux valeurs uniques. Si l'on considère le nombre total de coefficients requis, à savoir $N/2, N/4, \dots, 2, 1$ et un coefficient supplémentaire, il est facile de montrer que cela donne un total global de N coefficients [21].

➤ Reconstruction de la série temporelle

La reconstruction de la série temporelle à partir des coefficients d'ondelettes se fait en utilisant le processus inverse de la procédure de décomposition. Cette reconstruction est

illustrée dans la Figure I.14(C). À chaque étape, deux signaux sont sur-échantillonnés, filtrés, puis sommés [21].

➤ Fonctions de base et conception des filtres FIR

La dérivation ci-dessus de la DWT cache les fonctions de base utilisées dans la DWT. Ces fonctions sont entièrement spécifiées par le choix des filtres FIR. En fait, les fonctions de base sont la solution d'une équation de dilatation, dont les coefficients sont les coefficients du filtre. Les solutions de ces équations de dilatation peuvent avoir des propriétés indésirables, par exemple, elles peuvent être fractales. Dans une décomposition raisonnable, on souhaiterait que ces fonctions de base orthogonales soient bien comportées, au moins lisses dans une certaine mesure. Cela impose effectivement des contraintes supplémentaires à la procédure de conception du filtre [21].

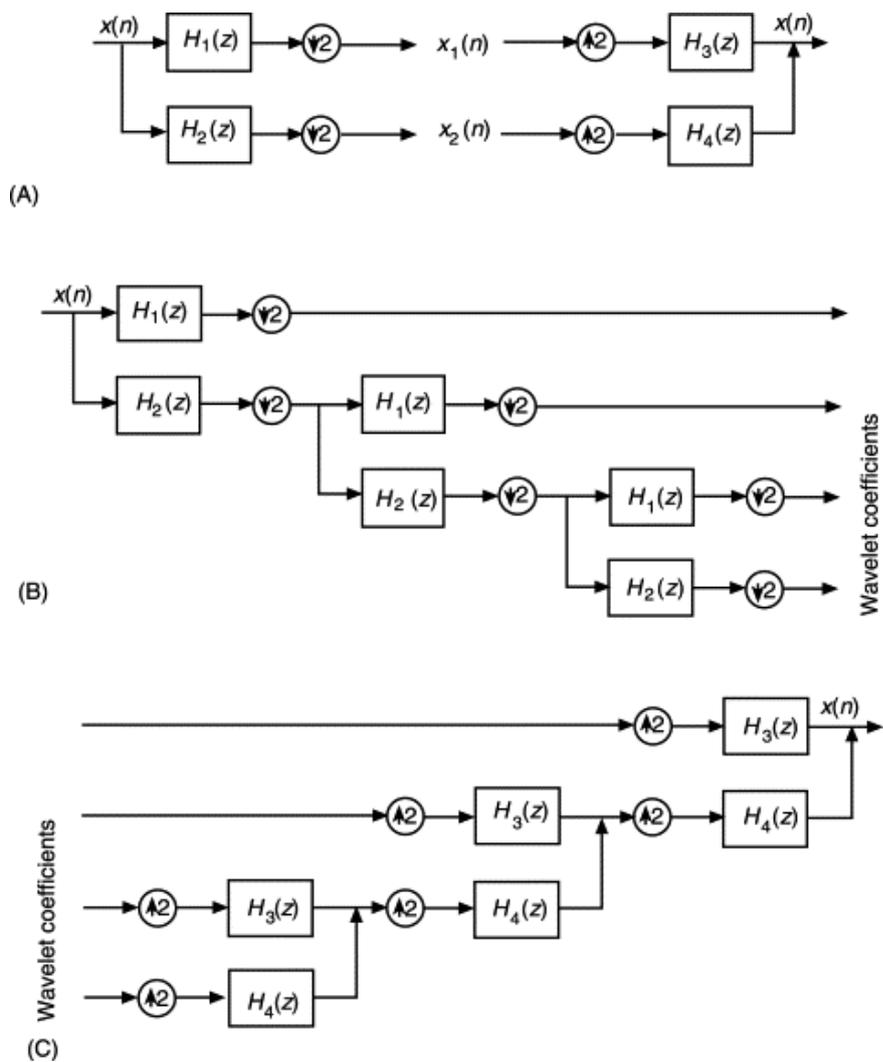


Figure I. 14: Implémentation d'une transformée en ondelettes dans un banc de filtres.

I.9 Intelligence artificielle

L'intelligence artificielle est un ensemble des théories et des techniques développant des programmes informatiques complexes capables de simuler certains traits de l'intelligence humaine (raisonnement, apprentissage...) [22].

I.10 Apprentissage automatique

L'intelligence artificielle (IA) et l'informatique se spécialisent dans l'apprentissage automatique (ML), qui se focalise sur l'exploitation de données et d'algorithmes afin de permettre à l'IA d'imiter la méthode d'apprentissage des humains, avec une amélioration progressive de sa précision [23].

I.11 Apprentissage profond

L'Apprentissage profond Learning est une technique d'intelligence artificielle qui permet à des ordinateurs d'acquérir des compétences en traitement des données en se basant sur le cerveau humain. Les algorithmes d'apprentissage profond peuvent détecter des modèles complexes dans des images, du texte, des sons et d'autres données pour générer des informations et des prévisions précises. Des techniques d'apprentissage profond peuvent être employées afin d'automatiser des tâches qui requièrent habituellement de l'intelligence humaine, comme la description d'images ou la transcription d'un fichier audio en texte [24], voir la figure I.15.

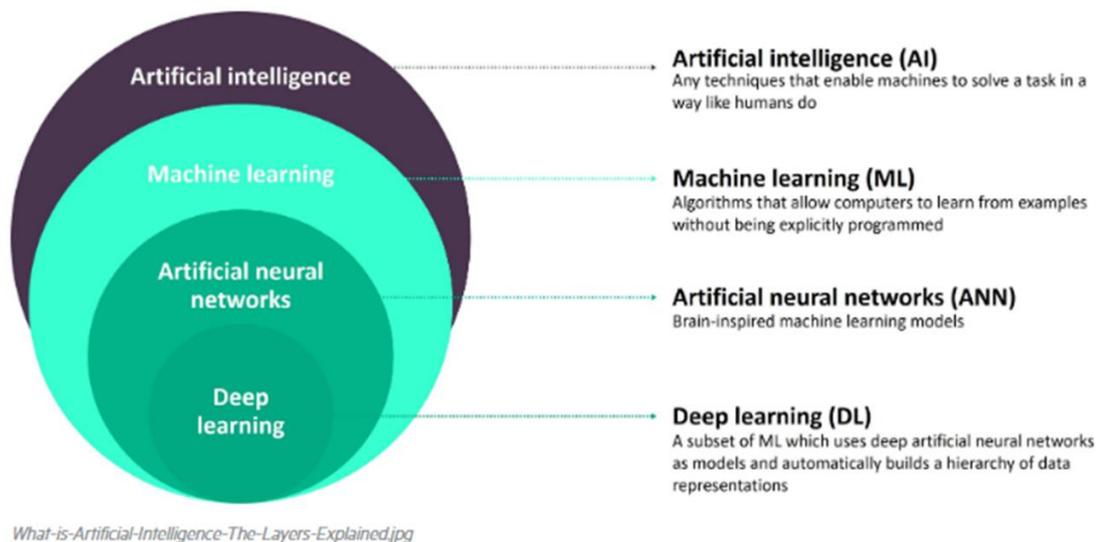


Figure I.15: La relation entre l'intelligence artificielle, le ML et le deep learning.

I.12 Réseau de neurone

Un réseau de neurones artificiels, ou réseau neuronal artificiel, est un système dont la conception est à l'origine schématiquement inspirée du fonctionnement des neurones biologiques [22], comme illustré dans figure I.16.

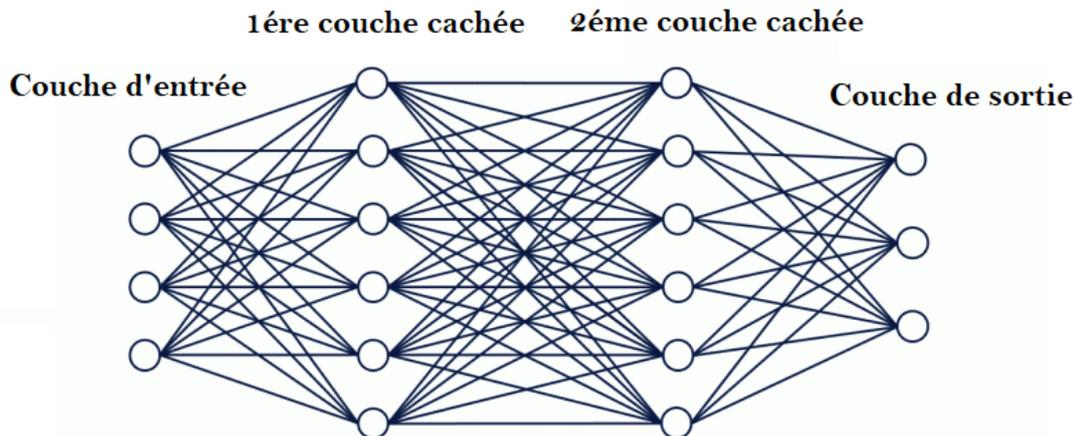


Figure I. 16: Réseau de neurone.

I.13 Perceptron monocouche

Un perceptron monocouche est un réseau de neurones contenant n neurones en entrée et m neurones en sortie. Ici pour simplifier l'écriture du code, nous prendrons un réseau $n-1$ (n neurones en entrée et 1 en sortie). Ce type de réseau est capable d'apprendre à classer des éléments en deux catégories. Par exemple, il est capable d'apprendre à détecter à partir des mots trouvés dans le sujet d'un mail s'il s'agit d'un spam ou non [22], (voir la figure I.17).

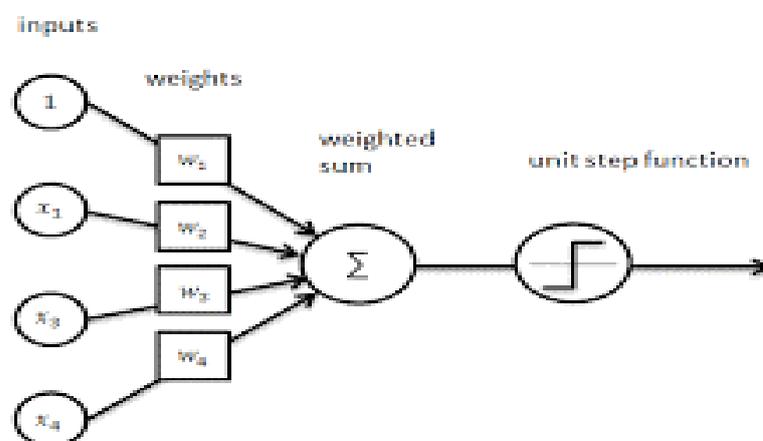


Figure I. 17: Perceptron Monocouche.

I.14 Perceptron multicouche

Le perceptron multicouche (multilayer perceptron MLP) est un type de réseau neuronal artificiel organisé en plusieurs couches au sein desquelles une information circule de la couche d'entrée vers la couche de sortie uniquement ; il s'agit donc d'un réseau à propagation directe (feedforward). Chaque couche est constituée d'un nombre variable de neurones, les neurones de la dernière couche (dite « de sortie ») étant les sorties du système global [22] ,(voir la figure I.18).

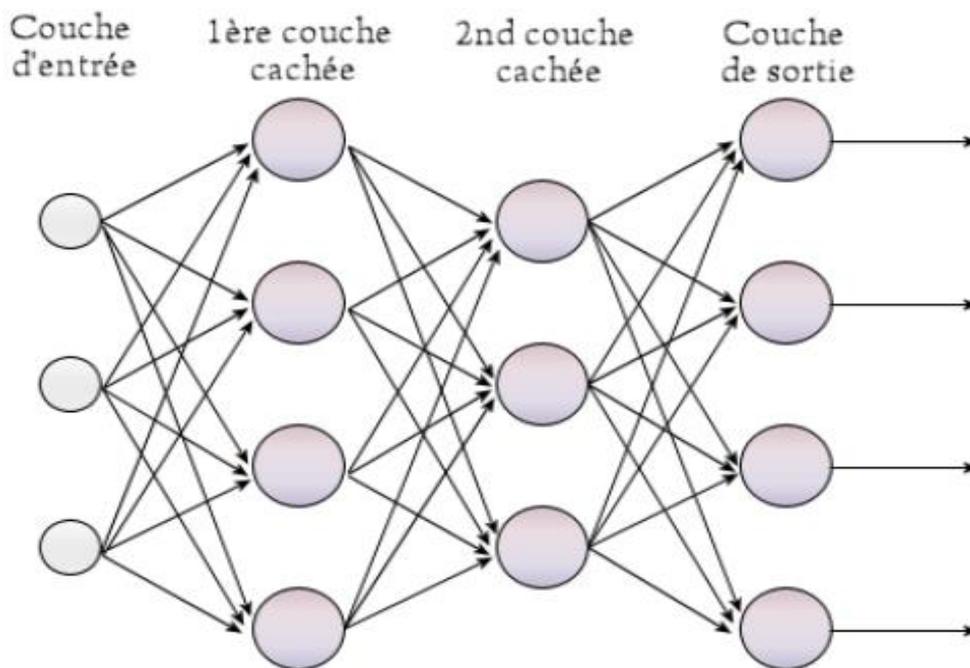


Figure I.18: Perceptron Multicouche.

I.15 Réseau de neurone U-NET

U-NET est un modèle de réseau de neurone convolutif utilisé pour la segmentation d'image en particule la segmentation sémantique, qui a pour objectif de traiter les tâches de vision par ordinateur [25].

I.15.1 Architecture de U-NET

L'architecture de U-NET est composée de deux parties la première appelé contraction (encodeur) et la seconde partie est l'expansion symétrique (décodeur) [26] ,comme illustré dans la figure I.19.

➤ Encodeur

Le réseau d'encodeurs est utilisé pour extraire des fonctionnalités et apprendre une représentation abstraite de l'image d'entrée à travers une séquence de blocs d'encodeurs. Chaque bloc d'encodeur se compose de deux convolutions 3×3 , et elles sont suivies d'une fonction d'activation ReLU. La fonction d'activation ReLU introduit la non-linéarité dans le réseau, ce qui aide à une meilleure généralisation des données d'entraînement. La sortie du ReLU agit comme une connexion de saut pour le bloc de décodeur correspondant. Appliquer ensuite l'opération de max pooling 2×2 pour diminuer la dimension spatiale. Cela réduit le coût de calcul en diminuant le nombre de paramètres pouvant être entraînés [25].

➤ Couches convolutives

Les premières couches de l'encodeur sont des couches convolutives standard qui extraient les caractéristiques spatiales de l'image d'entrée. Ces couches sont généralement composées de noyaux de convolution 3×3 ou 5×5 , suivis d'une fonction d'activation non linéaire (ReLU) et éventuellement d'une opération de normalisation (BatchNorm).

➤ Sous-échantillonnage (max-Pooling)

Après chaque bloc de couches convolutives, une opération de sous-échantillonnage (max-pooling) est appliquée pour réduire la résolution spatiale des caractéristiques, permettant de capturer des informations plus globales et d'augmenter le champ réceptif.

➤ Connexion de saut (connexions résiduelles)

Les connexions de saut se trouvent entre l'encodeur et le décodeur fournissent des informations supplémentaires qui aident le décodeur à un meilleur flux de gradient lors de la rétro-propagation, ce qui aide à l'apprentissage de meilleures représentations par le réseau.

➤ Pont

Le pont relie le codeur et le réseau de décodeurs et complète le flux d'informations. Il se compose de deux convolutions 3×3 , où chaque convolution est suivie d'une fonction d'activation ReLU [25].

➤ **Décodeur**

Le réseau de décodeur est utilisé pour permettre une localisation précise à l'aide de convolutions transposées. Il s'agit donc d'un réseau entièrement convolutif (FCN) qu'il ne contient que des couches convolutives et ne contient aucune couche dense grâce à laquelle il peut accepter des images de n'importe quelle taille [27].

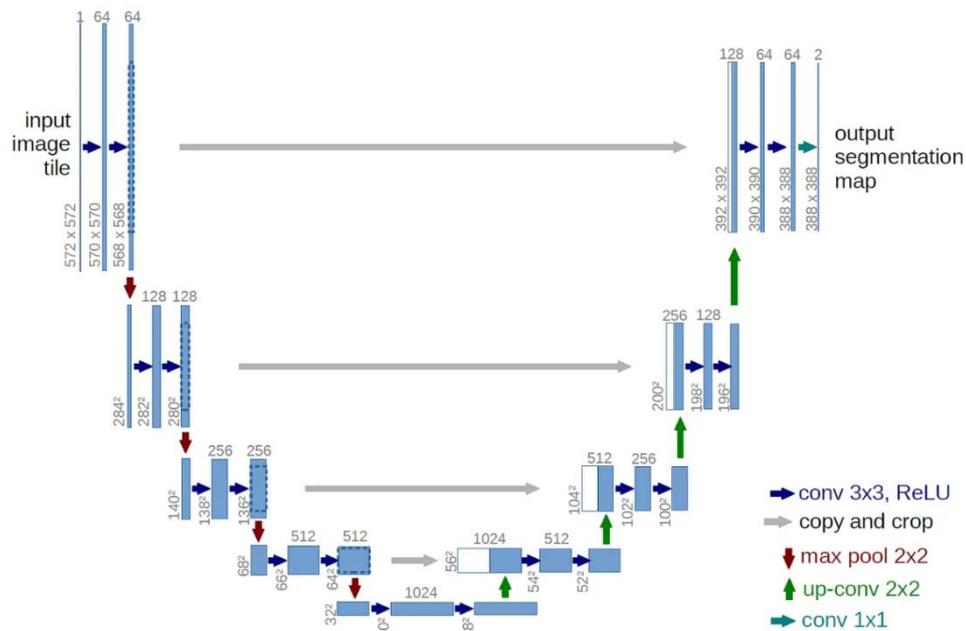


Figure I.19 : L'architecture UNET. [25]

I.16 Réseau antagoniste génératif (GAN)

Le réseau antagoniste génératif est un modèle de réseau de neurone artificiel utilisé dans l'apprentissage non supervisé pour la création des données réalistes. Il se compose de deux réseaux de neurones un générateur et un discriminateur qui sont entraînés en compétition [28]. Il est illustré dans la figure I.20.

I.17 Architecture du réseau antagoniste génératif (GAN)

I.17.1 Modèle générateur

C'est un réseau utilisé pour générer de nouveaux exemples plausibles à partir d'un bruit aléatoire ou d'un vecteur de code latent [28].

I.17.2 Modèle discriminateur

C'est un réseau utilisé pour classer correctement les données comme réels (étiquette 1) ou faux (étiquette 0) [28].

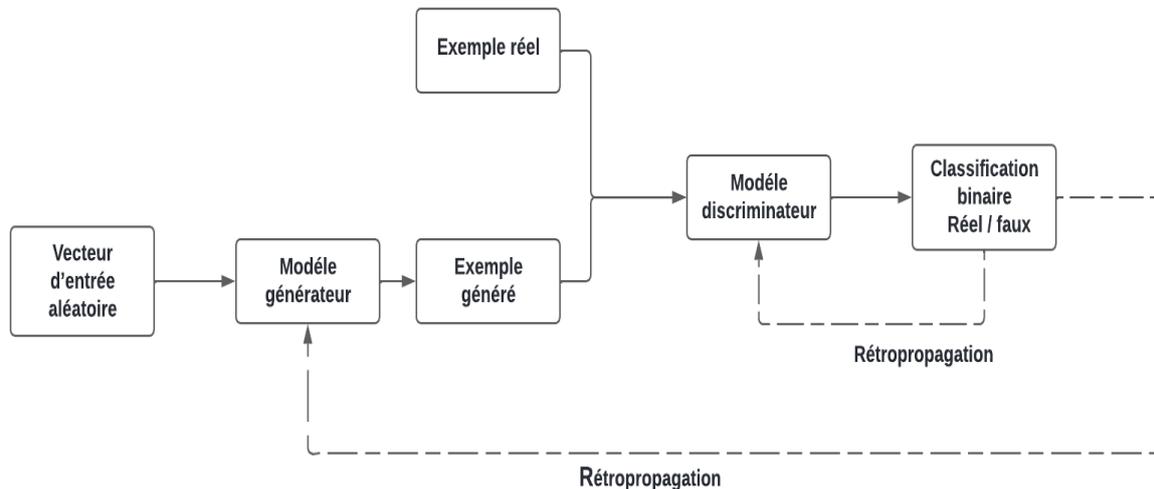


Figure I. 20: L'architecture du GAN.

I.18 Conclusion

Après toutes les études faites sur le flou, les techniques classiques de déconvolution et les notions de base des architectures U-NET et GAN, il est possible de créer un programme capable de corriger les images floues et d'améliorer leur netteté. Cette étude met en évidence l'importance de la restauration d'images floues pour diverses applications, notamment en photographie, en surveillance et en imagerie médicale. En combinant les forces des architectures U-NET et GAN, nous pouvons développer des solutions robustes et efficaces pour améliorer la qualité des images numériques et répondre aux besoins croissants de précision et de clarté dans le traitement d'images.

Chapitre II :
Méthodes utilisées

II.1 Introduction

La déconvolution d'images est un défi crucial dans le domaine du traitement d'images et de la vision par ordinateur, visant à restaurer des images nettes à partir de celles dégradées par le flou. Le flou peut survenir pour diverses raisons, notamment les secousses de l'appareil photo, le mouvement des objets ou la mise au point défectueuse lors de la capture de l'image. Cette dégradation affecte non seulement la qualité visuelle des images mais nuit également à la performance des tâches ultérieures telles que la reconnaissance d'objets, la segmentation d'images et la compréhension de scènes. Les techniques traditionnelles de la déconvolution reposaient sur des filtres prédéfinis et des méthodes de régularisation pour approximer l'image originale. Cependant, ces approches luttent souvent avec des motifs de flou complexes et nécessitent un ajustement intensif des paramètres. L'apprentissage profond a révolutionné l'approche de la déconvolution d'images, introduisant des modèles basés sur les données capables d'apprendre à inverser les effets du flou à partir de grands ensembles de données. Les approches récentes dans les réseaux de neurones convolutifs (CNN) et les réseaux antagonistes génératifs (GAN) ont démontré un succès remarquable dans la production d'images défloutées de haute qualité.

Dans ce chapitre, nous avons proposé une approche de déconvolution basée sur les réseaux antagonistes génératifs (GAN : Generative Adversarial Neural network en anglais). Cette méthode basée sur l'architecture U-Net [1], qui a montré de bons résultats dans plusieurs domaines, et permet de générer des images de haute qualité et réalistes. De plus, nous utilisons la transformée en ondelettes (WT : Wavelet Transform en anglais) pour l'analyse multi-résolution, ce qui facilite l'extraction et la reconstruction améliorées des caractéristiques.

II.2 Méthodologie

L'objectif de notre travail est de concevoir et de mettre en œuvre un modèle d'apprentissage profond capable de générer des images nettes de haute qualité tout en conservant les informations de l'image floue d'entrée, comme illustré dans la figure II.1. Pour cela, nous proposons une architecture basée sur le GAN et l'architecture M-Net, une extension de l'architecture U-Net, combinée avec la transformée en ondelettes (WT) dans le générateur.

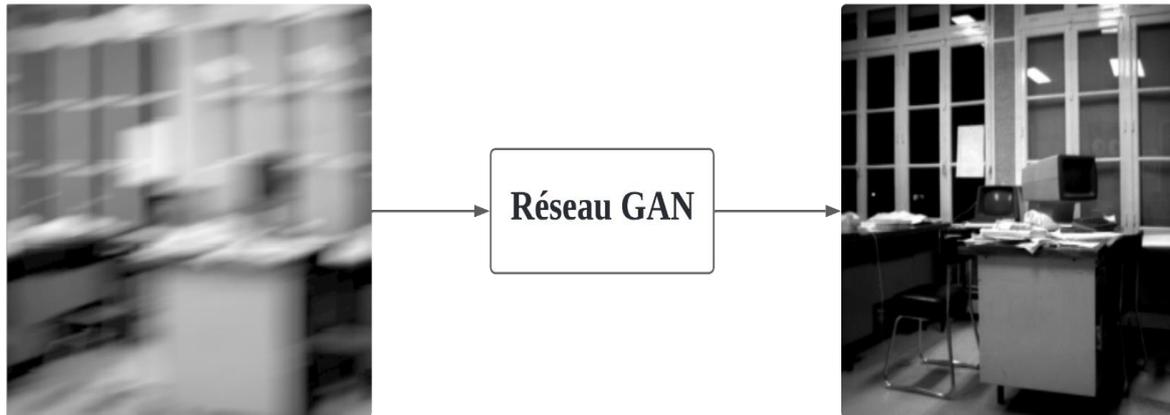


Figure II. 1: Déconvolution d'images basée sur GAN.

II.3 Générateur

II.3.1 Définition

Le générateur apprend à générer des données synthétiques en utilisant les feedbacks du discriminateur, visant à produire des données que ce dernier identifiera comme authentiques. Le générateur est basé sur l'architecture M-Net [29], illustrée dans la Figure II.2.

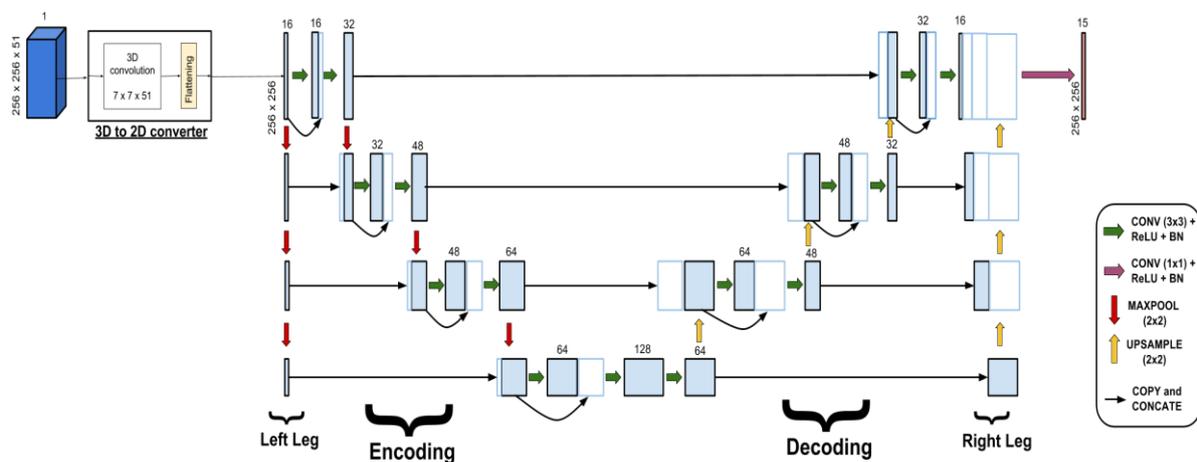


Figure II. 2: Architecture M-Net.

II.3.2 Architecture M-Net améliorée

L'architecture M-Net est une architecture de réseau neuronal en couches qui améliore l'U-Net original pour la segmentation d'images. Cette dernière est particulièrement efficace pour les tâches de segmentation, où la précision et la capacité à discerner les détails fins sont

essentielles. L'architecture M-Net, comme U-Net, introduit plusieurs modifications pour améliorer les performances, Il se compose d'une structure encodeur-décodeur avec des connexions de saut symétriques qui relient les couches correspondantes dans l'encodeur et le décodeur, ce qui est crucial pour reconstruire avec précision des images haute résolution [30]. Ainsi que deux voies latérales (gatepost feature paths) qui fournissent une supervision profonde à l'architecture. Chaque voie comprend quatre étapes. L'encodeur compresse l'image d'entrée en une représentation latente, capturant les caractéristiques essentielles, tandis que le décodeur reconstruit l'image haute résolution à partir de cette représentation.

L'architecture M-Net présente deux principaux défis. Premièrement, l'utilisation du max-pooling dans les chemins gatepost et U-Net pour le sous-échantillonnage entraîne une perte significative d'informations spatiales, ce qui est inapproprié pour les tâches de restauration pixel par pixel. Deuxièmement, l'ajout de chemins gatepost supplémentaires dans le décodeur augmente la complexité du processus de reconstruction en nécessitant la fusion de nombreuses caractéristiques, ce qui alourdit l'ensemble du réseau [31].

Pour adresser ces défis, notre proposition comprend deux améliorations clés. Premièrement, on adopte le sous-échantillonnage par pixel un-shuffle dans le chemin U-Net et le sous-échantillonnage bilinéaire dans le chemin gatepost (pieds). Deuxièmement, on introduit une méthode de fusion de caractéristiques différente, SKFF [32], pour synthétiser l'information dans le processus de reconstruction du décodeur. L'architecture proposée est illustrée dans la figure II.3.

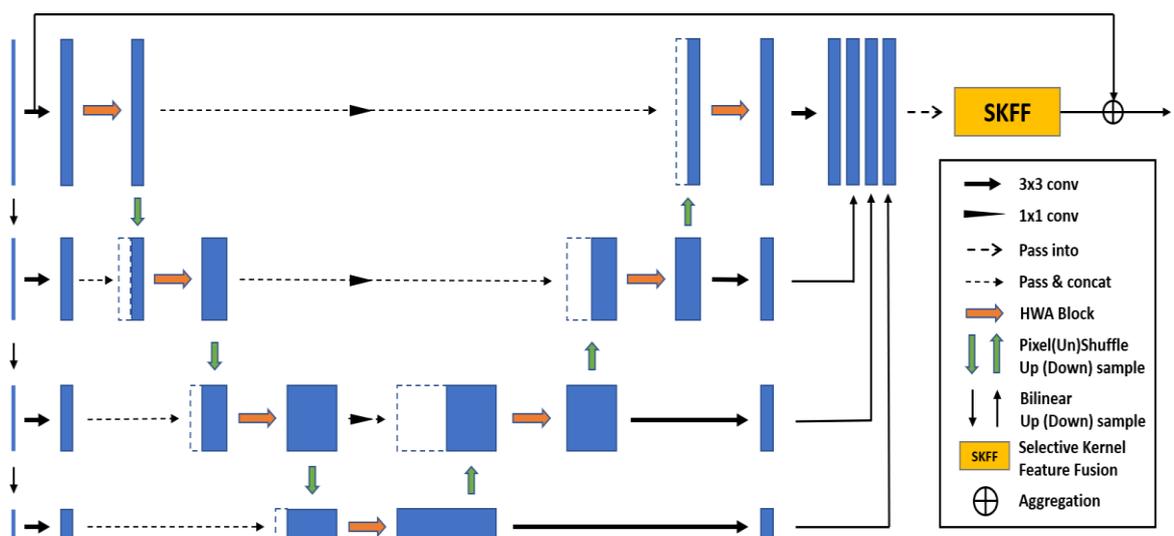


Figure II. 3: Architecture proposée.

II.3.2.1 Encodeur

L'encodeur suit un chemin d'encodage composé de blocs convolutionnels, où chaque niveau de l'encodeur intègre un Bloc d'Attention à Demi-Ondelette (HWAB : Half Wavelet Attention Block) [31], tel qu'illustré à la figure II.4. De plus, l'opération de sous-échantillonnage par pixel un-shuffle est employée pour réduire la taille des caractéristiques. Cette approche permet de capter efficacement les informations à différentes échelles tout en minimisant la perte d'information, un avantage crucial pour des applications avancées en traitement d'image et en apprentissage profond [31].

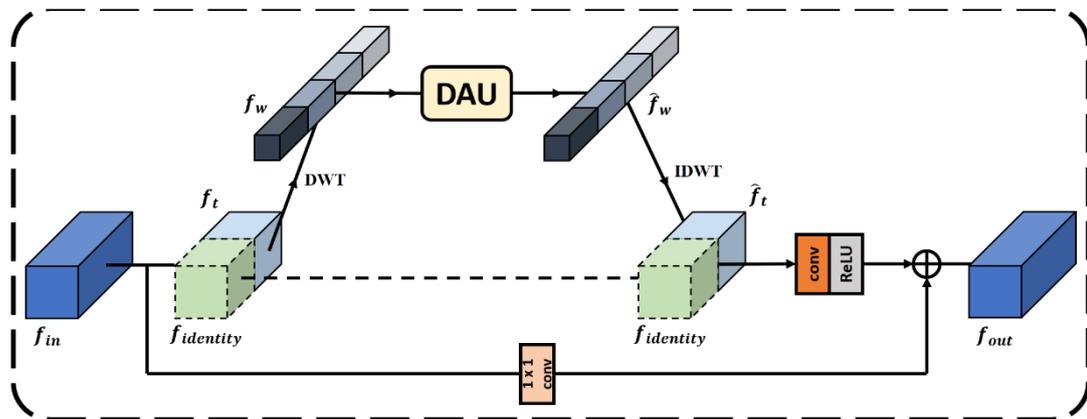


Figure II. 4: Half Wavelet Attention Block (HWAB).

-Bloc d'attention à demi- ondelette (HWAB : Half Wavelet Attention Block) : Le bloc HWAB combine l'extraction de caractéristiques spatiales et fréquentielles, utilisant des transformations par ondelettes et des mécanismes d'attention pour améliorer la représentation dans les réseaux neuronaux. Il reçoit des entrées à partir de plusieurs flux de résolution contenant des caractéristiques à diverses échelles, capturant différents niveaux de détails et de contexte. Chaque caractéristique d'entrée est divisée en deux parties : $f_{identity}$ conserve les caractéristiques normales du domaine spatial, tandis que f_t subit une transformation en ondelettes discrètes (DWT) pour obtenir des caractéristiques dans le domaine des ondelettes. Ces caractéristiques sont ensuite traitées par une unité d'attention double (DAU), permettant de focaliser sur les informations pertinentes des différents canaux de caractéristiques et de l'information spatiale cruciale.

Ce processus de division des caractéristiques d'entrée vise à réduire la complexité de calcul tout en préservant les informations contextuelles [31]. La partie $f_{identity}$ conserve les caractéristiques normales du domaine spatial, tandis que f_i subit une transformation en ondelettes discrètes (DWT) pour obtenir des caractéristiques dans le domaine des ondelettes. Cette transformation en DWT divise f_i en quatre sous-bandes de résolution inférieure, capturant les composantes haute et basse fréquence des caractéristiques d'entrée, essentielles pour une représentation détaillée de l'image. Les sous-bandes résultantes sont ensuite traitées par une Unité d'Attention Double (DAU) [32], illustré dans la figure II.5, qui utilise à la fois une attention sur les canaux et une attention spatiale pour concentrer sur les caractéristiques pertinentes des différents canaux de caractéristiques et de l'information spatiale cruciale [32].

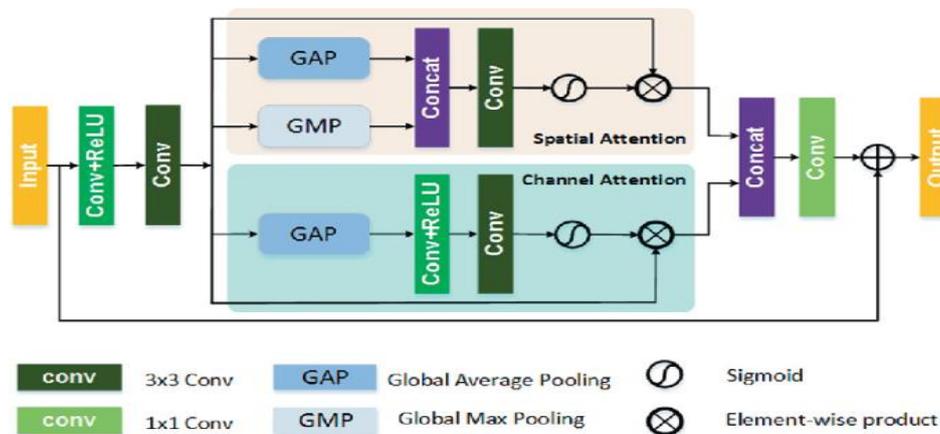


Figure II. 5: Dual attention unit (DAU).

Après ce traitement, les caractéristiques sont reconstruites à leur résolution d'origine à l'aide de la Transformée en Ondelettes Discrète Inverse (IDWT). Ensuite, les caractéristiques de sortie et $f_{identity}$ sont concaténées et passées à une couche de convolution 3×3 suivie d'une fonction d'activation ReLU. Ces caractéristiques sont ensuite ajoutées aux caractéristiques d'entrée qui ont passé par une convolution 1×1 pour obtenir des caractéristiques résiduelles, donnant ainsi les caractéristiques de sortie finales f_{out} .

II.3.2.2 Décodeur

Le décodeur est une composante essentielle permettant de reconstruire l'image de sortie en haute résolution à partir des caractéristiques compressées et extraites par l'encodeur. Sa structure est symétrique par rapport à celle de l'encodeur. Le décodeur utilise des blocs HWAB suivis par l'opération de pixel un-shuffle pour augmenter progressivement la résolution de

l'image. De plus, des connexions de saut (skip connections) intègrent les informations de l'encodeur, assurant ainsi une reconstruction précise et détaillée de l'image.

II.3.2.3 Connexions de saut (Skip Connections)

Les connexions de saut permettent de combiner les caractéristiques du décodeur avec celles correspondantes de l'encodeur. Dans l'architecture M-NET, les connexions de saut locales relient directement les couches profondes de l'encodeur aux couches correspondantes du décodeur. Grâce à ces connexions, le décodeur peut accéder aux informations provenant de différentes échelles spatiales extraites par l'encodeur. Ces voies améliorent le contexte multi-résolution de la reconstruction tout en maintenant les caractéristiques spatiales détaillées. En outre, des connexions de saut globales sont utilisées pour additionner l'image d'entrée avec l'image de sortie, afin de préserver les détails qui pourraient être perdus lors du processus de traitement avec M-NET. Cette technique assure une conservation optimale des informations essentielles, permettant une reconstruction fidèle et précise de l'image.

II.3.2.4 Pieds (Legs) (gauche et droite)

Dans l'architecture M-Net, les pides gauche et droit jouent un rôle essentiel dans le traitement et la reconstruction des images en conservant les informations à différentes résolutions. La leg gauche (Encodage) est responsable de la réduction progressive de la résolution de l'image d'entrée tout en capturant des caractéristiques importantes à chaque niveau. Elle opère sur une image d'entrée de taille $H \times W$ et utilise quatre couches de max-pooling de taille 2×2 pour réaliser cette tâche. Les sorties de ces couches de max-pooling sont ensuite fournies en entrée aux couches d'encodage correspondantes, ce qui permet d'extraire des caractéristiques à différentes échelles. La leg droite sur-échantillonne les sorties de chaque couche de décodage à la taille d'origine $H \times W$. Cela permet de réintégrer les caractéristiques extraites à différentes résolutions et de reconstruire l'image finale en haute résolution puis traiter par le block SKFF pour fusionner les caractéristiques.

II.3.2.5 Bloc de fusion des caractéristiques (SKFF)

Comme illustré dans la Figure II.6, Selective Kernel Feature Fusion (SKFF) est une méthode consiste à fusionner de manière adaptative les caractéristiques extraites par différentes branches d'un réseau de neurones, principalement dans le domaine du traitement d'images. Au lieu d'une simple concaténation, le SKFF apprend à projeter ces caractéristiques dans un espace

commun à l'aide de noyaux de convolution spécifiques à chaque branche, puis à les combiner par une somme pondérée où les poids sont également appris. De cette manière, le SKFF sélectionne et combine de manière astucieuse les informations pertinentes issues de différentes sources, en accordant une plus grande importance aux caractéristiques utiles pour la tâche visée. Il a été initialement proposé pour l'amélioration d'images et permet généralement d'obtenir de meilleures performances que des fusions naïves pour diverses tâches de vision par ordinateur [32].

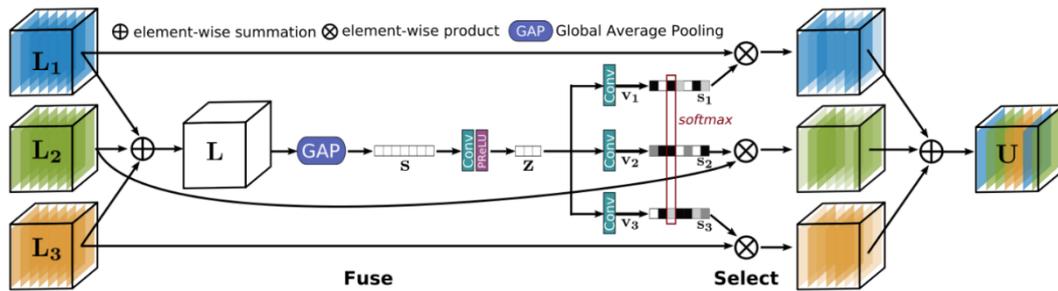


Figure II. 6: Selective kernel feature fusion (SKFF).

-Fusion (Fuse) : Le module SKFF reçoit des entrées de 4 flux de convolution parallèles portant des informations à différentes échelles. Nous combinons d'abord ces caractéristiques multi-échelles en utilisant une somme par élément comme suit :

$$(L = L1 + L2 + L3 + L4) \tag{II. 1}$$

Ensuite, nous appliquons un pooling moyen global (GAP) sur les dimensions spatiales de $L \in \mathbb{R}^{H \times W \times C}$ pour calculer les statistiques par canal $s \in \mathbb{R}^{1 \times 1 \times C}$. Ensuite, nous appliquons une couche de convolution pour réduire le nombre de canaux afin de générer une représentation compacte des caractéristiques $z \in \mathbb{R}^{1 \times 1 \times r}$, où $r = C/8$. Enfin, le vecteur de caractéristiques z passe à travers 4 couches de convolution parallèles (une pour chaque flux de résolution) pour générer 4 descripteurs de caractéristiques $v1, v2, v3$ et $v4$, chacun ayant des dimensions $1 \times 1 \times C$ [32].

-Sélectionner (Select) : Cet opérateur applique la fonction softmax à $v1, v2, v3$ et $v4$, ce qui produit des activations d'attention $s1, s2, s3$ et $s4$ que nous utilisons pour recalibrer de manière adaptative les cartes de caractéristiques multi-échelles $L1, L2, L3$ et $L4$ respectivement [32]. Le processus global de recalibrage et d'agrégation des caractéristiques est défini comme suit :

$$U = s1 \times L1 + s2 \times L2 + s3 \times L3 + s4 \times L4 \tag{II. 2}$$

II.3.2.6 Pixel shuffle et Unshuffle

Les opérations de pixel shuffle et pixel unshuffle sont utilisées pour adapter la résolution des caractéristiques à différentes échelles du réseau.

-Pixel Shuffle [33]: Le pixel shuffle est une opération de réarrangement des pixels qui transforme une image de basse résolution avec un nombre de canaux élevé en une image de haute résolution avec moins de canaux. Cette transformation est essentielle pour produire des détails fins tout en préservant les informations utiles encodées dans les canaux supplémentaires. Le processus consiste à réorganiser les canaux de l'image d'entrée en une nouvelle disposition où les pixels sont répartis de manière à augmenter la résolution spatiale.

-Pixel Unshuffle [33]: Le pixel unshuffle est l'opération inverse du pixel shuffle. Elle convertit une image de haute résolution avec moins de canaux en une image de basse résolution avec plus de canaux, en compressant les informations spatiales. Le processus consiste à regrouper les pixels de l'image haute résolution en blocs et à les réarranger en un ensemble de canaux supplémentaires dans une image de basse résolution.

II.4 Discriminateur

Le discriminateur est utilisé pour classer les images obtenues par le générateur comme vrais (réels) ou faux (générés) (voir la figure II.7). Le modèle discriminateur prend un exemple du domaine comme entrée (réelle ou générée) et prédit une étiquette de classe binaire de réel ou faux (généré) [28]. L'image réelle provient de l'ensemble de données d'apprentissage. Les exemples générés sont générés par le modèle de générateur.

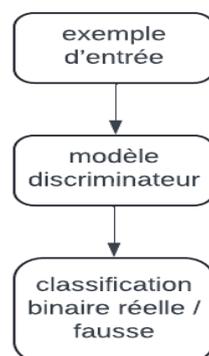


Figure II. 7: Exemple de modèle du Discriminateur.

Dans notre approche, nous avons adopté un discriminateur Markovien (PatchGAN) [34]. PatchGAN est un type de discriminateur pour les réseaux antagonistes génératifs qui ne pénalise la structure qu'à l'échelle des patches d'images locales. Le discriminateur PatchGAN tente de classer si chaque $N \times N$ patch dans une image est réel ou faux [35], comme illustré dans la figure II.8.

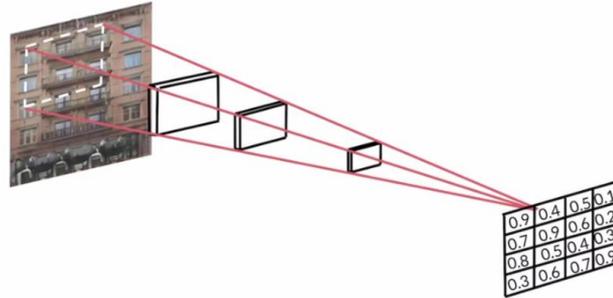


Figure II. 8: Discriminateur Markovien (PatchGAN).

Ce discriminateur impose des contraintes strictes pour préserver les détails nets à haute fréquence. Le PatchGAN est avantageux car il possède moins de paramètres et est plus rapide que les approches qui classifient l'image entière. Le discriminateur, illustré dans la Figure II.9, se compose de 5 blocs convolutionnels. Chaque bloc utilise des filtres de convolution 2D de taille 4x4 avec un pas de 2x2, suivis d'une normalisation par lots et d'une activation Leaky ReLU, à l'exception des premiers et derniers blocs. Le premier bloc utilise des filtres de taille 3x3 avec un pas de 2x2 et une activation Leaky ReLU. Le dernier bloc utilise des filtres de taille 4x4 avec un pas de 1x1, une normalisation par lots et une activation Leaky ReLU. La sortie du discriminateur est une matrice 2D de taille 20x20x1.

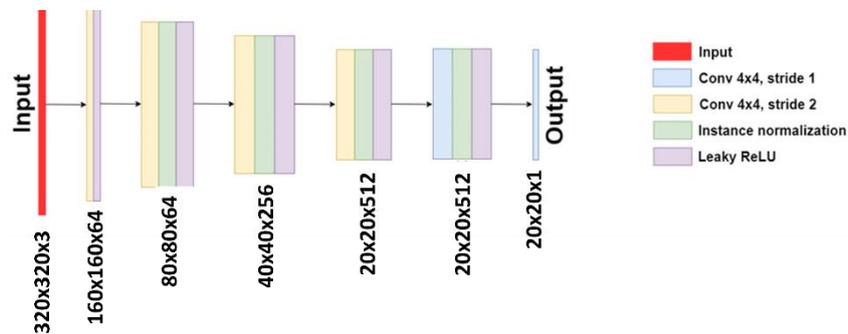


Figure II. 9: Architecture du discriminateur Markovien (PatchGAN).

II.5 Fonctions de perte (Loss function)

Nous employons des fonctions de perte spécifiques et adaptées à notre méthode. Nous décrivons ces fonctions de perte ci-dessous.

II.5.1 Perte du générateur et du discriminateur :

Nous utilisons le discriminateur basé sur le GAN moyen relativiste (RaGAN). Contrairement au discriminateur standard, qui estime la probabilité que les données réelles soient en moyenne plus réalistes que les données générées [35], les fonctions de perte pour le discriminateur D et le générateur G sont exprimées comme suit :

$$L_D^{RaGAN} = \mathbb{E}_{x_r \sim \mathbb{P}} [f_1 (C(x_r) - \mathbb{E}_{x_f \sim \mathbb{Q}} C(x_f))] + \mathbb{E}_{x_f \sim \mathbb{Q}} [f_2 (C(x_f) - \mathbb{E}_{x_r \sim \mathbb{P}} C(x_r))] \quad (\text{II. 3})$$

$$L_G^{RaGAN} = \mathbb{E}_{x_r \sim \mathbb{P}} [g_1 (C(x_r) - \mathbb{E}_{x_f \sim \mathbb{Q}} C(x_f))] + \mathbb{E}_{x_f \sim \mathbb{Q}} [g_2 (C(x_f) - \mathbb{E}_{x_r \sim \mathbb{P}} C(x_r))] \quad (\text{II. 4})$$

Où f_1, f_2, g_1 et g_2 sont des fonctions scalaire-à-scalaire. x_r, x_f sont respectivement les données réelles et générées.

II.5.2 Perte de Charbonnier

La fonction de perte de Charbonnier [31] est une fonction de perte employée dans diverses opérations de traitement d'images, telles que le débruitage d'archives. Il s'agit d'une forme lisse de la perte de Huber, développée pour être moins sensible aux valeurs aberrantes que la perte de l'erreur quadratique moyenne.

$$l_{Char} = \sqrt{(x_r + x_f)^2 + \epsilon^2} \quad (\text{II. 5})$$

Où x_r, x_f sont respectivement les données réelles et générées respectivement. La constante ϵ est fixée à 10^{-3} .

II.5.3 Fonction de Perte Perceptuelle

La fonction de perte de perception [36] a pour objectif de détecter les disparités de perception entre les images, comme les différences de contenu et de style, qui ne sont pas toujours visibles au niveau des pixels. Ils sont fréquemment utilisés dans des activités visant à produire des images qui sont agréables visuellement pour les êtres humains, telles que le

transfert de style neuronal, la super-résolution et la synthèse d'images. Les connexions à destination locale. Elle est définie comme suit :

$$l_{per} = \|\varphi(x_r) - \varphi(x_f)\|_2^2 \quad (\text{II. 6})$$

Où ϕ est les représentations des caractéristiques extraites par le VGG19.

II.6 Conclusion

Dans ce deuxième chapitre, nous avons exposé de manière approfondie l'architecture et les caractéristiques techniques de notre projet, notamment l'architecture M-Net améliorée pour le générateur et le Patch GAN pour le discriminateur. Nous avons également exploité différents blocs utilisés dans notre approche, tels que le bloc de fusion des caractéristiques (SKFF) et le bloc d'attention à demi-ondelette (HWAB), ce qui permet à notre modèle de capturer et de restaurer des détails fins et des textures complexes. Dans notre projet, nous avons utilisé différentes fonctions de perte pour renforcer la performance de notre modèle en termes de fidélité et de réalisme des images reconstruites.

Chapitre III :

Résultats et Implémentations

III.1 Introduction

Ce chapitre est entièrement dédié à la présentation des résultats de notre système. Notre approche consiste à entraîner un modèle de réseau de neurones convolutifs basé sur les GAN (réseaux antagonistes génératifs) dans le but de déconvoluer les images floues. Plusieurs métriques de qualité sont utilisées pour évaluer les performances de notre système, telles que les métriques de qualité avec référence et les métriques de qualité sans référence. Nous avons également employé diverses méthodes pour comparer les performances du système.

III.2 Outils utilisés

- **Langage de programmation (python):** Python est un langage de programmation orienté objet de haut niveau avec une syntaxe simple et élégante. Il est suffisamment puissant pour les applications complexes, allant du développement web à l'analyse de données, et il dispose d'une vaste bibliothèque standard ainsi que de nombreuses bibliothèques tierces, ce qui le rend accessible et largement adopté. Compatible avec tous les types de systèmes, Python favorise l'innovation et la croissance rapide des projets. Ce langage est particulièrement prisé pour l'apprentissage automatique (Machine Learning) et la science des données (Data Science) [37].
- **Pytorch :** PyTorch est un projet open source axé sur l'apprentissage automatique, une bibliothèque dérivée de Python développée par le laboratoire de recherche en IA de Facebook (FAIR). Il offre une interface intuitive en Python et une alternative C++ pour des performances accrues. PyTorch fournit une plateforme simple et efficace pour créer et former des modèles de deep learning. Sa conception modulaire et sa syntaxe en font un outil idéal pour les chercheurs dans les disciplines de l'IA. La flexibilité et l'ouverture de PyTorch en ont fait un choix populaire pour les chercheurs et les ingénieurs travaillant sur les applications d'intelligence artificielle [38].

III.3 Datasets

Dans notre projet, nous avons utilisé le dataset GoPro pour la déconvolution des images [39], qui est couramment utilisé dans la recherche en traitement d'image et en vision par ordinateur. Il se compose de 3214 images floues avec une résolution de 1280x720 pixels. Ces images sont divisées en 2103 images d'entraînement et 1111 images de test. Chaque image floue

a une image de vérité terrain nette correspondante obtenue à l'aide d'une caméra à haute vitesse. Ce dataset est précieux pour développer et évaluer des algorithmes de déconvolution.

III.4 Détails de l'implémentation

Nous avons implémenté nos architectures en utilisant le framework PyTorch sur NVIDIA Quadro RTX 6000. Nos modèles sont entraînés à partir de zéro pendant 100 époques avec un taux d'apprentissage (learning rate) de $1e-4$. Nous utilisons l'optimiseur Adam (Adam optimizer) et la taille de lot (batch size) est fixée à 10. Pendant la phase d'entraînement, nous découpons aléatoirement les images à une taille de 320×320 .

III.5 Métriques d'évaluation de la qualité de l'image

Les métriques de qualité d'image peuvent être divisées en deux grandes catégories : les métriques à référence complète (Full-Reference, FR) et les métriques sans référence (No-Reference, NR). Chacune de ces catégories mesure différents aspects de la qualité d'une image, en fonction de la disponibilité d'une image de référence.

❖ Métriques de qualité avec référence :

- **Indice de similarité structurale (SSIM)**: SSIM est une métrique de qualité d'image qui compare deux images pour déterminer leur similarité visuelle, il compare la luminance, le contraste et la structure entre l'image originale et l'image restaurée [40]. Un score élevé indique une grande similarité, plus alignée avec la perception humaine.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3.1)$$

- μ_x et μ_y sont les moyennes des images x et y .
- σ_x^2 et σ_y^2 sont les variances des images x et y .
- σ_{xy} est la covariance entre x et y .
- $C_1 = (k_1L)^2$ et $C_2 = (k_2L)^2$ sont deux constantes de stabilisation pour éviter la division par zéro, où L est la plage dynamique des valeurs de pixel et k_1 et k_2 sont des constantes par défaut (typiquement $k_1=0.01$ et $k_2=0.03$).

- **Erreur quadratique moyenne (MSE) :** MSE mesure l'erreur quadratique moyenne entre les pixels de l'image originale et de l'image restaurée [41]. Une valeur faible est désirée pour une meilleure compression. MSE est donnée par l'équation ci-dessous :

$$\text{MSE} = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (f(i, j) - g(i, j))^2 \quad (3.2)$$

- $f(i,j)$: l'image originale de dimension $M \times N$.
- $g(i,j)$: l'image restaurée de dimension $M \times N$.

- **Peak Signal-to-Noise Ratio (PSNR):** Le PSNR mesure le rapport entre la valeur maximale possible d'un pixel et l'erreur quadratique moyenne [42]. Une valeur élevée indique une bonne qualité d'image débruitée. Le PSNR est régulièrement exprimé en termes d'échelle logarithmique en décibels et est donné par l'équation suivante :

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right) \quad (3.3)$$

Où MAX est la valeur maximale possible d'un pixel de l'image.

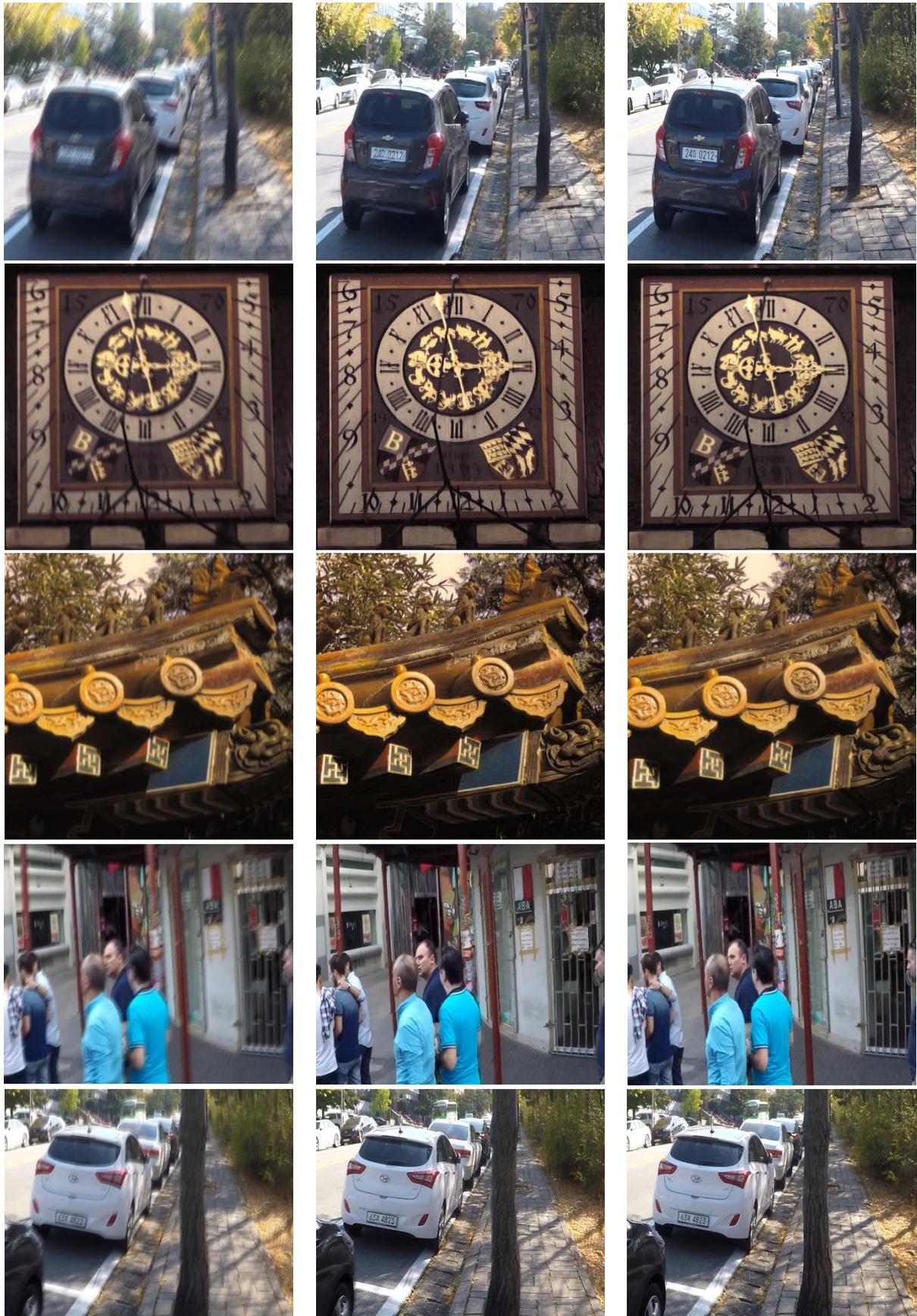
❖ Métriques de qualité sans référence :

- **Évaluateur de qualité spatiale d'image aveugle/sans référence (BRISQUE):** BRISQUE est une métrique de qualité d'image sans référence qui évalue la qualité en analysant les caractéristiques spatiales naturelles des patchs locaux d'une image. Elle modélise les déviations de ces caractéristiques par rapport à des modèles statistiques. Un score faible indique une meilleure qualité d'image [43].
- **Évaluateur de qualité d'image basé sur la perception (PIQE) :** PIQE est une métrique de qualité d'image sans référence qui évalue la qualité perçue des images en analysant les blocs locaux de l'image pour détecter des artefacts tels que le flou et le bruit, en se basant sur la variance des gradients. Un score faible indique une meilleure qualité perçue [43].
- **Évaluateur de qualité d'image naturelle (NIQE) :** Le NIQE est une métrique sans référence qui évalue la qualité d'une image en se basant sur des caractéristiques statistiques naturelles extraites de patchs locaux. Un score faible indique une meilleure qualité [44].

III.6 Résultats et discussion

- **Comparaison qualitative :** nous procédons à une comparaison qualitative entre une image floue, l'image restaurée par notre système, et l'image réelle correspondante. Cette analyse vise à évaluer la capacité de notre architecture à restaurer fidèlement les détails et à se rapprocher visuellement des images réelles de haute qualité.

Comme illustré dans la Figure 3.1, les résultats obtenus grâce à notre approche démontrent la capacité de notre système à produire des performances compétitives dans le domaine de la déconvolution d'images. Notre méthode permet une amélioration significative de la visibilité des détails fins ainsi qu'une restauration efficace des caractéristiques visuelles initialement altérées par le flou. Elle assure également la reconstruction précise des contours et des textures des objets, rétablissant ainsi la clarté et la netteté des scènes capturées.

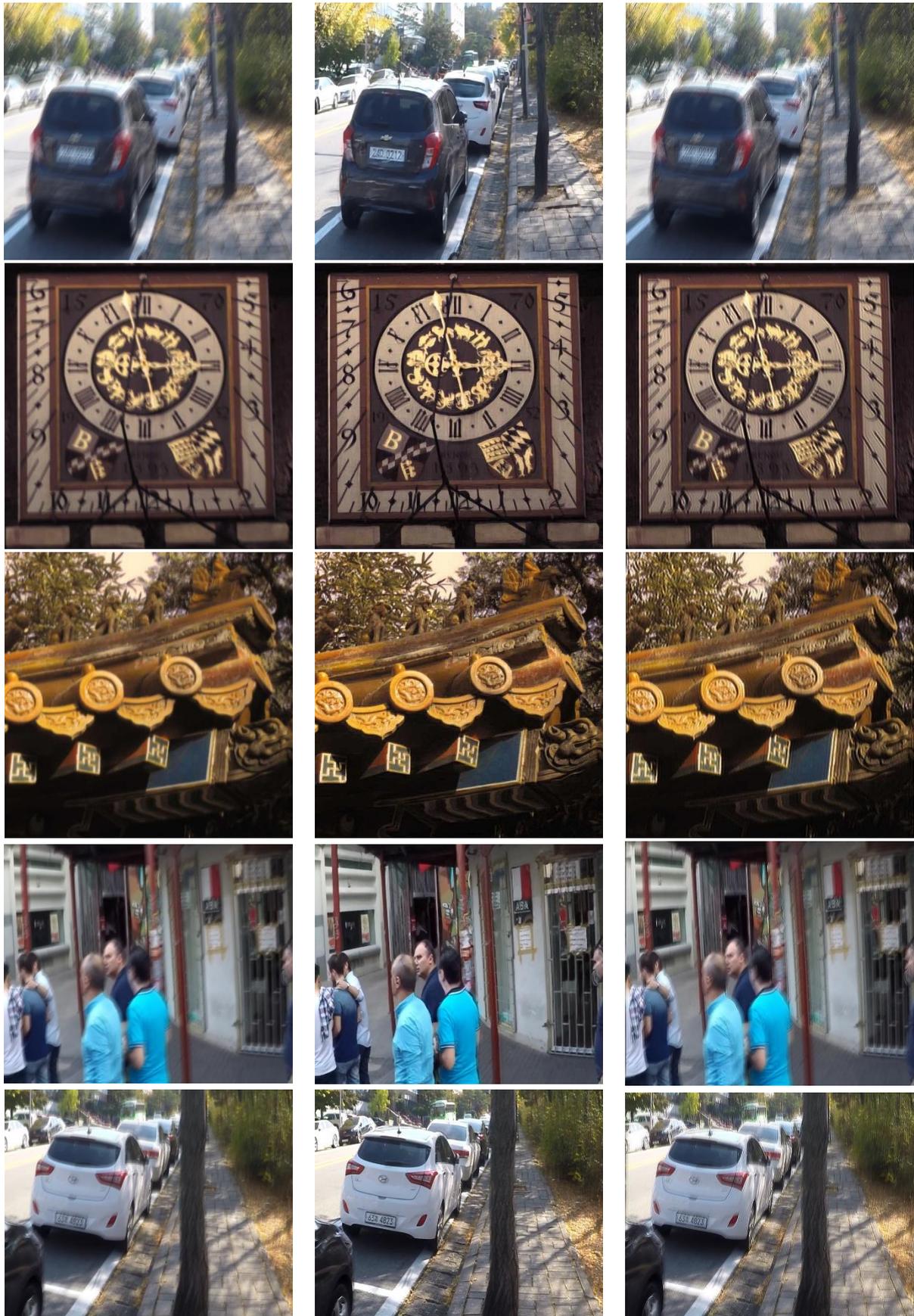


a) Image flou

b) Image restaurée

c) Image réelle

Figure III.1 : Résultats de notre système.

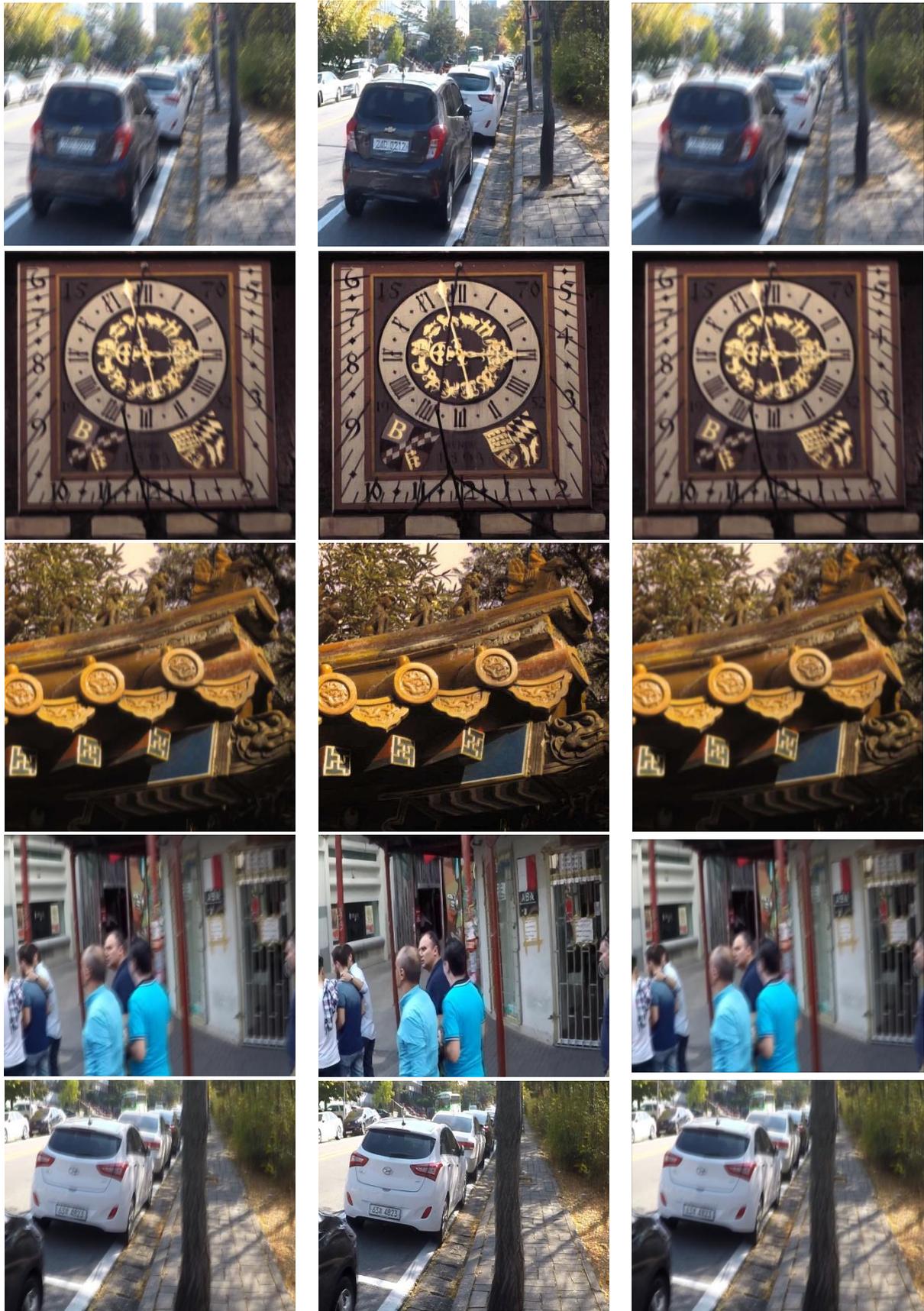


a) Image flou

b) Our

c) Déconvolution par aveugle

Figure III.2 : Comparaison qualitative des résultats obtenus avec la déconvolution par aveugle.

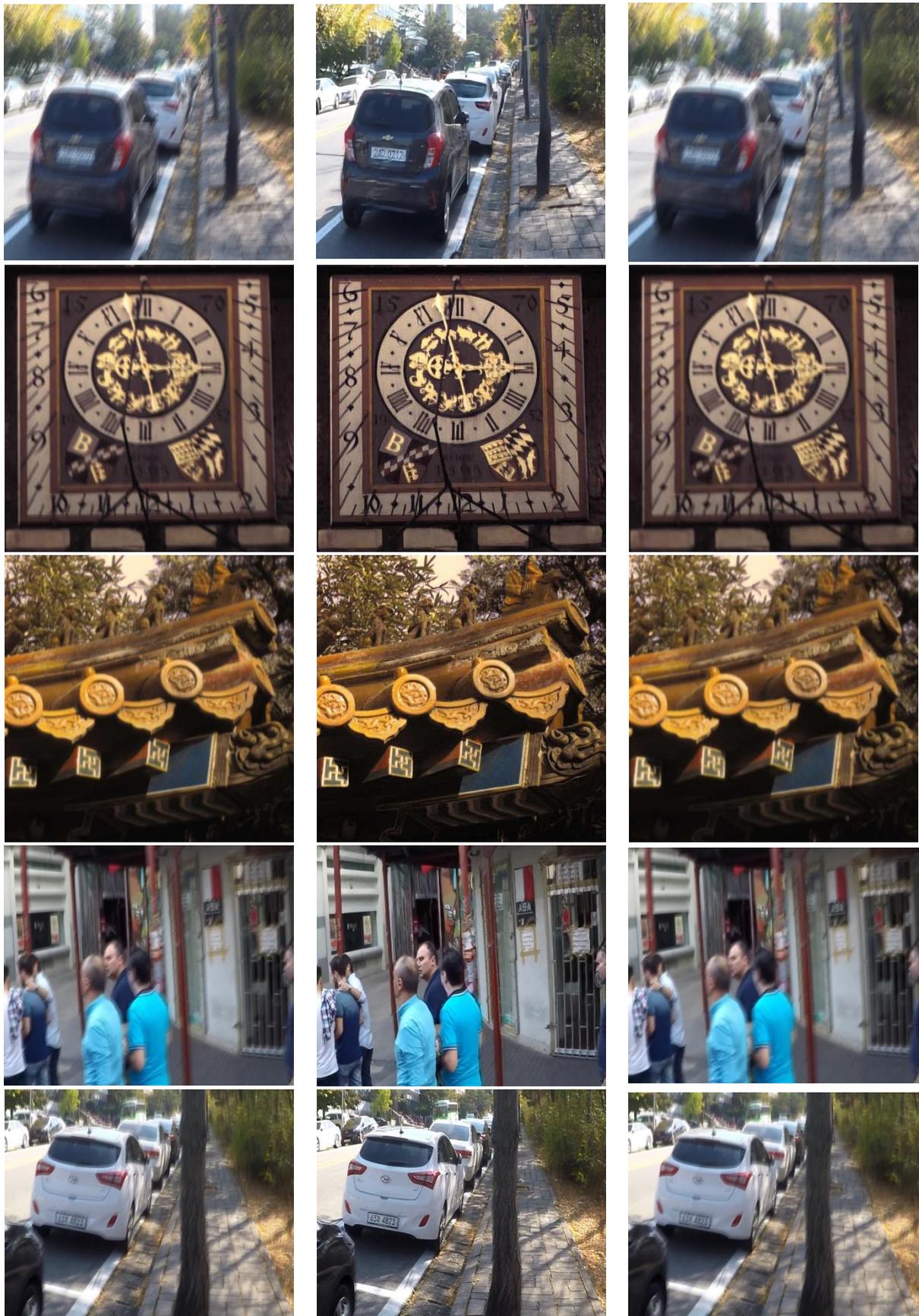


d) Image flou

e) Our

f) Filtre Gaussien

Figure III.3 : Comparaison qualitative des résultats obtenus avec la déconvolution par le filtre de Gaussien.



a) Image flou

b) Our

c) Filtre wiener

Figure III.4 : Comparaison qualitative des résultats obtenus avec la déconvolution par le filtre de wiener.

Notre approche est comparée à trois méthodes traditionnelles telles que la déconvolution par aveugle (figure III.2), le filtre gaussien (figure III.3) et le filtre Wiener (figure III.4).

Les résultats illustrés dans les figures III.2, III.3 et III.4 mettent en évidence les limitations des méthodes traditionnelles en termes de la déconvolution des images et de restauration de leur qualité visuelle, notamment en ce qui concerne la visibilité et la restauration des détails.

➤ **Comparaison quantitative :**

Tableau III.1 : Comparaison quantitative des résultats sur des ensembles de données. (+) indique que le score le plus élevé est meilleur, (-) indique le score le plus bas est meilleur. Les deux meilleurs résultats sont en rouge et bleu respectivement.

| Méthode | MSE- | PSNR+ | SSIM+ | NIQE- | PIQE- | BRISQUE- |
|---------------------------|--------|---------|--------|--------|---------|----------|
| Déconvolution par aveugle | 0.0177 | 66.6355 | 0.5049 | 4.5134 | 48.2923 | 37.1766 |
| Filtre gaussien | 0.0150 | 67.6900 | 0.6000 | 5.1208 | 84.3682 | 54.9319 |
| Filtre Wiener | 0.0113 | 59.5600 | 0.6285 | 5.2338 | 80.4488 | 54.6586 |
| Notre Méthode | 0.0086 | 71.4871 | 0.7468 | 3.3425 | 37.7420 | 31.5014 |

Les résultats illustrés dans le tableau III.1 montrent que notre méthode surpasse significativement les méthodes traditionnelles (déconvolution par aveugle, filtre gaussien, filtre Wiener) dans toutes les métriques évaluées. Cela souligne son efficacité supérieure pour la déconvolution des images, aussi bien en termes de précision technique que de qualité visuelle perçue.

➤ **Comparaison globale :**

Notre étude comparative repose sur une évaluation quantitative et qualitative, ce qui nous permet de confirmer l'avantage significatif de notre méthode exploitant les techniques avancées de deep learning par rapport aux approches traditionnelles pour la déconvolution des images. En analysant les données à travers une série de métriques standard telles que le MSE, le PSNR, le SSIM, le NIQE, le PIQE et le BRISQUE, nous avons observé de manière cohérente que notre méthode surpassait ses concurrentes. Notre méthode a démontré une capacité supérieure à reconstruire les images avec une précision plus proche de l'originale ainsi que garantissant ainsi une restauration plus fidèle des détails et des contours, préserver les textures et les contours des objets de manière nettement améliorée par rapport aux méthodes traditionnelles.

III.7 Conclusion

Dans ce chapitre, nous avons présenté et analysé les résultats obtenus à partir de notre système comparativement à d'autres méthodes existantes. Les résultats démontrent clairement que notre architecture parvient à restaurer les images floues avec une fidélité et des détails comparables, voire supérieurs, à ceux des autres méthodes, se rapprochant significativement des images réelles de haute qualité. L'évaluation des performances de notre système s'est appuyée sur une variété de métriques et a été conduite sur divers types d'images, illustrant ainsi sa robustesse et sa fiabilité dans des conditions variées. Ces analyses détaillées ont permis de mettre en évidence les avantages de notre approche, notamment sa capacité à améliorer la netteté et à préserver les détails essentiels des images restaurées. En conclusion, notre étude apporte une contribution significative au domaine du traitement d'images, en proposant une solution efficace et prometteuse pour la restauration d'images floues, avec des implications potentielles dans divers secteurs nécessitant des standards élevés de qualité visuelle.

Conclusion générale

Conclusion générale

Ce projet de recherche met en évidence l'efficacité remarquable des réseaux antagonistes génératifs (GANs) combinés à d'autres architectures de réseaux neuronaux pour la déconvolution d'images floues. En surpassant les méthodes traditionnelles, cette approche se distingue par sa capacité à améliorer significativement la netteté et la qualité des images traitées, les rapprochant ainsi des originaux de haute qualité. Cette avancée technologique ouvre de nouvelles perspectives dans des secteurs critiques comme la photographie et l'imagerie médicale, où la précision et la clarté des images sont essentielles pour des applications telles que le diagnostic médical précis et la surveillance détaillée.

Les résultats obtenus démontrent non seulement les performances supérieures des GANs dans la restauration d'images dégradées par le flou, mais également leur flexibilité à traiter différents types de flou. Cette polyvalence est rendue possible par la capacité des GANs à apprendre directement à partir des données et à générer des images réalistes et détaillées. En intégrant ces capacités avec d'autres architectures de réseaux neuronaux, il devient possible de développer des modèles robustes capables de s'adapter aux conditions variables de captation d'image et aux besoins spécifiques des applications.

En conclusion, cette étude met en évidence la capacité significative des techniques d'apprentissage profond à relever les défis complexes du traitement d'image. Ces avancées ne se contentent pas d'améliorer la qualité des images numériques ; elles révolutionnent également nos approches de capture, d'analyse et d'exploitation de l'information visuelle à travers une diversité de domaines professionnels et scientifiques.

Références

Références

- [1] <https://fr.wikipedia.org/wiki/Image>.consulter le 17/05/2024.
- [2] <https://www.encyclopedia.com/computing/news-wires-white-papers-and-books/digital-images>.consulter le 17/05/2024.
- [3] K. KHAMMAR, «Robot mobile avec suivi de ligne,»,Mémoire Master 2 Biskra, 2013.
- [4] <https://www.larousse.fr/dictionnaires/francais/flou/34248#locution>. consulter le 11/05/2024.
- [5] <https://ressources.univlemans.fr/AccesLibre/UM/Pedago/physique/02b/coursoptique/abergeo.html>. consulter le 18/05/2024.
- [6] <https://www.shopmoment.com/reviews/how-to-create-motion-blur-with-the-moment-camera-app>. consulter le 18/05/2024.
- [7] <https://www.nicephore.ch/six-facons-de-jouer-avec-le-flou/>. consulter le 18/05/2024.
- [8] <https://photographylife.com/what-is-depth-of-field>. consulter le 18/05/2024.
- [9] <https://www.designworldonline.com/understanding-accuracy-and-precision-for-mems-pressure-sensors/>. consulter le 19/05/2024.
- [10] <https://www.1point2vue.com>. Consulter le 10/06/2024.
- [11] <https://phlearn.com/magazine/managing-digital-noise-good-bad-ugly/>. consulter le 16/05/2024.
- [12] A. Branco, «Photo : qu'est-ce que le bruit numérique ?,» 19 novembre 2014. <https://www.01net.com/astuces/quest-ce-que-le-bruit-numerique-632881.html>. [Accès le 19 06 2024].
- [13] <https://imagify.io/blog/lossless-vs-lossy-image-compression/>. consulter le 19/05/2024.
- [14] <https://www.photoshopessentials.com/essentials/resizing-vs-resampling/>. consulter le 19/05/2024.
- [15] <https://www.educative.io/answers/what-is-image-blurring>.consulter le 15/05/2024.
- [16] https://en.wikipedia.org/wiki/Gaussian_blur. consulter le 20/05/2024.
- [17] <https://docs.gimp.org/2.10/en/gimp-filter-median-blur.html>. consulter le 20/05/2024.
- [18] https://fr.wikipedia.org/wiki/Flou_cin%C3%A9matographique.consulter le 16/05/2019.
- [19] B. Perret., 2017. <https://perso.esiee.fr/~perretb/I5FM/TAI/convolution/index.html#id3>.

- [20] G. S. -. N. Roumaissa., «Image from Image Restoration by Filtering (Perona-Malik Model)».
- [21] P. White, «Transforms, Wavelets,» chez *Encyclopedia of Vibration*, 2001, pp. 1419-1435.
- [22] M. N. Abderrahmane, «intelligence Artificiel,» Université Saad Dahleb BLIDA, 2022.
- [23] <https://www.ibm.com/topics/machine-learning>.consulter le 31/05/2024.
- [24] <https://datascientest.com/deep-learning-definition>. consulter le 10/06/2024..
- [25] <https://medium.com/analytics-vidhya/what-is-unet-157314c87634>. consulter le 9/5/2024.
- [26] <https://datascientest.com/U-net>.consulter le 12/5/2024.
- [27] <https://towardsdatascience.com/understanding-semantic-segmentation-with-unet-6be4f42d4b47>.consulter le 10/5/2024.
- [28] J. Brownlee, *Generative Adversarial Networks with Python,deep learning generative models for*, 2019.
- [29] Raghav Mehta and Jayanthi Sivaswamy, «“M-net: A convolutional neural network for deep brain structure segmentation,» *IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017).*,» p. pp. 437–440., 2017.
- [30] P. F. a. T. B. Olaf Ronneberger, «“Unet: Convolutional networks for biomedical image segmentation,»in *International Conference on Medical image computing and computer-assisted,*» p. pp. 234–241., 2015, .
- [31] C.-M. T.-J. L. a. K.-H. L. Fan, «"Half wavelet attention on M-Net+ for low-light image enhancement.",» *2022 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2022..
- [32] A. A. S. K. M. H. F. S. K. M.-H. Syed Waqas Zamir, «“Learning enriched features for real image restoration and enhancement,»» pp. Part XXV 16. Springer, 2020, pp. 492–511., *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020.*,.
- [33] J. C. F. H. J. T. A. P. A. R. B. ´. D. R. Z. W. W. Shi, «Real-time single image and video superresolution using an efficient sub-pixel convolutional neural network,» pp. *Proceedings of the IEEE conference on computer vision and pattern recognition*,1874–1883,, 2016..

- [34] P. e. a. Isola, «"Image-to-image translation with conditional adversarial networks."», Proceedings of the IEEE conference on computer vision and pattern recognition. 2017..
- [35] A. Jolicoeur-Martineau, «"The relativistic discriminator: a key element missing from standard GAN."», arXiv preprint arXiv:1807.00734 (2018)..
- [36] C. e. a. Ledig, «"Photo-realistic single image super-resolution using a generative adversarial network."», Proceedings of the IEEE conference on computer vision and pattern recognition. 2017..
- [37] <https://www.futura-sciences.com/tech/definitions/informatique-python-19349/>.
conculter le 8/6/2024.
- [38] <https://datascientest.com/pytorch-tout-savoir.consuter le 10/06/2024>.
- [39] S. T. H. K. a. K. M. L. Nah, «"Deep multi-scale convolutional neural network for dynamic scene deblurring."», *Proceedings of the IEEE conference on computer vision and pattern* , 2017.
- [40] M. I. A. C. B. F. I. H. R. S. S. M. I. a. E. P. S. S. M. I. Zhou Wang, «Image Quality Assessment: From Error Visibility to Structural Similarity», pp. IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 13, NO. 4, , APRIL 2004 .
- [41] N. D. N. a. L. J. Karam, «A No-Reference Image Blur Metric Based on the Cumulative Probability of Blur Detection (CPBD)», *IEEE Transactions on Image Processing*, Vols. %1 sur %2vol. 20, no. 9, pp. pp. 2678-2683, 2011.
- [42] R. C. G. a. R. E. Wood, Digital Image Processing, Digital Image Processing, 2018.
- [43] A. K. M. a. A. C. B. F. I. Anish Mittal, «No-Reference Image Quality Assessment in the Spatial Domain», DECEMBER 2012.
- [44] R. S. a. A. C. B. F. I. Anish Mittal, «Making a ‘Completely Blind’ Image Quality».
- [45] «Algorithms and architectures for 2D discrete wavelet transform», *The Journal of Supercomputing Article*, vol. Volume 62, p. pages 1045–1064, 2012.