

Republique Algérienne Democratique Et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université du 8 mai 1945-Guelma-
Faculté des mathématiques, de l'informatique et des sciences de la
matière Département d'informatique



Mémoire de Master

Spécialité : Informatique
Option:

Science et technologie de l'information et de la communication

Thème

Du Visage Vers Une Image De Qualite Naturelle

Présentée par : Hind BOUBIDI

Membres du jury :

N	Nom et prénom	Qualité
1	BORDJIBA Yamina	Président
2	GUERROUI Nadia	Superviseur
3	DERDAR Salah	Examineur

soutenue le 25 juin 2023

Remerciment

Cher comité d'évaluation,

Je tiens à exprimer ma sincère gratitude pour l'opportunité qui m'a été offerte de présenter ma mémoire devant vous. Votre présence et votre expertise ont été d'une valeur inestimable dans l'évaluation de mon travail de recherche.

Je souhaite également exprimer ma profonde reconnaissance envers madame GEUROUI NADIA pour son soutien constant, ses conseils avisés et son expertise précieuse tout au long de ce processus. Votre guidance a été déterminante dans la réalisation de cette mémoire, et je vous suis sincèrement reconnaissante de votre investissement dans mon développement académique.

Je tiens à remercier chaleureusement mes amis proches, ma famille et mes proches pour leur soutien indéfectible tout au long de cette aventure. Vos encouragements, vos mots d'encouragement et votre amour ont été une source inépuisable de motivation et de confiance en moi-même.

Enfin, je tiens à exprimer ma gratitude envers toutes les personnes qui ont contribué à ma recherche, que ce soit en tant que participants, en partageant leurs connaissances ou en offrant leur soutien logistique. Votre contribution a été d'une importance capitale pour l'enrichissement de mon travail.

En conclusion, je voudrais remercier sincèrement tous ceux qui ont contribué de près ou de loin à la réalisation de cette mémoire. Votre engagement, vos encouragements et vos conseils ont fait de cette expérience une réussite et ont marqué une étape importante dans mon parcours académique.

Je vous suis profondément reconnaissante pour votre évaluation attentive et pour l'opportunité de partager mes recherches avec vous. Vos commentaires et suggestions constructifs seront d'une grande valeur pour ma croissance académique future.

Avec une gratitude sincère

Dédicace

À mon cher père,

Tes encouragements constants, ta sagesse et ton amour inconditionnel ont été les piliers de ma vie. Tu as été mon modèle et mon guide, et je te suis infiniment reconnaissante pour tout ce que tu as fait pour moi. Cette réussite est aussi la tienne.

À ma chère mère,

Ton amour inébranlable, ta patience infinie et ton dévouement sans faille ont été mes sources d'inspiration tout au long de cette aventure. Tu as été mon roc, me soutenant dans les moments difficiles. Je te suis extrêmement reconnaissant(e) pour ta présence constante et ton amour inconditionnel.

À mon petit frère,

Tu es ma fierté et ma joie de vivre. Ta présence joyeuse et tes encouragements ont illuminé mon chemin tout au long de cette thèse. Merci d'avoir été mon soutien indéfectible et mon compagnon de route. Je suis fière d'être ton grand frère/grande sœur.

À ma précieuse copine Karima,

Ton amour, ton soutien inconditionnel et ta confiance en moi m'ont donné la force et le courage de persévérer. Tu as été mon rocher, mon épaule sur laquelle je pouvais m'appuyer. Merci d'avoir été là pour moi à chaque étape de cette thèse. Notre amour est ma plus grande inspiration.

À mes chères amies Hiba, Wiam, Majda et Zahra,

Vos sourires, votre présence et votre soutien indéfectible ont illuminé ma vie. Vous avez été mes piliers, mes confidents et mes compagnons de voyage. Vos encouragements et votre amitié sincère m'ont donné la force de surmonter les défis. Je suis honorée de vous avoir à mes côtés.

À Loufi,

je tiens à te souhaiter une continuité remarquable et une multitude de bonheur dans ta vie. Ta présence illumine nos journées de travail et ta positivité est contagieuse. Que chaque nouvelle étape te conduise vers des succès et des réalisations extraordinaires. Je suis convaincue que ton avenir sera brillant. Que tes rêves se concrétisent et que chaque jour t'apporte de nouvelles opportunités de croissance et d'épanouissement. Tu mérites tout le bonheur du monde. Continue de rayonner et de réaliser tes ambitions.

À vous tous, mes êtres chers,

Cette thèse n'aurait pas été possible sans votre amour, votre soutien et votre confiance en moi. Vos encouragements constants, vos mots d'encouragement et votre présence ont été des cadeaux inestimables. Je vous suis profondément reconnaissante pour votre impact positif dans ma vie et dans la réalisation de cette thèse.

Avec tout mon amour et ma gratitude,

Hind

من الوجه إلى صور طبيعية عالية الجودة

الملخص

تركز هذه الأطروحة على تطوير نظام تحسين دقة صور الوجوه، بهدف إعادة بناء الصور ذات الدقة المنخفضة إلى صور عالية الدقة ذات الجودة الطبيعية. باستخدام تقنيات التعلم العميق، وبالأخص شبكات الأعصاب الترددية المرتبطة، توصلت أبحاثنا إلى نظام عالي الأداء

تتمكن المساهمة الرئيسية في تحسين التقنيات الحالية وتصميم نظام قادر على الحصول على صور وجهية عالية الدقة بزيادة الدقة والمصدقية

بشكل عام، أظهرت نتائج دراستنا تحسناً ملحوظاً من حيث الجودة البصرية وحدة حدود الشكل والحفاظ على التفاصيل الوجهية. تمكنا من التفوق على الأساليب الحالية بفضل تصميم شبكة الأعصاب الترددية المرتبطة المحددة لتحسين دقة الصور الوجهية. كما يمكن للأبحاث المستقبلية التركيز بالتالي على تحسين هذه الجوانب واستكشاف تصاميم شبكات جديدة

دراستنا قد تحمل آثاراً مملوسة، خصوصاً في مجالات التعرف على الوجوه والمراقبة عبر الفيديو والتلفزيون عالي الدقة وأنظمة الأمن. يمكن أن تحسّن الصور الوجهية المعاد بناؤها دقة أنظمة التعرف على الوجوه وتقدّم تجربة بصرية أفضل. بالفعل، يكون وقت الحساب والطاقة اللازمين للحصول على صور فائقة الدقة في الوقت الحقيقي غالباً غير كافي لإعادة إنتاج الصور الوجهية

بشكل ملخص، قدمت هذا البحث مساهمة وإن كانت متواضعة في تطوير نظام تحسين دقة الصور الوجهية المعتمد على شبكات الأعصاب الترددية المرتبطة. النتائج المتحصل عليها مثيرة للاهتمام وتعرض أملاً لتحسين العديد من التطبيقات التي تتطلب صوراً عالية الدقة. ومع ذلك، ينبغي إجراء المزيد من الأبحاث للتغلب على القيود المحددة والاستمرار في استكشاف إمكانيات تحسين دقة الصور

الكلمات الدالة:

صورة عالية الدقة، صورة وجه، صورة منخفضة الجودة، صورة عالية الجودة، شبكة عصبونية تحويلية

(CNN).

Du visage vers une image de qualité naturelle

Résumé :

Ce mémoire de master porte sur le développement d'un système de super-résolution d'images de visages, visant à reconstruire des images basse résolution en images haute résolution de qualité naturelle. En utilisant des techniques d'apprentissage profond, en particulier les réseaux neuronaux convolutifs (CNN), notre recherche a donné naissance à un système performant.

Notre étude peut avoir des implications considérables, notamment dans les domaines de la reconnaissance faciale, de la vidéosurveillance, de la télévision à haute définition et des systèmes de sécurité. Les images faciales reconstruites peuvent améliorer la précision des systèmes de reconnaissance faciale et offrir une meilleure expérience visuelle. En effet, le temps de calcul et la puissance nécessaires à la super-résolution en temps réel sont souvent insuffisants pour permettre aux images faciales de se reproduire.

La principale contribution réside dans l'amélioration des techniques existantes ainsi que dans la conception d'un système permettant d'obtenir des images faciales de haute résolution avec une précision et une fidélité accrues.

Dans l'ensemble, les résultats de notre étude ont montré une forte amélioration en termes de qualité visuelle, de netteté des contours et de préservation des détails du visage. Nous avons pu surpasser les méthodes existantes grâce à la conception d'un CNN spécifique pour la super-résolution des images faciales.

Les axes de recherche futurs pourraient donc se concentrer sur l'amélioration de ces aspects et l'exploration de nouvelles architectures de réseau.

En synthèse, cette recherche a apporté une contribution significative au développement d'un système de super-résolution basé sur le CNN pour les images faciales. Les résultats obtenus sont intéressants et laissent espérer des améliorations dans diverses applications nécessitant des images à haute résolution. Toutefois, des recherches supplémentaires sont nécessaires pour surmonter les limites identifiées et continuer à explorer les possibilités de la super-résolution d'images.

Mots clefs : Image en super résolution, image de visage ,image basse résolution ,image haute résolution, réseau de neurones convolutifs CNN

From the face to high-quality natural images

Abstract :

From the face to a natural quality image Summary This master's thesis focuses on the development of a face image super-resolution system aimed at reconstructing low-resolution images into high-resolution natural quality images. Using deep learning techniques, particularly convolutional neural networks (CNN), our research has led to a high-performing system.

Our study may have significant implications, particularly in the fields of facial recognition, video surveillance, high-definition television, and security systems. Reconstructed facial images can improve the accuracy of facial recognition systems and offer a better visual experience.

However, the computational time and power required for real-time super-resolution are often insufficient for facial images to reproduce.

The main contribution lies in improving existing techniques and designing a system that can obtain high-resolution facial images with increased accuracy and fidelity. Overall, the results of our study have shown a strong improvement in terms of visual quality, edge sharpness, and preservation of facial details. We were able to surpass existing methods by designing a specific CNN for face image super-resolution.

The future research directions could therefore focus on improving these aspects and exploring new network architectures.

In summary, this research has made a significant contribution to the development of a super-resolution system based on CNN for facial images. The results obtained are interesting and give hope for improvements in various applications requiring high-resolution images. However, further research is needed to overcome the identified limitations and continue exploring the possibilities of image super-resolution.

Keywords : Super resolution image, face image ,low image ,high image, convolutional neural network CNN.

Liste des abréviations

HR Haute résolution

LR Basse résolution

SR Super résolution

IA Intelligence Artificielle

ML Apprentissage automatique

DL Apprentissage profond

CNN Réseau de neurones convolutif

RNN Réseau de neurones récurrent

LSTM Mémoire à court terme à long terme

GRU Unité récurrente à portes

DEEP RNN Réseau de neurones récurrent profond

VAE Autoencodeur variationnel

AAE Autoencodeur antagoniste

GAN Réseau antagoniste génératif

MSE Erreur quadratique moyenne

RMSE Erreur quadratique moyenne racine

PSNR Rapport signal sur bruit maximal

SSIM Indice de similarité structurale

VIF Facteur d'inflation de la variance

LFW Visages étiquetés dans la nature

VM Machine virtuelle

GPU Unité de traitement graphique

TPU Unité de traitement de tenseurs

Table des matières

Liste des tableaux	8
Table des figures	9
Introduction générale	1
1 État de l’art	3
1.1 Introduction	3
1.2 Concepts généraux et définitions	3
1.2.1 Qu’est ce qu’une image?	3
1.2.2 La super résolution d’image (Super image resolution)	5
1.2.3 La dégradation d’image	6
1.2.4 Intelligence artificielle, l’apprentissage automatique et l’appren- tissage profond	6
1.2.5 L’intelligence artificielle (IA)	7
1.2.6 Le machine learning (ML) :	7
1.2.7 L’apprentissage automatique	7
1.2.8 L’apprentissage profond (<i>DL :Deep learning</i>)	8
1.3 Les travaux connexes	16
1.4 Conclusion	20
2 La conception d’un CNN spécifique pour la super-résolution d’images faciales	21
2.1 Introduction :	21
2.2 Solution proposée pour la super résolution	21
2.2.1 Hyper paramètre utilisé	23
2.2.2 Aperçu général	23
2.2.3 Description détaillée de l’approche :	24
2.3 Conclusion :	25
3 Implémentation et résultats expérimentaux	27
3.1 Introduction :	27
3.2 Les outils de développement	27
3.2.1 Google Colab	27
3.2.2 Python	28
3.3 Les bibliothèques utilisées	28
3.3.1 Keras	28
3.3.2 Scikit-image	29
3.3.3 OpenCv	29

3.3.4	Numpy	29
3.3.5	Matplotlib	30
3.3.6	Conv2D	30
3.3.7	Adam	30
3.4	Résultats expérimentaux et discussions	30
3.4.1	Données d'expérimentation et les métriques d'évaluation de qualité	30
3.4.2	Tests et discussions	33
3.5	Conclusion	36
	Conclusion générale	37
	Bibliographie	39

Liste des tableaux

1.1	Récapitulation des réseaux d'apprentissage profond (DL :Deep learning)	9
1.2	Récapitulation des travaux connexes	18
2.1	Les paramètres d'apprentissage utilisés	26
3.1	Les exemples de résultats obtenus sur quelques images	35

Table des figures

1.1	Image basse et image haute résolution [J.M Field,]	4
1.2	IA et ses sous domaines.	7
1.3	Les techniques de l'apprentissage automatique.	8
1.4	Architecture typique d'un GAN [Vint et al., 2021]	10
1.5	L'architecture générale d'un RNN [Mishra et al., 2018]	11
1.6	L'architecture générale d'un auto-encodeur [Stewart, 2019]	12
1.7	L'architecture typique du CNN.	12
1.8	Évolution des CNN profonds illustrant les innovations architecturales depuis ConvNet jusqu'aux architectures actuelles.	16
1.9	Techniques d'apprentissage en profondeur avec leurs différents types.	17
2.1	L'architecture générale du système proposé.	22
2.2	L'architecture du modèle CNN.	24
3.1	Environnement du Google Colab [Carneiro et al., 2018]	28
3.2	Logo Python [Howse and Minichino, 2020]	28
3.3	Logo Keras [Kim et al., 2022]	28
3.4	Logo [Kuwahara et al., 2023]	29
3.5	Logo OpenCV [Intel, 2000]	29
3.6	Logo Numpy [Andersen, 2022]	30
3.7	Aperçu de l'architecture	31
3.8	Les étapes de l'architecture du CNN proposée pour la super-résolution	32
3.9	Résultats de l'apprentissage par epochs	33
3.10	Les résultats obtenues	33
3.11	La représentation graphique de l'évolution de l'erreur	34

Introduction générale

Contexte de la thèse

Les technologies de l'information et de la communication (TIC) sont au cœur de l'actualité. Il est donc essentiel de pouvoir améliorer la qualité des images à faible résolution pour mieux les exploiter et les analyser. La super-résolution d'images (**SRI**) est une solution prometteuse pour atteindre cet objectif, en utilisant des techniques d'apprentissage profond.

La **SRI** constitue un domaine de recherche passionnant qui vise à améliorer la qualité des images à faible résolution en les convertissant en images à haute résolution. Dans le contexte de cette thèse de fin d'études (**PFE**), nous souhaitons réaliser un système de super-résolution d'images conçu pour obtenir une image de haute qualité à partir d'une image à faible résolution. Pour cela, nous nous appuyons généralement sur une technique d'apprentissage profond appelée CNN (Convolutional Neural Network, ou réseau neuronal convolutif).

Problématique de la thèse

Bien que des progrès aient été réalisés dans ce contexte, il reste des défis à relever. La super-résolution d'images à partir d'une seule image à faible résolution est un problème complexe, car il est difficile de trouver les détails manquants et de produire ainsi une image réaliste de haute qualité. Nous souhaitons surmonter ces difficultés et concevoir un système de super-résolution d'images qui soit capable de fournir des résultats précis et convaincants.

Contributions de la thèse

Ce travail de recherche porte sur la conception et le développement d'un système de super-résolution d'images basé sur le CNN. Nous étudions différentes architectures CNN, ainsi que des mécanismes d'apprentissage et d'optimisation spécifiques pour améliorer la qualité des images générées. De plus, nous examinons la performance de notre système en utilisant des mesures objectives telles que l'erreur quadratique

moyenne (MSE), le rapport signal-bruit (PSNR) et l'indice de similarité structurelle (SSIM).

Organisation de la thèse

Outre cette introduction générale, qui présente le contexte dans lequel se place l'étude, la problématique de la recherche, les apports de la thèse et l'organisation du manuscrit, ce dernier est divisé en trois chapitres, comme indiqué ci-dessous.

Chapitre 1 : État de l'art. Il offre au lecteur un aperçu du domaine de recherche, ainsi que des explications rapides mais complètes. Il constitue une étude importante de la super-résolution d'images, couvrant les différentes techniques ainsi que les travaux qui s'y rapportent.

Chapitre 2 : Conception d'un CNN spécifique pour la super-résolution d'images faciales. se focalise sur la conception détaillée du CNN pour la super-résolution d'images, où les différentes couches et les architectures utilisées sont clairement définies.

Chapitre 3 : Implémentation et résultats expérimentaux. Il traite les essais et les évaluations des résultats obtenus avec le système, en utilisant des mesures objectives et en comparant les performances avec celles d'autres méthodes existantes.

Enfin, la conclusion générale souligne la valeur de l'approche, ainsi que ses limites, tout en proposant des perspectives de recherche future basées sur les principaux résultats de cette thèse.

État de l'art

1.1 Introduction

Le présent chapitre se concentre principalement sur la recherche de techniques de super-résolution permettant de transformer des images à faible résolution en images à haute résolution. En effet, les chercheurs souhaitent transformer les images à faible résolution en images à haute résolution, dans le but d'obtenir une perception plus détaillée et plus réaliste des visages.

Nous nous pencherons ici sur les techniques de super-résolution spécifiquement adaptées aux images faciales. Ensuite, il est essentiel de faire le point sur les travaux réalisés dans le domaine de la super-résolution d'images.

Les travaux réalisés jusqu'à présent ont exploré diverses approches pour améliorer la résolution des images. Cependant, dans les dernières années, les réseaux neuronaux profonds, en particulier les réseaux de convolution (CNN) et les réseaux adversaires génératifs (GAN), sont apparus comme des outils puissants pour la super-résolution d'images. Ces modèles exploitent les avantages de l'apprentissage automatique en utilisant de grandes quantités de données d'entraînement pour apprendre à reconstruire les détails manquants sur les images à faibles résolutions.

1.2 Concepts généraux et définitions

1.2.1 Qu'est ce qu'une image ?

Une image est une représentation visuelle d'un objet ou d'une scène du monde réel, capturée à l'aide d'un dispositif tel qu'un appareil photo ou un scanner. Une image est constituée d'une grille de minuscules éléments d'image appelés pixels, disposés en lignes et en colonnes. Chaque pixel contient une valeur qui représente la couleur ou la luminosité de cette partie particulière de l'image. En termes techniques, une image peut être décrite comme un tableau bidimensionnel de valeurs numériques utilisées pour représenter l'intensité et la couleur de chaque pixel.

Les images sont utilisées dans un large éventail d'applications, de la photographie numérique à la conception graphique, en passant par l'imagerie médicale et la

recherche scientifique. Elles constituent un élément essentiel de la communication visuelle et jouent un rôle crucial dans la transmission d'informations et d'émotions. Dans le domaine de la médecine, les images sont utilisées pour diagnostiquer et traiter diverses conditions médicales, telles que le cancer, en fournissant des informations détaillées sur les organes et les tissus internes du corps. Dans le domaine de la vision artificielle et de l'apprentissage automatique, les images sont utilisées pour entraîner des algorithmes à reconnaître des objets et des modèles, ce qui permet aux machines d'effectuer des tâches telles que la reconnaissance faciale et la conduite autonome. En résumé, les images sont un élément essentiel de notre monde moderne, et leurs applications sont diverses et d'une grande portée.

1.2.1.1 L'image basse Vs l' image haute (low-image Vs high image)

Les concepts d'image de basse (*faible*) qualité et d'image de haute qualité peuvent être expliqués en termes techniques. La notion d'image faible fait référence à la qualité d'une image avec une faible résolution, une pixellisation ou une perte de détails (voir la figure 1.1) [J.M Field,]. Elle se caractérise généralement par des images floues qui manquent de netteté et de clarté. D'autre part, l'image de haute qualité fait référence à la qualité d'une image avec une résolution élevée, une grande netteté et une grande clarté. Elle se caractérise généralement par des images aux couleurs vives, des détails fins et un excellent contraste.

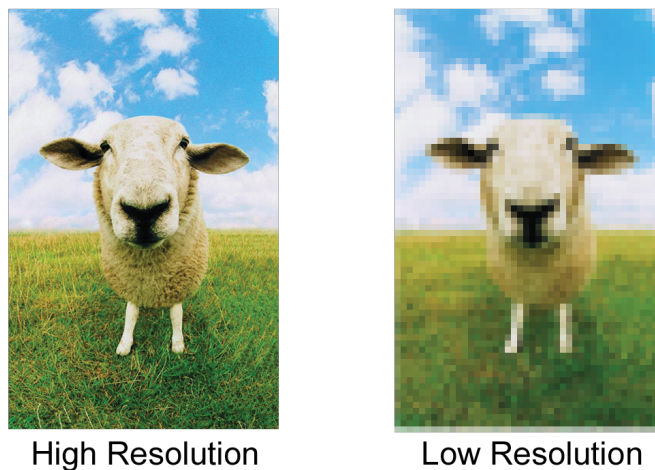


FIGURE 1.1 – Image basse et image haute résolution [J.M Field,]

La différence entre une image faible et une image haute réside dans le nombre de pixels qui composent l'image. Une image basse a un nombre inférieur de pixels, tandis qu'une image haute a un nombre supérieur de pixels. Plus le nombre de pixels est élevé, plus l'image est claire et nette. Par conséquent, une image de haute qualité nécessite un appareil photo ou un dispositif à haute résolution capable de capturer plus de pixels.

1.2.1.2 L'image aveugle (Blind image)

L'imagerie aveugle est une technique utilisée en traitement d'images dont l'objectif est de récupérer une image à partir de mesures incomplètes ou corrompues. Le principe de base de cette technique est l'utilisation d'algorithmes mathématiques pour reconstruire une image à partir d'informations limitées. Dans le cas de l'imagerie en aveugle, l'image n'est pas connue à l'avance et le processus consiste donc à estimer l'image à partir des données observées. Cette technique repose sur l'hypothèse que les images possèdent certaines propriétés statistiques, qui peuvent être exploitées pour récupérer les informations manquantes.

1.2.2 La super résolution d'image (Super image resolution)

La super résolution d'image (SRI) est une technologie de pointe qui utilise des algorithmes avancés et des techniques d'apprentissage automatique pour produire des images à haute résolution à partir d'entrées à faible résolution. Le principe de la SRI est de générer une image haute résolution en déduisant les détails et les informations manquants d'images basse résolution. Pour ce faire, l'image basse résolution est analysée et des modèles d'apprentissage automatique sont utilisés pour estimer les informations manquantes. Ces modèles sont entraînés sur un vaste ensemble de données d'images à haute résolution, ce qui leur permet de prédire avec précision les informations manquantes dans l'image à basse résolution

La super-résolution d'image est une technique utilisée pour améliorer la résolution d'une image, ce qui permet d'obtenir une image de meilleure qualité et plus détaillée. Le processus implique l'utilisation d'algorithmes qui analysent l'image et prédisent les données manquantes pour augmenter la résolution. Ce processus comporte différentes étapes :

- **Acquisition de l'image** : Elle consiste à capturer une image à l'aide d'un appareil photo ou d'un scanner. L'image est généralement de faible résolution en raison des limitations de l'appareil d'imagerie ou d'un sous-échantillonnage pour le stockage ou la transmission [Szeliski, 2022].
- **Prétraitement de l'image** : ce stade consiste à traiter l'image à faible résolution pour le processus de super-résolution. Cette étape peut consister à supprimer le bruit ou le flou de l'image, à recadrer l'image ou à aligner plusieurs images [Szeliski, 2022].
- **Super-résolution de l'image** : Dans cette étape, un modèle et/ou un algorithme est utilisé pour augmenter la résolution de l'image. Différentes méthodes permettent de le faire, telles que l'interpolation, la reconstruction ou les méthodes basées sur l'apprentissage profond [Milanfar, 2017].
- **Post-traitement de l'image** : ce stade implique d'affiner l'image de sortie en supprimant les artefacts ou en améliorant la qualité de l'image. Cela peut inclure le flou, le débruitage ou la compensation des couleurs [Szeliski, 2022].
- **Évaluation de l'image** : Ce processus permet l'évaluation de la qualité de l'image de sortie par rapport à l'image haute résolution d'origine, si cette dernière est

déjà connue. Cela peut se faire à l'aide de mesures spécifiques telles que le PSNR [Milanfar, 2017], le SSIM [Milanfar, 2017] ou le VIF [Wang et al., 2004].

Au cours de la première étape, l'image à faible résolution est acquise à l'aide de n'importe quel dispositif d'imagerie. Au cours de la phase de prétraitement, l'image est filtrée afin d'éliminer tout bruit ou artefact susceptible d'en altérer la qualité. L'extraction des caractéristiques consiste à identifier les caractéristiques significatives de l'image, telles que les bords, les coins et les textures. Enfin, l'image est mise à l'échelle à l'aide de diverses techniques telles que l'interpolation, l'apprentissage profond et les réseaux adversaires génératifs (GAN).

Cependant, il convient de noter que le processus de super-résolution ne peut pas ajouter de nouvelles informations à une image qui n'a pas été capturée à l'origine, et que la qualité de sortie est limitée par la qualité d'entrée.

1.2.3 La dégradation d'image

La dégradation de l'image se réfère à la perte de qualité (et V ou) de détails d'une image à cause de divers facteurs, dont la compression, le bruit, le flou et la réduction de la résolution. Cela peut affecter la qualité visuelle d'une image, en la rendant moins nette, floue, ou encore moins précise.

- *Compression* : les algorithmes de compression d'image peuvent entraîner une perte d'informations et d'artefacts dans les images, en particulier lorsque la compression est élevée.
- *Bruit* : Un bruit aléatoire peut être introduit dans une image en raison de diverses sources, telles que le bruit électronique dans les capteurs d'image ou les opérations de traitement d'image.
- *Flou* : le flou de l'image peut se produire en raison d'un mouvement, d'une optique floue ou d'autres facteurs, ce qui rend l'image moins nette et réduit la visibilité des détails fins.
- *Réduction de la résolution* : Réduire la résolution d'une image peut entraîner une perte de détails et peut rendre l'image plus grossière et moins claire.

En général, la dégradation de l'image est considérée comme un problème inverse, car l'image originale doit être récupérée à partir des données dégradées.

1.2.4 Intelligence artificielle, l'apprentissage automatique et l'apprentissage profond

L'apprentissage profond (*Deep Learning*), l'apprentissage automatique (*Machine Learning*) et l'intelligence artificielle sont trois concepts interdépendants qui ont révolutionné le domaine de l'informatique (voir la figure 1.9). L'apprentissage profond est un sous-ensemble de l'apprentissage automatique, qui est un type d'intelligence artificielle (IA) [Prasad Lokulwar, 2022], [Zhang, 2010].

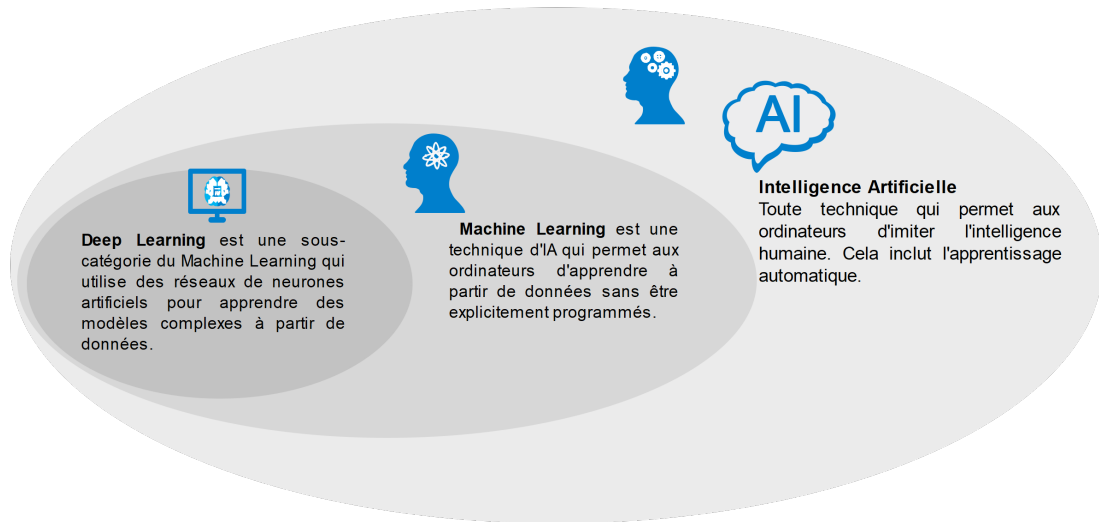


FIGURE 1.2 – IA et ses sous domaines.

1.2.5 L'intelligence artificielle (IA)

L'IA est la capacité des machines à imiter les fonctions cognitives humaines telles que l'apprentissage, la perception, le raisonnement et l'interaction sociale.

1.2.6 Le machine learning (ML) :

Le ML est une technique d'IA qui permet aux ordinateurs d'apprendre à partir de données sans être explicitement programmés.

1.2.7 L'apprentissage automatique

L'apprentissage automatique est un sous-domaine de l'intelligence artificielle (IA) qui implique le développement d'algorithmes et de modèles statistiques permettant aux systèmes informatiques d'apprendre et de s'améliorer à partir de l'expérience sans être explicitement programmés. Le principe fondamental de l'apprentissage automatique est de permettre aux ordinateurs d'apprendre à partir de données, d'identifier des modèles et de faire des prédictions ou de prendre des décisions sur la base de cet apprentissage. Il s'agit d'entraîner le système informatique à reconnaître des modèles, des relations et des corrélations dans de vastes ensembles de données, puis d'utiliser ces connaissances pour faire des prédictions ou prendre des décisions concernant de nouvelles données [Zhang, 2010].

Au cœur de la ML se trouve le concept d'apprentissage piloté par les données, qui consiste à alimenter un système informatique avec de grandes quantités de données et à utiliser des techniques statistiques pour identifier des schémas et des relations dans ces données [Mahesh, 2020]. Ce processus comporte plusieurs étapes clés, notamment la collecte de données, le prétraitement des données, la sélection des caractéristiques, le

développement d'algorithmes, l'entraînement des modèles et l'évaluation des modèles. Les algorithmes de ML peuvent être classés en quatre grandes catégories :

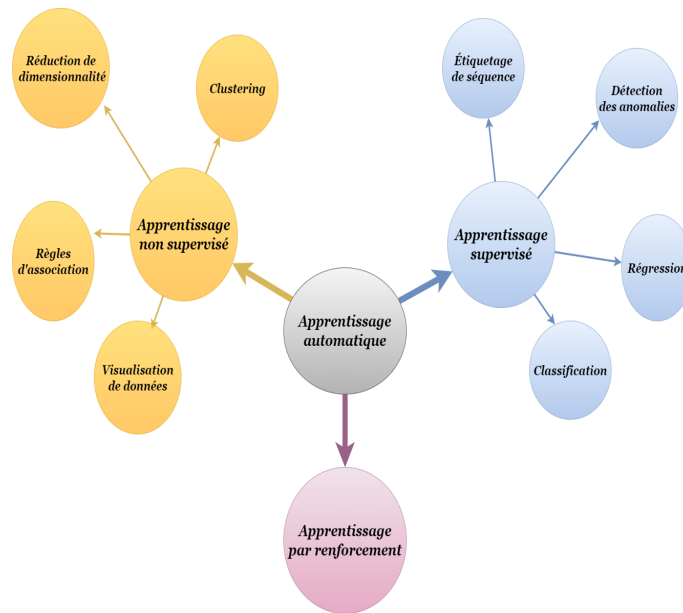


FIGURE 1.3 – Les techniques de l'apprentissage automatique.

- **Apprentissage supervisé** : Dans cette approche, un modèle est entraîné à partir de données d'entrée (caractéristiques) et des étiquettes correspondantes (résultats souhaités). Le modèle apprend à prédire les résultats pour de nouvelles données en généralisant à partir des exemples fournis [Bishop and Nasrabadi, 2006].
- **Apprentissage non supervisé** : Cette technique est utilisée lorsque les données ne sont pas étiquetées. L'objectif est de découvrir des structures ou des patterns inhérents aux données sans informations a priori. Les techniques courantes incluent la classification non supervisée, la réduction de dimensionnalité et le regroupement (clustering) [Radford et al., 2015].
- **Apprentissage par renforcement** : Cette technique implique un agent qui interagit avec un environnement dynamique. L'agent prend des actions et reçoit des récompenses ou des punitions en fonction de ses actions. L'objectif est d'apprendre une politique (stratégie) optimale qui maximise la récompense à long terme [Sutton and Barto, 2018].

Chacun de ces algorithmes possède ses propres forces et faiblesses et est adapté à différents types de problèmes.

1.2.8 L'apprentissage profond (DL :Deep learning)

Le DL est une sous-catégorie du Machine Learning qui utilise des réseaux de neurones artificiels pour apprendre des modèles complexes à partir de données voir la figure 1.9

Les techniques d'apprentissage profond, également connues sous le nom de Deep Learning, sont basées sur l'utilisation de réseaux de neurones artificiels pour modéliser et

résoudre des problèmes complexes. En utilisant une méthode d'apprentissage supervisé, le réseau est entraîné à partir d'un grand ensemble de données d'entraînement pour identifier des motifs et des relations cachées entre les entrées et les sorties. Ensuite, le modèle est utilisé pour prédire des sorties pour de nouvelles entrées en utilisant les motifs appris à partir des données d'entraînement [Chinnasamy et al., 2023].

Réseaux d'apprentissage profond	Caractéristiques majeures	Documents de références
RvNN	Utilisation d'une structure arborescente. Il est conseillé pour le NLP	[Goller and Kuchler, 1996]; [Socher et al., 2011]
RNN	Adapté aux informations séquentielles. Recommandé pour le NLP et le traitement de la parole	[Cho et al., 2014]; [Li and Wu, 2015]
CNN	Au départ, cette technique a été utilisée pour la reconnaissance d'images. Étendue au NLP, au traitement de la parole, et la vision par ordinateur	[LeCun et al., 1995]; [Krizhevsky et al., 2017]; [Zhang and Wallace, 2015]; [Abdel-Hamid et al., 2014]
DBN	Apprentissage non supervisé-Connexions orientées	[Hinton, 2009]; [Hinton et al., 2012]
DBM	Apprentissage non supervisé. Un modèle composé de RBM Connexions non orientées	[Hinton and Salakhutdinov, 2012]
GAN	Apprentissage non supervisé-Le framework de la théorie des jeux	[Goodfellow et al., 2014]; [Radford et al., 2015]
VAE	Apprentissage non supervisé Modèle graphique probabiliste- Graphique probabiliste	[Kingma and Welling, 2013]

TABLE 1.1 – Récapitulation des réseaux d'apprentissage profond (DL :Deep learning)

Les techniques d'apprentissage profond ont de nombreuses applications dans divers domaines tels que la vision par ordinateur, le traitement du langage naturel et la reconnaissance de la parole. Dans la vision par ordinateur, les réseaux de neurones profonds sont utilisés pour la détection d'objets, la segmentation d'images et la recon-

naissance faciale. Dans le traitement du langage naturel, les modèles d'apprentissage profond sont utilisés pour la traduction automatique, la classification de texte et la génération de texte. Enfin, dans la reconnaissance de la parole, les réseaux de neurones profonds sont utilisés pour la transcription de la parole en texte et la synthèse de la parole. Les techniques d'apprentissage profond sont également utilisées pour des applications telles que la recommandation de produits, la détection de fraude et la prédiction de l'activité sismique.

Les techniques d'apprentissage profond incluent les **réseaux de neurones convolutifs**, les **réseaux de neurones récurrents**, les **réseaux de neurones adverses génératifs**, et les **réseaux de neurones à mémoire à court et long terme** [Pouyanfar et al., 2018].

- **Les GAN (Generative Adversarial Networks)** peuvent être utilisés en combinaison avec le CNN pour améliorer la qualité de la super résolution des images de visage. Dans un GAN, le générateur, basé sur le CNN, génère des images haute résolution à partir d'un espace latent. Le discriminateur, également basé sur le CNN, évalue la qualité des images générées en les comparant aux images réelles haute résolution. Au fur et à mesure de l'entraînement, le générateur CNN s'améliore en générant des images de visage de haute résolution qui ressemblent de plus en plus à celles du jeu de données d'entraînement. Les GAN permettent ainsi d'obtenir des images de visage super résolues réalistes et de haute qualité [Goodfellow et al., 2014].

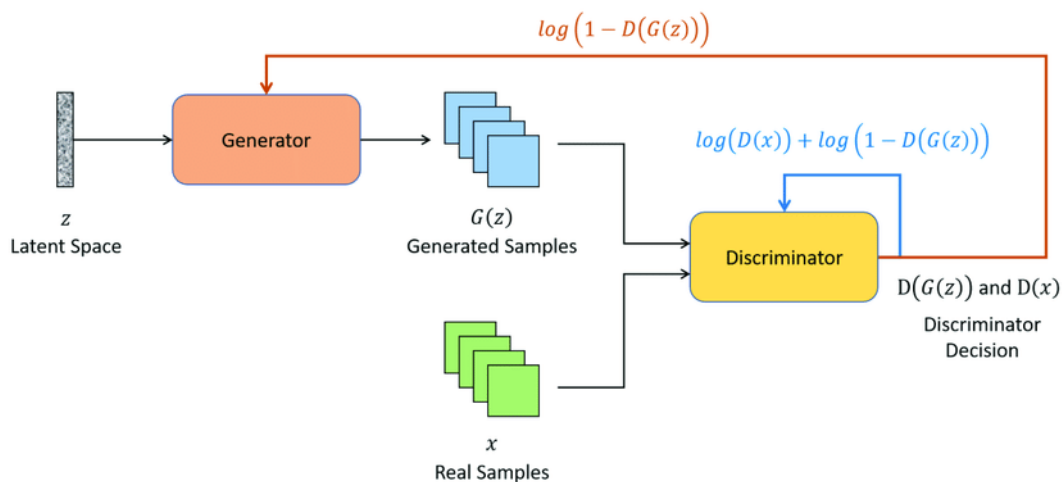


FIGURE 1.4 – Architecture typique d'un GAN [Vint et al., 2021]

- **Le RNN (Recurrent Neural Network)** est généralement utilisé pour modéliser des séquences, telles que des séquences temporelles ou des séquences de mots. Dans le contexte de la super résolution des images de visage, le RNN peut être utilisé pour améliorer la qualité et la cohérence des détails super résolus. Par exemple, il peut être utilisé pour générer des descriptions textuelles détaillées des caractéristiques du visage, qui peuvent ensuite être utilisées pour guider le processus de super résolution. Cela permet de mieux aligner les détails super résolus avec les caractéristiques attendues des visages [Rumelhart et al., 1986] [Graves and Graves, 2012].

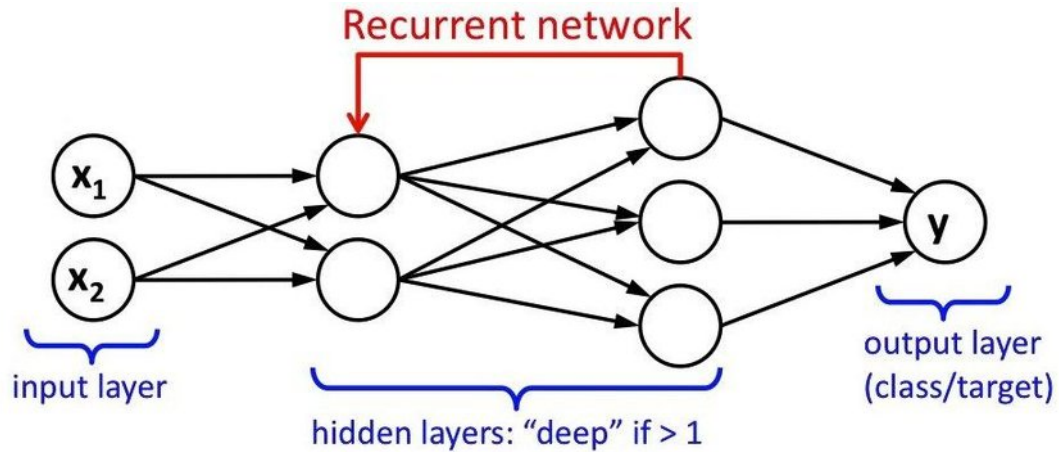


FIGURE 1.5 – L’architecture générale d’un RNN [Mishra et al., 2018]

- **L’autoencodeur** est une autre architecture de réseau de neurones utilisée pour la compression et la reconstruction des données. L’autoencodeur CNN peut être utilisé pour extraire les caractéristiques visuelles des images de visage, puis les reconstruire pour obtenir une version super résolution. L’autoencodeur peut être entraîné à reconstruire les images de visage originales à partir de leur version basse résolution, en utilisant le CNN pour extraire les caractéristiques et le décodeur pour reconstruire les images. Cela permet d’obtenir une version super résolution des images de visage tout en conservant les caractéristiques et les détails importants [Bengio et al., 2006].
- **Un CNN (Convolutional Neural Network)** est une architecture de réseau de neurones qui utilise des couches convolutives pour extraire automatiquement des caractéristiques pertinentes à partir des données d’entrée, en particulier des images. Il est spécifiquement conçu pour exploiter la structure spatiale des données en utilisant des opérations de convolution pour détecter des motifs locaux. Le CNN est généralement composé de : **Couches de convolution, Couches de pooling, Couches entièrement connectées, Fonctions d’activation et Couches de sortie**

☞ **Caractéristiques**

- Les CNNs sont principalement utilisés pour traiter des données visuelles telles que des images et des vidéos.
- Celles-ci se composent de couches convolutives, de couches de regroupement et de couches entièrement connectées.
- Les CNNs sont couramment utilisés pour la classification d’images, la détection d’objets, la segmentation sémantique, etc.
- Ceux-ci sont formés à l’aide d’un algorithme de rétropropagation de gradient pour minimiser l’erreur entre les prédictions du réseau et les étiquettes de sortie.

☞ **La structure générale du CNN** le CNN est composé de plusieurs couches, notamment

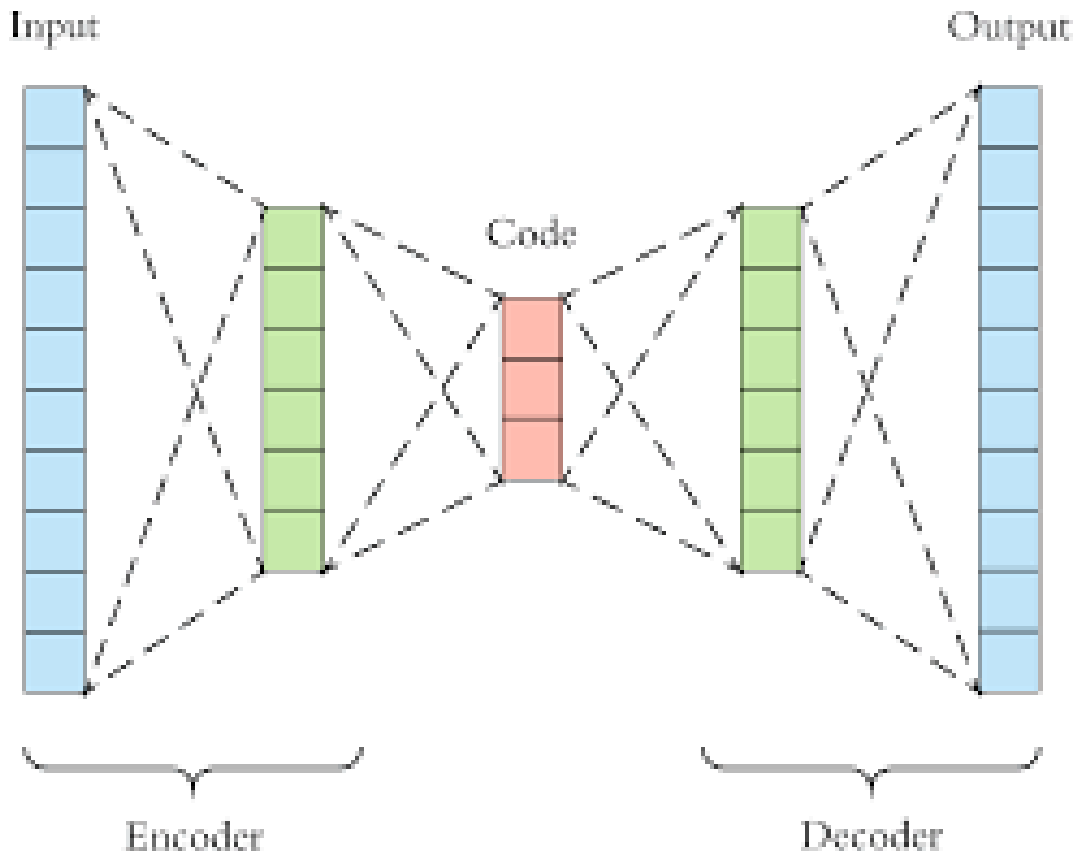


FIGURE 1.6 – L'architecture générale d'un auto-encodeur [Stewart, 2019]

des couches de convolution, de pooling et de couches entièrement connectées.

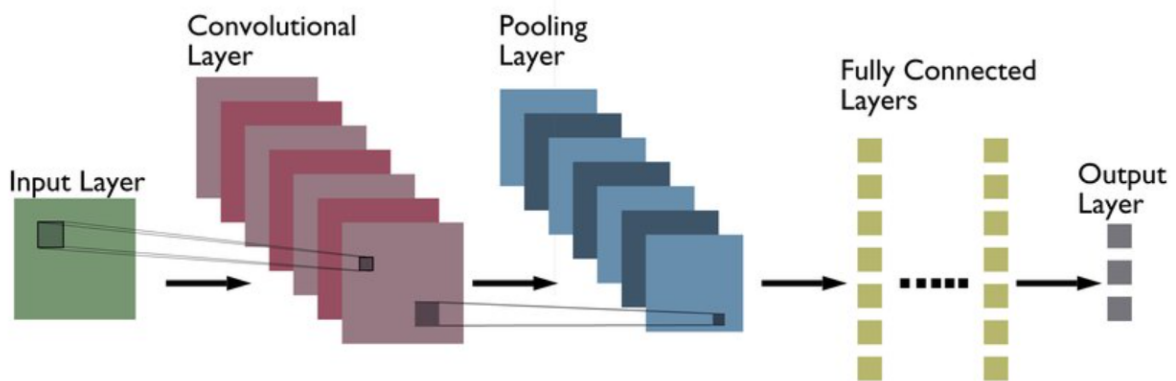


FIGURE 1.7 – L'architecture typique du CNN.

- **La couche de convolution** : Les couches de convolution sont un élément clé des CNN, qui nous permettent d'extraire des caractéristiques visuelles des images d'entrée. Le filtre de convolution est appliqué localement à une région de l'image et effectue une opération de convolution pour calculer la sortie. Les couches de convolution peuvent conte-

nir plusieurs filtres, chacun détectant des caractéristiques telles que les contours, les textures et des formes spécifiques. En combinant les résultats de ces filtres, les CNN capturent des caractéristiques complexes et génèrent des cartes de caractéristiques multidimensionnelles. Ces cartes servent de données d'entrée pour les couches suivantes du réseau [LeCun et al., 1998a] [Goodfellow et al., 2016].

- **Activation non linéaire** : Une fois que les couches de convolution sont réalisées, une fonction d'activation non linéaire est appliquée à la sortie de chaque filtre afin d'introduire la non-linéarité dans le réseau. Le type de fonction d'activation le plus couramment utilisé est la fonction ReLU (Rectified Linear Unit, unité linéaire rectifiée). Cette fonction permet de modéliser des relations complexes entre les caractéristiques extraites par les filtres de convolution. Grâce à la fonction ReLU, les CNN reconnaissent les informations pertinentes dans les images de manière plus robuste et plus sélective en éliminant les valeurs négatives. De plus, la fonction ReLU est souvent choisie pour l'activation non linéaire des CNN en raison de sa simplicité et de son efficacité de calcul [Glorot et al., 2011].
- **Couche pooling** : Les couches de regroupement constituent une importante partie des CNN, car elles permettent de réduire la dimensionnalité spatiale des caractéristiques extraites tout en préservant les informations importantes. Le pooling est réalisé en appliquant des opérations. Le regroupement maximum sur des régions qui ne se chevauchent pas dans les cartes de caractéristiques. Par exemple, le regroupement maximal conserve la valeur maximale dans chaque région, tandis que le regroupement moyen prend la moyenne des valeurs. Le regroupement réduit le nombre de paramètres du réseau, ce qui permet d'économiser de la mémoire et des calculs. Le pooling réduit le nombre de paramètres du réseau, ce qui permet d'économiser de la mémoire et des calculs. En bref, les couches de mise en commun jouent un rôle important dans la réduction de la dimensionnalité des caractéristiques extraites tout en préservant les informations importantes [Krizhevsky et al., 2017] [LeCun et al., 2015].
- **couche entièrement connectée** : Une fois les étapes de convolution et de regroupement effectuées dans le CNN, les caractéristiques extraites sont aplaties et introduites dans des couches entièrement connectées. Chaque neurone de cette couche dense est connecté à tous les neurones de la couche précédente. Les couches entièrement connectées sont utilisées pour la classification finale et certaines autres tâches, en combinant les caractéristiques précédemment extraites pour déterminer les classes et les catégories d'images. Toutefois, l'utilisation de couches entièrement connectées peut entraîner une complication du modèle et nécessiter davantage de ressources informatiques. Les nouvelles architectures de CNN, telles que le Global Average Pooling, réduisent l'utilisation de couches entièrement connectées en faveur de couches de classification plus légères [Szegedy et al., 2015a].
- **La couche de sortie** : La couche de sortie est la dernière couche d'un

CNN et son type dépend de la tâche spécifique à laquelle le réseau est destiné. Pour une tâche de classification, la couche de sortie utilise généralement une fonction d'activation, comme la fonction softmax, qui attribue des probabilités à chaque classe possible. La somme des probabilités de toutes les classes est égale à 1, permettant ainsi d'obtenir une distribution de probabilités pour la classification. Pour d'autres tâches, telles que la détection d'objets, la couche de sortie peut utiliser différentes fonctions d'activation appropriées. Par exemple, la fonction sigmoïde peut être utilisée pour la détection d'objets où chaque neurone de sortie est associé à une classe spécifique et indique la probabilité de présence de cette classe. Il est important de choisir la fonction d'activation appropriée en fonction de la tâche spécifique pour obtenir des résultats précis et interprétables [Chollet, 2021].

☞ **Application du CNN** : Les CNN (Convolutional Neural Networks) ont trouvé de nombreuses applications dans divers domaines. Voici quelques-unes des principales applications des CNN [Chollet, 2021], [LeCun et al., 2015], [Krizhevsky et al., 2017] :

- *Classification des images* : Les CNN sont largement utilisés pour la classification des images et peuvent détecter et classer automatiquement les objets présents dans les images. Les CNN peuvent être formés sur de grands ensembles de données pour distinguer différentes classes d'objets telles que les animaux, les véhicules et les aliments.
- *Détection d'objet* : Les CNNs sont utilisés pour détecter et localiser des objets spécifiques dans les images. Vous pouvez trouver des régions d'intérêt (ROI) et prédire les classes d'objets présentes dans ces régions. Cette capacité de détection d'objets est utilisée dans des applications telles que la vidéosurveillance, la conduite autonome et la réalité augmentée.
- *Segmentation sémantique* : CNN est utilisé pour effectuer une segmentation sémantique. C'est -à-dire divisez l'image en différentes régions et attribuez une étiquette sémantique à chaque région. La structure et le contenu des images peuvent être compris plus en détail, ce qui est utile dans des domaines tels que la médecine (analyse d'images médicales) et la création de cartes.
- *Reconnaissance de visage* : Les CNN sont couramment utilisés pour la reconnaissance faciale et aident à identifier et à vérifier les visages dans les images et les vidéos. Les CNN peuvent extraire des caractéristiques uniques des visages, telles que : Placement des contours, des textures et des caractéristiques. Cela permet une reconnaissance précise même lorsque l'éclairage, les expressions faciales et les poses changent.
- *Traduction automatique et traitement du langage naturel* : Les CNNs sont également utilisés dans le domaine du traitement du langage naturel, en particulier la traduction automatique et la génération de texte. Les CNNs peuvent être utilisés pour extraire des caractéristiques de séquences de mots afin d'obtenir des informations contextuelles et sémantiques importantes pour les tâches de traitement du langage.

- ☞ **Les différents types de réseaux de neurones convolutionnels (CNN)** La figure 1.8 fait référence aux différents types de réseaux neuronaux convolutifs (CNN), conçus pour des tâches spécifiques. Parmi les types de CNN les plus courants, certains sont présentés ci-dessous :
- **LeNet** : Développé par Yann LeCun dans les années 1990 pour la reconnaissance de caractères manuscrits [LeCun et al., 1998b].
 - **AlexNet** : Développé en 2012 par Alex Krizhevsky pour la classification d’images [Krizhevsky et al., 2012].
 - **VGGNet** : Développé en 2014 par le groupe Visual Geometry Group (VGG) de l’université d’Oxford pour la classification d’images [Simonyan and Zisserman, 2015].
 - **GoogLeNet (Inception)** : Développé en 2014 par Google pour la classification d’images. Il utilise des blocs Inception qui combinent plusieurs opérations de convolution à différentes échelles spatiales [Szegedy et al., 2015b].
 - **ResNet** : Développé en 2015 par Microsoft pour la classification d’images. Il utilise des blocs de résidus qui permettent aux réseaux de neurones de s’entraîner plus efficacement pour des tâches plus complexes [He et al., 2015].
 - **DenseNet** : Développé en 2017 par Gao Huang et al. pour la classification d’images. Il utilise des blocs de connexion dense pour améliorer la propagation de l’information dans le réseau [Huang et al., 2017].

De manière générale, le CNN est essentiel pour l’extraction des caractéristiques visuelles et la génération d’images super-résolues de haute qualité dans le système d’images faciales à super-résolution. Les GAN, RNN et auto-encodeurs peuvent être utilisés en combinaison avec le CNN pour améliorer la qualité, la cohérence et la précision des détails des images faciales super-résolues. Ces architectures permettent d’obtenir des résultats impressionnants en matière de super-résolution pour les images faciales.

En effet, l’apprentissage profond évolue très rapidement, avec de nombreux nouveaux réseaux et architectures apparaissant tous les quelques mois, ce qui dépasse notre intention et notre capacité à traiter ce problème dans le présent travail. La table 1.1 contient une synthèse des réseaux d’apprentissage profond les plus populaires, tels que le réseau neuronal récurrent (RvNN), le RNN, le CNN et les modèles génératifs profonds, leurs principaux points forts et les publications les plus significatives [Chinnasamy et al., 2023], [Pouyanfar et al., 2018].

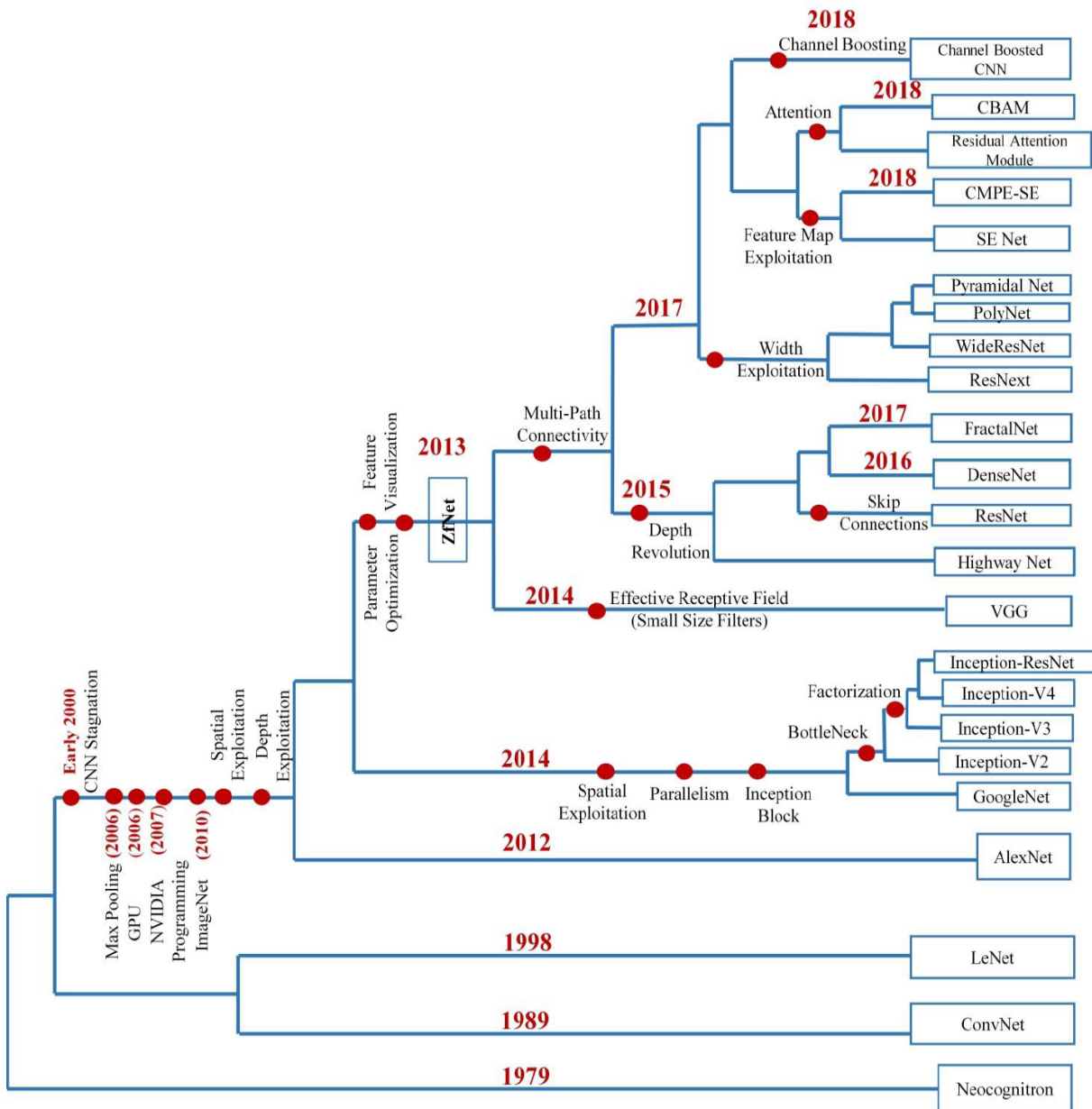


FIGURE 1.8 – Évolution des CNN profonds illustrant les innovations architecturales depuis ConvNet jusqu'aux architectures actuelles.

1.3 Les travaux connexes

De nombreuses approches ont été proposées pour améliorer la résolution et le niveau de détail des images, ce qui constitue un sujet de recherche qui évolue sans cesse. Parmi les contributions à ce domaine, on peut citer la super-résolution d'images, la déconvolution, la restauration d'images et la super-résolution vidéo. Dans chaque cas, différentes méthodes apportent leur contribution unique à ces domaines, ce qui a permis des avancées significatives. Une récapitulation des travaux connexes est donnée par la table 1.2

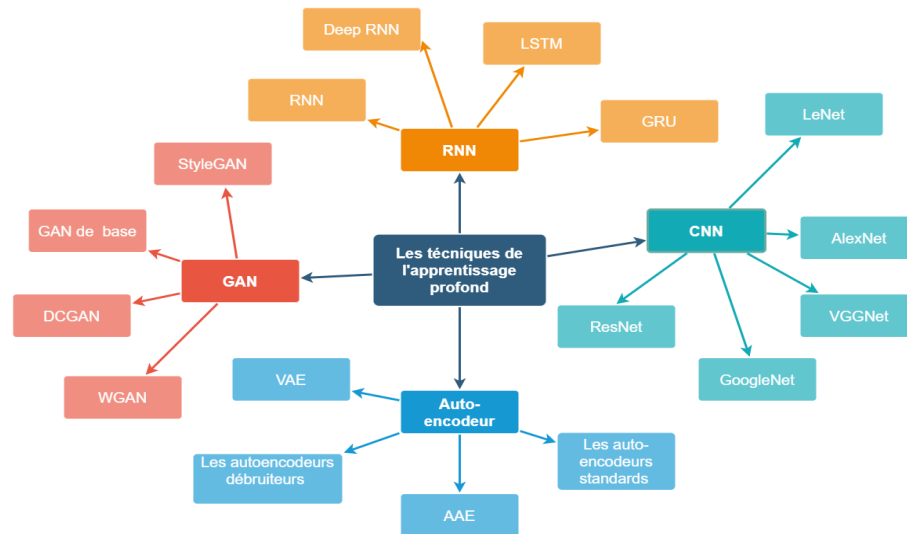


FIGURE 1.9 – Techniques d'apprentissage en profondeur avec leurs différents types.

[Dong et al., 2015], ont présenté comment exploiter les réseaux de convolution profonde pour la super-résolution d'images. Les résultats atteints sont prometteurs, car ils permettent d'augmenter la résolution tout en préservant les détails importants.

une autre méthode, développée par [Shocher et al., 2017], consiste à effectuer une approche de " *super-résolution sans apprentissage* ". Cette méthode exploite les informations internes de l'image pour améliorer sa résolution, sans nécessiter de jeu de données d'apprentissage.

[Zhang et al., 2018] sont arrivés à la mise en place du réseau résiduel dense, qui utilise les connexions résiduelles pour améliorer les performances du réseau de super-résolution d'images. A noter que cette méthode se distingue tout particulièrement par sa capacité à produire des résultats précis et détaillés.

Pour la restauration d'images, [Tai et al., 2017] ont proposé **MemNet**, un réseau de mémoire persistante. A travers cette approche, la mémoire à long terme est utilisée pour conserver les informations utiles et améliorer la qualité de la restauration.

L'utilisation d'auto-encodeurs convolutifs avec des connexions de saut symétriques pour la restauration d'images a été explorée par Mao et al [Mao et al., 2016]. Cette méthode permet de restaurer des images en conservant les détails fins et en réduisant les effets de flou.

À propos de la déconvolution, Zhang et al [Zhang et al., 2017] ont utilisé l'apprentissage résiduel CNN pour améliorer la qualité des images déconvoluées, avec des performances supérieures à celles des méthodes traditionnelles basées sur des dé-bruiteurs gaussiens.

Approche basée sur	Référence	Architecture	Base de données	SSIM
Super-résolution d'images aveugles	[Dong et al., 2015]	DCNN :L'utilisation des réseaux de neurones convolutionnels profonds pour la super-résolution d'images.	Set5	97.25 %
			Set14	93.76 %
			BSD200	92.61 %
	[Shocher et al., 2017]	ZSSR :L'utilisation de l'apprentissage interne profond pour atteindre des résultats sans exemple d'apprentissage spécifique.	Set5	87.96 %
			Set14	76.51 %
			BSD100	72.11 %
Détection de dégradation d'image	[Tai et al., 2017]	MemNet :un réseau de mémoire persistante pour la restauration d'images.	Set5	88.93 %
	[Zhang et al., 2018]	RDN :un réseau résiduel dense pour la super-résolution d'images.	Set5	90.03 %
			Set14	78.93 %
			B100	74.34 %
			Urban100	80.69 %
			Manga109	91.84 %
Restauration d'image	[Mao et al., 2016]	Auto-encodeur convolutif avec des connexions sautées symétriques pour la restauration d'images.	Set5	88.69 %
			Set14	77.18 %
			BSD200	65.44 %
	[Zhang et al., 2017]	DnCNN :une méthode de débruitage d'images basée sur l'apprentissage résiduel d'un réseau de neurones convolutifs profonds.	BSD68	70.76 %
			Set14	76.72 %
			Urban100	75.21 %

TABLE 1.2 – Récapitulation des travaux connexes

À la différence des réseaux neuronaux convolutifs (CNN), le GAN vise à générer des images plus réalistes ou HR qui correspondent davantage à la perception de l'œil humain, tandis que l'objectif du CNN est de reconstruire de manière fiable

des informations à haute fréquence à partir d'images. De nombreux réseaux améliorés basés sur le GAN ont été proposés ces dernières années [Wang et al., 2018]. On peut dire que, dans une certaine mesure, ces réseaux ont résolu au mieux les problèmes susmentionnés et obtenu de bonnes performances en termes de restitution.

En réalité, il est intéressant de noter que ces différentes approches contribuent de manière remarquable à l'amélioration de la résolution, de la qualité et du niveau de détail des images obtenues. Des progrès significatifs ont été réalisés en matière de super-résolution et de déconvolution d'images, grâce à la combinaison de techniques telles que les réseaux de convolution profonds, les GAN, l'apprentissage de la distribution du gradient et l'estimation du noyau du gradient. Les recherches se poursuivent en vue d'obtenir des images de haute qualité et de haute résolution.

Une méthode basée sur des réseaux de convolution profonde est présentée dans l'article de [Dong et al., 2015]. Ils se concentrent sur la préservation des détails fins tout en augmentant la résolution de l'image. Cette approche a permis de fournir les bases de nombreuses autres méthodes dans ce domaine.

En effet, une autre approche populaire consiste à exploiter les réseaux adversaires génératifs (GAN) pour la super-résolution d'images. La contribution de [Ledig et al., 2017] illustre une idée originale et montre comment les GANs peuvent générer des images à haute résolution avec des détails réalistes. Cette méthode ouvre le chemin à de nombreux progrès ultérieurs en ce domaine.

Plusieurs méthodes sont axées sur la super-résolution aveugle, c'est-à-dire la restauration d'images à faible résolution sans connaître les dégradations spécifiques présentes. Selon un article de [Yue et al., 2022], les auteurs proposent une approche élaborée pour modéliser les dégradations, en prenant en compte le bruit et le flou. En revanche, l'article de [Luo et al., 2022] se concentre sur l'apprentissage de la distribution des dégradations pour la super-résolution aveugle des images. Ils exploitent les informations relatives aux dégradations dans les images à faible résolution afin de renforcer la qualité ainsi que la résolution des images en sortie.

Une solution efficace de convolution sous-pixel pour la super-résolution en temps réel d'images et de vidéos a été introduite par Shi et al [Shi et al., 2016]. Plus précisément, cette méthode utilise des convolutions sous-pixel pour générer des détails plus fins et améliorer la résolution de l'image.

De leur côté, [Jo et al., 2018] ont proposé un réseau de super-résolution vidéo profonde qui exploite les corrélations spatiales et temporelles entre les images consécutives, sans nécessiter de compensation de mouvement explicite.

En résumé, ces différentes approches témoignent des progrès significatifs réalisés dans le domaine de la vision par ordinateur. Elles soulignent l'importance des

réseaux à convolution profonde, de l'apprentissage résiduel, des connexions par saut symétrique et des convolutions sous-pixel pour améliorer la résolution, la qualité et la restauration des images et des vidéos. Ces travaux ont lancé de nouvelles perspectives pour les avancées ultérieures dans ces domaines.

1.4 Conclusion

Le premier chapitre de ce mémoire a examiné l'état de l'art dans le domaine de la super-résolution des images. Nous avons constaté que la super-résolution est un domaine de recherche actif qui vise à améliorer la qualité des images en augmentant leur résolution. Une approche prometteuse pour résoudre ce problème est l'utilisation de techniques basées sur l'intelligence artificielle (IA), en particulier l'apprentissage automatique et l'apprentissage profond.

Dans le contexte de l'IA, l'apprentissage automatique est un sous-domaine essentiel qui permet aux machines d'apprendre à partir des données et de prendre des décisions intelligentes. L'apprentissage profond, qui est une branche de l'apprentissage automatique, est devenu particulièrement populaire ces dernières années en raison de ses performances exceptionnelles dans diverses tâches de vision par ordinateur. Les réseaux de neurones convolutifs (CNN) sont l'architecture de base utilisée dans de nombreux modèles d'apprentissage profond, y compris pour la super-résolution des images.

En outre, nous avons étudié les travaux connexes dans le domaine des CNN et des GAN. Les CNN ont été largement utilisés pour résoudre des problèmes de vision par ordinateur, y compris la super-résolution des images. Ils ont montré leur capacité à extraire des caractéristiques pertinentes des images et à améliorer la qualité des images reconstruites. Les GAN, quant à eux, sont des modèles génératifs qui utilisent un réseau discriminatif et un réseau générateur pour produire des images de haute qualité. Ils ont été appliqués avec succès à la super-résolution des images pour générer des images réalistes et détaillées à partir d'images de basse résolution.

En conclusion, ce chapitre a mis en évidence l'importance de la super-résolution des images et son application dans le domaine de l'intelligence artificielle. L'utilisation des CNN et des GAN a montré des résultats prometteurs pour améliorer la résolution des images. Les travaux connexes ont fourni un aperçu des différentes approches et des avancées récentes. Dans le chapitre suivant, nous présenterons notre méthodologie de recherche et détaillerons notre approche proposée pour la super-résolution des images.

La conception d'un CNN spécifique pour la super-résolution d'images faciales

2.1 Introduction :

Dans le domaine de la super-résolution et du traitement des images, le CNN, le GAN, le RNN et l'autoencodeur peuvent tous jouer un rôle important, le CNN possédant une importance particulière dans le contexte de la super-résolution des images faciales. Nous allons donc explorer ici comment le CNN peut être utilisé pour améliorer la super-résolution des images de visages.

Le CNN (Convolutional Neural Network) est souvent utilisé pour extraire les caractéristiques visuelles des images, ce qui en fait un choix naturel pour la super-résolution des images. Le CNN peut apprendre à reconstruire les détails manquants dans une image basse résolution en analysant des paires d'images basse et haute résolution. En se basant sur les caractéristiques visuelles apprises, le CNN peut générer une version haute résolution d'une image de visage à partir de sa version basse résolution correspondante.

2.2 Solution proposée pour la super résolution

Dans cette section, nous décrivons notre approche pour résoudre au mieux notre problème. Tout d'abord, nous commençons par une vue globale, puis nous passons à une description détaillée. Un schéma descriptif de notre système est présenté par la figure 2.1.

Pour concevoir une architecture d'un réseau CNN capable d'atteindre un bon taux de détection (précision) avec un temps d'apprentissage raisonnable, nous devons constamment vérifier le taux de détection et ajuster progressivement certains paramètres pertinents.

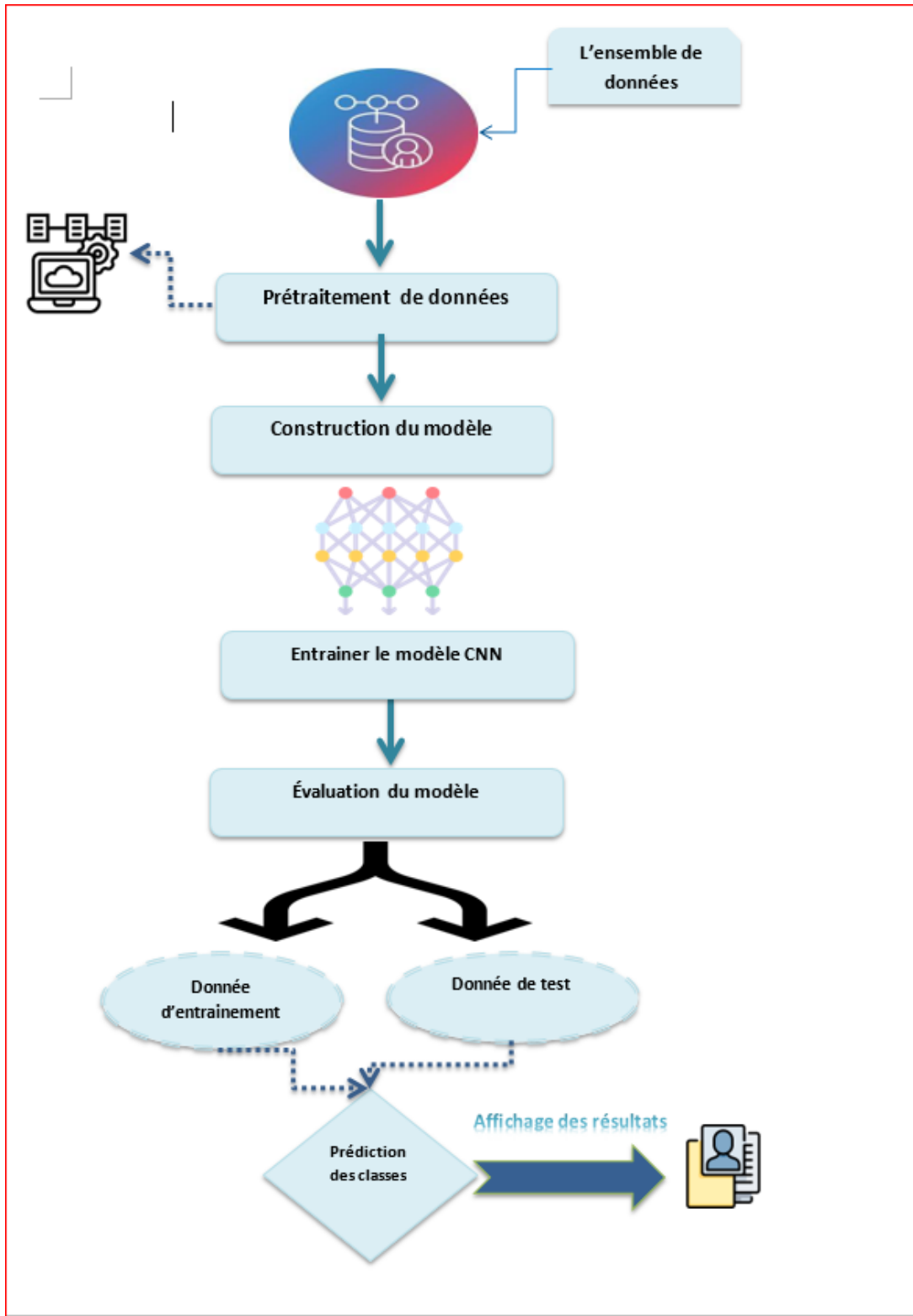


FIGURE 2.1 – L'architecture générale du système proposé.

2.2.1 Hyper paramètre utilisé

L'ensemble d'hyper-paramètres pris en compte et ajustés étape par étape dans ce projet sont :

- *Relatif aux images :*
 - La taille de l'image,
 - Paramètres de prétraitement,
 - Taille de la base de données.
- *Relatif à l'apprentissage :*
 - La taille de l'image,
 - Nombre d'époques (epochs)
 - Taille du lot (batch size),
 - Taux d'apprentissage (learning rate),
 - Critère de mise à jour des poids,
 - Expression de la perte d'apprentissage (training loss),
 - Critère d'initialisation des poids
- *Relatif à l'architecture :*
 - Nombre de couches,
 - Type de couches,
 - Nombre de filtres,

2.2.2 Aperçu général

Cette section présente les grandes lignes de la solution proposée, comme l'illustre la figure???.Le système repose sur six (06) étapes :

1. *Préparation des données :* Les images d'entraînement doivent être préparées sous forme de tenseurs pour être utilisées par le modèle. Cela peut impliquer le redimensionnement des images, la normalisation des valeurs de pixel, etc.
2. *Construction du modèle :* Le modèle CNN est construit à l'aide de la fonction `build model`, qui crée une architecture avec des couches de convolution, des couches d'activation et une couche de sortie.
3. *Compilation du modèle :* Avant de pouvoir entraîner le modèle, il doit être compilé en spécifiant une fonction de perte (loss function), un optimiseur et des métriques pour évaluer les performances du modèle.
4. *Entraînement du modèle :* Le modèle est entraîné en utilisant un ensemble de données d'entraînement. Pendant l'entraînement, les poids du modèle sont ajustés pour minimiser la fonction de perte et améliorer les performances du modèle sur les données d'entraînement.

5. *Évaluation du modèle* : Une fois l'entraînement terminé, le modèle est évalué sur un ensemble de données de validation ou de test pour évaluer ses performances sur des données qu'il n'a pas vues pendant l'entraînement. Cela permet de vérifier si le modèle généralise bien et peut effectuer des prédictions précises sur de nouvelles images.
6. *Prédiction des classes* : Une fois le modèle entraîné et évalué, il peut être utilisé pour effectuer des prédictions sur de nouvelles images. Les images de test ou d'autres images d'entrée sont fournies au modèle, qui produit des scores de classe pour chaque image. La classe prédite est généralement déterminée en prenant la classe avec le score le plus élevé.

2.2.3 Description détaillée de l'approche :

Nous allons procéder ci-dessous à une description détaillée de notre système :

2.2.3.1 Architecture du modèle CNN

L'architecture globale de notre modèle CNN est illustrée à la figure 2.2

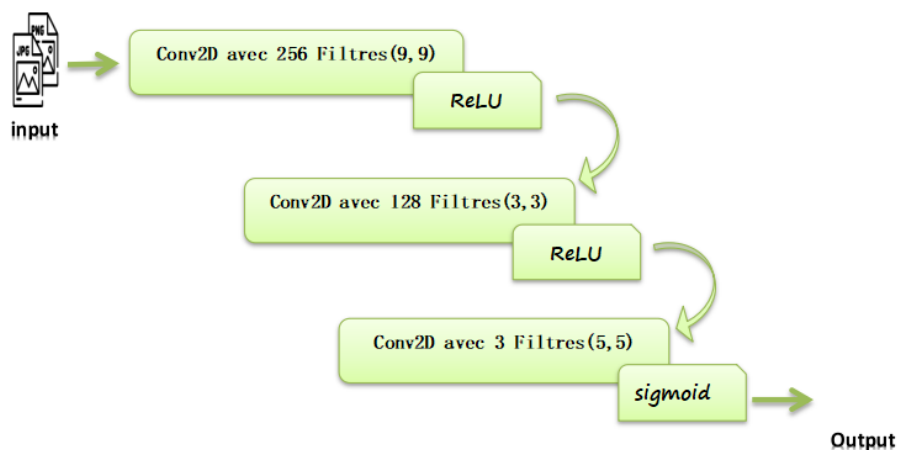


FIGURE 2.2 – L'architecture du modèle CNN.

1. **Couche de convolution** : La première couche ajoutée est une couche de convolution avec 256 filtres et une taille de noyau de (9, 9). Cette couche est utilisée pour extraire des caractéristiques à partir de l'image d'entrée.
2. **Couche d'activation** : Après chaque couche de convolution, une couche d'activation ReLU est ajoutée. La fonction ReLU est couramment utilisée car elle permet d'introduire de la non-linéarité dans le modèle.
3. **Couche de convolution** : Une deuxième couche de convolution est ajoutée avec 128 filtres et une taille de noyau de (3, 3). Cette couche permet de continuer à extraire des caractéristiques plus complexes de l'image.
4. **Couche d'activation** : Une autre couche d'activation ReLU est ajoutée après la deuxième couche de convolution.

5. **Couche de convolution finale** : La dernière couche de convolution est ajoutée avec 3 filtres et une taille de noyau de (5, 5). Cette couche produit les sorties finales du modèle.
6. **Couche d'activation finale** : Une couche d'activation sigmoïde est ajoutée après la dernière couche de convolution. Cela permet de produire des sorties entre 0 et 1, qui peuvent être interprétées comme des probabilités pour une classification binaire. La formule de la fonction *sigmoïde* : $(sigmoid(x) = \frac{1}{1+\exp(-x)})$

La conception du CNN adoptée dans notre application suit un schéma commun pour la classification des images. Les couches de convolution successives extraient progressivement des caractéristiques de plus en plus complexes de l'image. Les couches d'activation **ReLU** introduisent la non-linéarité, ce qui permet au modèle d'apprendre des relations complexes entre les caractéristiques extraites. La dernière couche de convolution produit les sorties finales du modèle, qui sont activées par une couche d'activation sigmoïde pour obtenir les probabilités de classification.

2.2.3.2 Apprentissage du modèle

Un optimiseur est un algorithme ou une méthode utilisée pour modifier les attributs du réseau neuronal tels que les poids et le taux d'apprentissage afin de réduire les pertes. Cet exemple utilise l'optimiseur Adam comme optimiseur. Un optimiseur qui implémente l'algorithme d'Adam [Kingma and Ba, 2014]. Il s'agit d'une méthode **Stochastic Gradient Decent (SGD)** basée sur l'estimation adaptative des premiers et seconds moments. Selon [Kingma and Ba, 2014], la méthode est efficace sur le plan informatique, nécessite peu de mémoire, Le redimensionnement diagonal dégradé fonctionne bien pour les problèmes de données volumineux. Le paramètre le plus important à ce stade est le taux d'apprentissage. Le taux d'apprentissage a été initialisé à 0,001, ce qui est un bon point de départ. la fonction de perte utilisée pour l'entraînement du modèle est l'erreur quadratique moyenne (**Mean Squared Error, MSE**). qui correspond à l'erreur quadratique moyenne. L'erreur quadratique moyenne est couramment utilisée pour les tâches de régression, où l'objectif est d'approximer une valeur continue. De plus, le nombre d'échantillons d'entraînement utilisés dans une itération avant de mettre à jour les poids du modèle. Par défaut, la taille du batch est définie à 64. Le nombre d'époques. Il s'agit du nombre de fois que l'ensemble des données d'entraînement est parcouru lors de l'entraînement du modèle. Par défaut, le nombre d'époques est défini à **3000** (voir la table 2.1).

2.3 Conclusion :

Dans ce chapitre, nous avons abordé le CNN (réseau neuronal convolutif) dans son ensemble, en mettant l'accent sur son architecture et son fonctionnement. Le CNN est une architecture de réseau neuronal profond spécialement conçue pour

Paramètre	Valeur
Training Loss	MSE
Optimiseur	Adam
Learning rate	0.001
Nombre d'epochs	3000
Batch size	64

TABLE 2.1 – Les paramètres d'apprentissage utilisés

le traitement d'images et largement utilisée dans les tâches de vision par ordinateur telles que la classification, la détection d'objets et la segmentation sémantique. Nous avons également décrit le modèle CNN envisagé pour la réalisation d'un système de super-résolution d'images. On a souligné l'importance de l'apprentissage du CNN à l'aide de techniques telles que la rétropropagation du gradient et la régularisation afin de minimiser l'erreur et d'éviter le *sur-ajustement*. Ensuite, le concept de pré-entraînement et de transfert de connaissances a été abordé, ce qui permet d'utiliser des CNN pré-entraînés sur de grandes bases de données pour améliorer les performances sur des tâches similaires avec des ressources limitées. En bref, le CNN est une architecture puissante largement utilisée dans le traitement des images. En exploitant les caractéristiques apprises par le réseau, notre modèle CNN suggère une approche pour la super-résolution d'images afin d'améliorer la qualité des images à haute résolution. Toutefois, il convient de noter que l'utilisation du CNN dans des applications spécifiques nécessite une adaptation et une expérimentation appropriées afin d'obtenir les meilleurs résultats, ce que nous aborderons dans le chapitre suivant.

Implémentation et résultats expérimentaux

3.1 Introduction :

Ce chapitre porte sur les outils de développement et les logiciels utilisés pour le développement des modèles. Pour comprendre l'environnement technique et les ressources utilisées pour le développement et l'évaluation des modèles, il est important de se familiariser avec tous les outils et les logiciels utilisés. Il est donc nécessaire de présenter les outils utilisés, d'expliquer leur importance et de donner une synthèse des résultats obtenus grâce à leur utilisation. Des mesures de performance ont été utilisées pour évaluer la qualité des images reconstruites en utilisant notre modèle, notamment MSE, PSNR et SSIM. Les résultats font l'objet d'une discussion détaillée qui met en évidence l'amélioration significative du modèle par rapport aux techniques traditionnelles de super-résolution d'images.

3.2 Les outils de développement

3.2.1 Google Colab

GOOGLE COLAB est une plateforme en ligne développée par Google qui permet aux utilisateurs d'exécuter et de collaborer sur du code Python via un navigateur web. Il s'agit d'un environnement de développement intégré (IDE) basé sur le cloud qui offre un accès gratuit à une machine virtuelle (VM) préconfigurée pour l'exécution de tâches d'apprentissage automatique (machine learning), d'analyse de données et de création de modèles.

GOOGLE COLAB permet aux utilisateurs d'écrire et d'exécuter du code Python, de compiler et d'exécuter des notebooks Jupyter, et de partager facilement leur travail avec d'autres personnes. Les utilisateurs peuvent accéder aux ressources informatiques puissantes de Google, y compris les processeurs graphiques (GPU) et les unités de traitement tensoriel (TPU), pour accélérer leurs calculs et leurs tâches intensives en ressources [Carneiro et al., 2018].

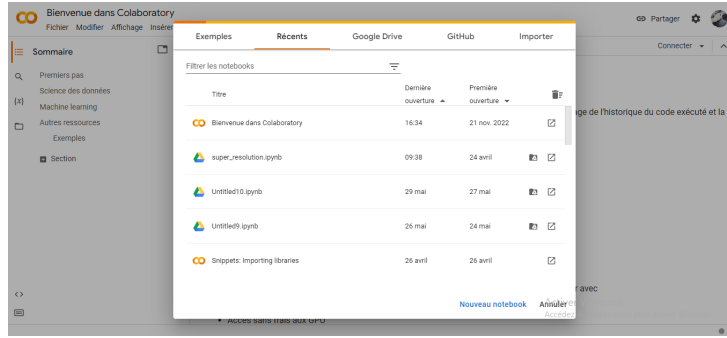


FIGURE 3.1 – Environnement du Google Colab [Carneiro et al., 2018]

3.2.2 Python

Python est un langage de programmation puissant et facile à apprendre. Il été développé depuis 1989 par GuidoVan Rossumet .Il présente des structures de données de haut niveau et permet une approche simple et efficace de la programmation orientée objet. La syntaxe élégante, le typage dynamique et l’interprétabilité de Python en font un langage idéal pour les scripts et le développement rapide d’applications dans de nombreux domaines et sur la plupart des plates-formes. Il dispose d’une bibliothèque standard étendue et complète [Howse and Minichino, 2020].



FIGURE 3.2 – Logo Python [Howse and Minichino, 2020]

3.3 Les bibliothèques utilisées

3.3.1 Keras

Keras est une API d’apprentissage en profondeur écrite en Python et exécutée sur la plate-forme d’apprentissage automatique TensorFlow. Conçu dans le but de permettre une expérimentation rapide. La clé d’une bonne recherche est de passer de l’idée aux résultats le plus rapidement possible [Kim et al., 2022]



FIGURE 3.3 – Logo Keras [Kim et al., 2022]

3.3.2 Scikit-image

Scikit-image est une collection d'algorithmes de traitement d'images, gratuit et illimité [Van der Walt et al., 2014]. Nous sommes fiers d'un code de haute qualité, évalué par des pairs, écrit par une communauté active de bénévoles [Kuwahara et al., 2023].



FIGURE 3.4 – Logo [Kuwahara et al., 2023]

3.3.3 OpenCv

OpenCV est une énorme bibliothèque open source de vision par ordinateur, d'apprentissage automatique et de traitement d'images [Howse and Minichino, 2020]. OpenCV prend en charge divers langages de programmation tels que Python, C++ et Java. Les images et les vidéos peuvent être traitées pour identifier des objets, des visages et même l'écriture humaine. Intégrez-vous à diverses bibliothèques [Intel, 2000]



FIGURE 3.5 – Logo OpenCV [Intel, 2000]

3.3.4 Numpy

Numpy est un package de calcul scientifique de base pour Python. Il offre un objet de type tableau multidimensionnel, divers objets dérivés (comme les tableaux masqués et les matrices) et un ensemble d'opérations rapides sur les tableaux telles que les mathématiques, la logique, la manipulation de formes, le tri, la sélection...etc

Cette bibliothèque Python comprend des routines pour les transformées de Fourier discrètes, l'algèbre linéaire de base, les opérations statistiques de base, les simulations aléatoires,.. etc [Andersen, 2022].



FIGURE 3.6 – Logo Numpy [Andersen, 2022]

3.3.5 Matplotlib

Matplotlib est une bibliothèque complète permettant de créer des visualisations statiques, animées et interactives en Python [Nelli, 2015]

3.3.6 Conv2D

Cette couche crée un noyau de convolution qui fonctionne avec la couche d'entrée pour produire le tenseur de sortie. Si cela est vrai, un vecteur de biais est créé et ajouté à la sortie. Sinon, cela s'appliquera également à Outputs [Chenebuah et al., 2023].

Si vous souhaitez utiliser ce niveau comme premier niveau dans le modèle, spécifiez un argument de mot-clé (tuple entier ou , sans inclure les exemples d'axe).

3.3.7 Adam

Un optimiseur qui implémente l'algorithme d'Adam est une descente de gradient stochastique basée sur une estimation adaptative des premiers et deuxièmes moments [Jais et al., 2019]

3.4 Résultats expérimentaux et discussions

Le système que nous avons implémenté est exécuté dans l'environnement de programmation sur un ordinateur doté d'un processeur Intel(R) Core(TM) i5-4200U CPU @ 1.60GHz 2.29 GHz et d'une capacité de mémoire de 4 Go.

3.4.1 Données d'expérimentation et les métriques d'évaluation de qualité

L'architecture du CNN proposé est illustrée par la figure 3.8 suivante :

3.4.1.1 Base de données

Le modèle sera entraîné sur un ensemble d'images sélectionnées dans la base de données Labeled Faces in the Wild (LFW) [Learned-Miller et al., 2016]. Cette

```

model_1
Model: "model_1"

```

Layer (type)	Output Shape	Param #
conv2d_6 (Conv2D)	(None, 64, 64, 256)	62464
activation_6 (Activation)	(None, 64, 64, 256)	0
conv2d_7 (Conv2D)	(None, 64, 64, 128)	295040
activation_7 (Activation)	(None, 64, 64, 128)	0
conv2d_8 (Conv2D)	(None, 64, 64, 3)	9603
activation_8 (Activation)	(None, 64, 64, 3)	0

```

Total params: 367,107
Trainable params: 367,107
Non-trainable params: 0

```

FIGURE 3.7 – Aperçu de l’architecture

dernière contient 13233 images de visages. La base de données Labeled Faces in the Wild (LFW) ne contient pas de classes prédéfinies et a été créée pour fournir un ensemble de données standardisées pour l’évaluation des systèmes de reconnaissance faciale dans des conditions réelles, y compris des variations d’éclairage, d’expression faciale, de pose, d’âge, de maquillage, d’accessoires,.. etc.

3.4.1.2 Les métriques d’évaluation de qualité

- **Erreur quadratique moyenne MSE** : Le Mean Squared Error (MSE) est une mesure qui calcule la moyenne des écarts au carré entre les valeurs prédites par un modèle et les valeurs réelles. Il est couramment utilisé pour évaluer la précision d’un modèle de régression en quantifiant l’erreur moyenne au carré [Girod, 1993].

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{\text{pred}_i} - y_{\text{true}_i})^2 \quad (3.1)$$

Où **SSIM** : l’indice de similarité structurale; μ_x, μ_y : Moyennes des blocs de l’image; σ_x, σ_y : Écarts-types des blocs de l’image; σ_{xy} : Covariances des blocs de l’image; $c1, c2$: les constantes pour la stabilité numérique

- **Rapport signal sur bruit de crête PSNR**

Le Peak Signal-to-Noise Ratio (PSNR) est une mesure de la qualité d’une image ou d’un signal, qui évalue le rapport entre la puissance du signal et la puissance du bruit. Il permet de comparer la qualité d’une image dégradée à l’image d’origine et est exprimé en décibels (dB). Un PSNR élevé indique une meilleure qualité de l’image ou du signal [Wang et al., 2002].

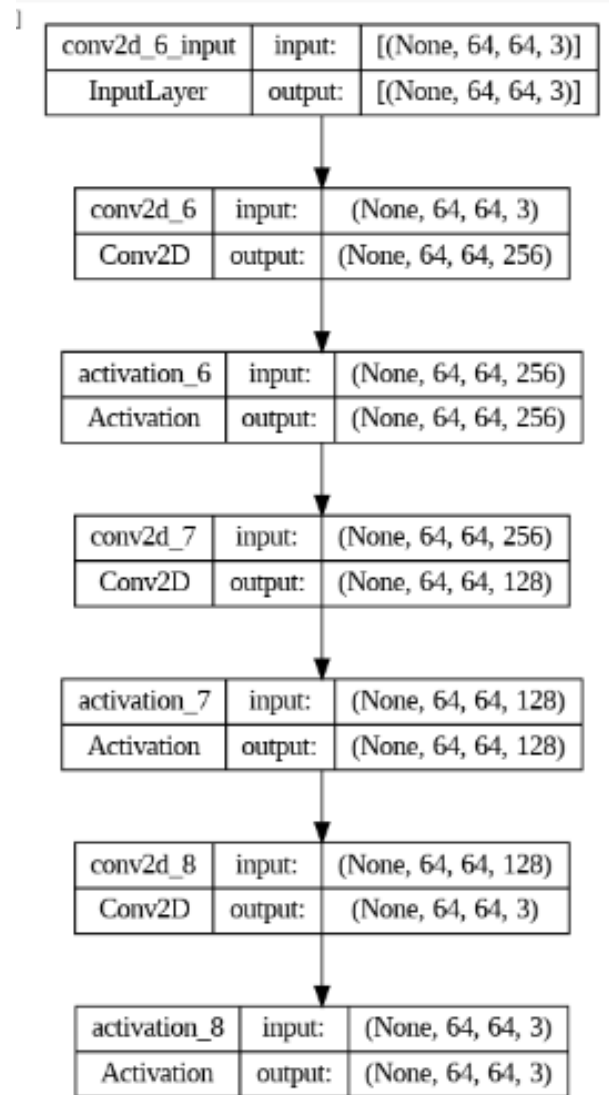


FIGURE 3.8 – Les étapes de l’architecture du CNN proposée pour la super-résolution

$$\text{PSNR} = 20 \cdot \log_{10}(\text{max_val}) - 10 \cdot \log_{10}(\text{MSE}) \quad (3.2)$$

Où **MAX** : Valeur maximale possible pour les données; **MSE** : Erreur quadratique moyenne

- **Similarité structurale SSIM** :Le Structural Similarity Index (SSIM) est une mesure de la similarité entre deux images qui prend en compte la luminance, le contraste et la structure spatiale. Il permet d’évaluer dans quelle mesure les détails structurels d’une image sont préservés lors d’une compression ou d’une transformation. Le SSIM varie de -1 à 1, où une valeur de 1 indique une similitude parfaite entre les images [Wang et al., 2004]

$$\text{SSIM} = \frac{(2 \cdot \mu_x \cdot \mu_y + c1) \cdot (2 \cdot \sigma_{xy} + c2)}{(\mu_x^2 + \mu_y^2 + c1) \cdot (\sigma_x^2 + \sigma_y^2 + c2)} \quad (3.3)$$

Où μ_x, μ_y : sont les moyennes des blocs de l'image ; σ_x, σ_y : sont les écarts-types des blocs de l'image ; σ_{xy} : est la covariance des blocs de l'image ; $c1, c2$: sont des constantes utilisées pour assurer la stabilité numérique.

3.4.2 Tests et discussions

Sur la base d'images redimensionnées à 64x64, nous obtenons un taux égal à 92,03 sur les données de test en environ 3000 itérations (c.-à-d. epochs= 3000). Les résultats de l'exécution des dernières itérations (epochs) sont illustrés dans la figure 3.9.

```

24/24 [=====] - 4s 150ms/step - loss: 8.9737e-04 - val_loss: 8.8868e-04
Epoch 2988/3000
24/24 [=====] - 3s 142ms/step - loss: 9.0185e-04 - val_loss: 8.7945e-04
Epoch 2989/3000
24/24 [=====] - 3s 143ms/step - loss: 9.0436e-04 - val_loss: 8.9495e-04
Epoch 2990/3000
24/24 [=====] - 3s 141ms/step - loss: 9.4671e-04 - val_loss: 8.9517e-04
Epoch 2991/3000
24/24 [=====] - 3s 146ms/step - loss: 9.1587e-04 - val_loss: 8.9839e-04
Epoch 2992/3000
24/24 [=====] - 3s 145ms/step - loss: 9.0590e-04 - val_loss: 8.9716e-04
Epoch 2993/3000
24/24 [=====] - 3s 141ms/step - loss: 9.1269e-04 - val_loss: 8.9546e-04
Epoch 2994/3000
24/24 [=====] - 4s 149ms/step - loss: 9.0121e-04 - val_loss: 8.9276e-04
Epoch 2995/3000
24/24 [=====] - 4s 151ms/step - loss: 9.0091e-04 - val_loss: 9.0348e-04
Epoch 2996/3000
24/24 [=====] - 3s 141ms/step - loss: 9.0109e-04 - val_loss: 8.9709e-04
Epoch 2997/3000
24/24 [=====] - 4s 147ms/step - loss: 8.9353e-04 - val_loss: 8.8130e-04
Epoch 2998/3000
24/24 [=====] - 4s 148ms/step - loss: 8.9458e-04 - val_loss: 8.8175e-04
Epoch 2999/3000
24/24 [=====] - 4s 150ms/step - loss: 9.0027e-04 - val_loss: 8.8986e-04
Epoch 3000/3000
24/24 [=====] - 3s 146ms/step - loss: 9.0384e-04 - val_loss: 8.8665e-04

```

FIGURE 3.9 – Résultats de l'apprentissage par epochs

La figure 3.9 montre que le l'erreur quadratique moyenne (MSE) de quelque images avec le modèle implémenter, ici le MSE est égale **0,0032** qu'est calculer avec la fonction de calcule MSE.

```

MSE: 0.0031073589061926055
RMSE: 0.05574368938447298
Coefficient of Determination (R2): 0.9587655936239317
Mean SSIM: 0.925441936821124
PSNR: 25.076085819632688

```

FIGURE 3.10 – Les résultats obtenues

Nos résultats montrent que le modèle fonctionne bien. Les faibles valeurs MSE et RMSE indiquent que les prédictions de grade sont proches des valeurs réelles. Un R2 élevé suggère que le modèle explique une grande partie de la variance des données. En outre, un SSIM moyen élevé et un PSNR élevé indiquent une meilleure qualité de l'image prédite par rapport à l'image réelle. En combinant

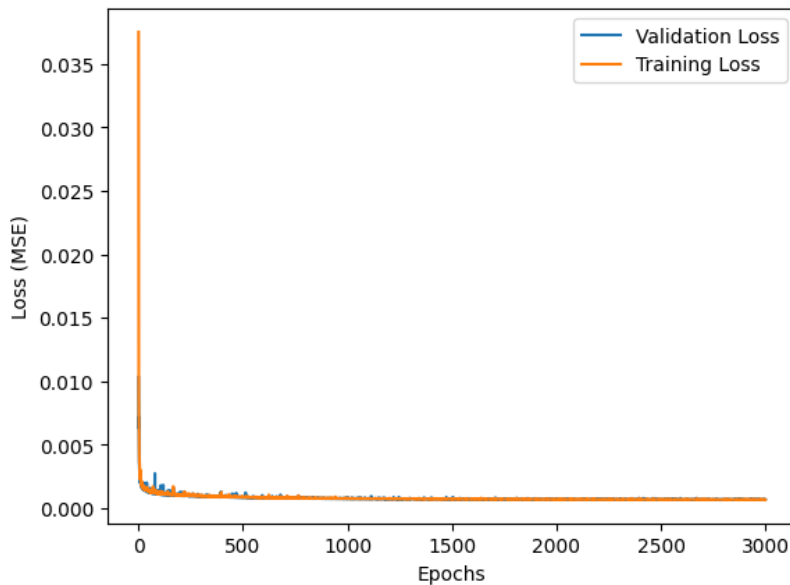


FIGURE 3.11 – La représentation graphique de l'évolution de l'erreur

ces valeurs, on peut obtenir une indication claire de la performance du modèle.

La courbe (**La figure 3.11**) représentant le nombre d'époques (epochs) sur l'axe des abscisses(x) et l'erreur quadratique moyenne (MSE) sur l'axe des ordonnées (y) est une représentation graphique de l'évolution de l'erreur du modèle au fil des itérations d'apprentissage.

L'axe des abscisses x représente le nombre d'époques, qui correspond au nombre de fois où le modèle a parcouru l'ensemble des données d'entraînement lors de l'apprentissage. Chaque époque consiste en une série d'exécutions du modèle sur les données d'apprentissage, au cours desquelles les poids du modèle sont ajustés pour minimiser l'erreur.

L'axe des ordonnées représente l'erreur quadratique moyenne (MSE). Le MSE est une mesure de l'écart quadratique moyen entre les valeurs prédites par le modèle et les valeurs réelles de la variable cible. Il est calculé en faisant la moyenne des carrés des différences entre les valeurs prédites et les valeurs réelles.

La courbe MSE vs Epochs permet de visualiser comment l'erreur du modèle évolue au fur et à mesure de son apprentissage. En général, on s'attend à ce que l'erreur diminue progressivement à chaque époque, car le modèle ajuste ses paramètres pour mieux s'adapter aux données. De plus la courbe obtenue est une courbe MSE décroissante qui indique une amélioration de la performance du modèle au fur et à mesure de l'apprentissage. Cependant, il est important de trouver le bon équilibre entre l'apprentissage et la généralisation, en évitant un ajustement excessif aux données d'entraînement.

Après avoir défini notre modèle, il nous suffit de l'appliquer à certaines de nos

images d'entrée pour effectuer la super-résolution et calculer le PSNR et le SSIM pour chacune d'entre elles (un exemple est donné par la table 3.1).

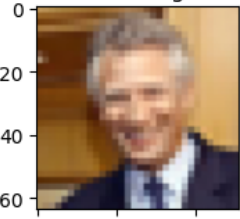

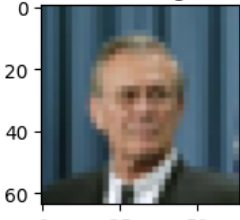

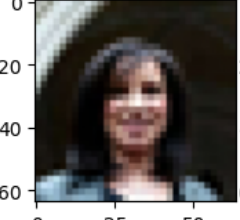

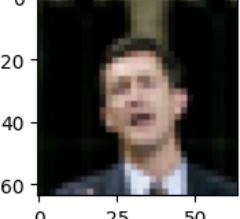
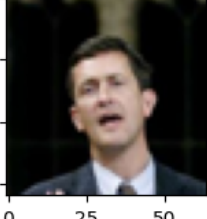
Image	Les paramètres	PSNR	SSIM	résultat
Image 1		72.64	0.86	<div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <p>Low image</p>  </div> <div style="text-align: center;"> <p>High image</p>  </div> </div>
Image 2		75.15	0.92	<div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <p>Low image</p>  </div> <div style="text-align: center;"> <p>High image</p>  </div> </div>
Image 3		72.13	0.91	<div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <p>Low image</p>  </div> <div style="text-align: center;"> <p>High image</p>  </div> </div>
Image 4		74.21	0.93	<div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <p>Low image</p>  </div> <div style="text-align: center;"> <p>High image</p>  </div> </div>

TABLE 3.1 – Les exemples de résultats obtenus sur quelques images

Les résultats PSNR et SSIM peuvent varier d'une image à l'autre. Ces mesures sont utilisées pour évaluer la qualité de la super-résolution d'une image en comparant la différence entre l'image super-résolue et l'image de référence à plus haute résolution.

3.5 Conclusion

Dans ce chapitre, nous avons fait le point sur les outils de développement et les logiciels utilisés pour explorer les modèles. Nous avons également procédé à une évaluation détaillée des performances liées au modèle à l'aide de certains indicateurs de performance. Cette évaluation a permis de constater l'efficacité du modèle dans le cadre de la super-résolution d'images. Les résultats expérimentaux obtenus sur les ensembles de données bien connus montrent que l'approche proposée apporte une valeur ajoutée au domaine et qu'elle est plus performante que les méthodes existantes.

Les résultats obtenus constituent une bonne approche pour la progression de l'étude et de la mise en œuvre des modèles dans les applications réelles de traitement d'images à haute résolution.

Conclusion générale

Ce travail de recherche vise à présenter nos résultats dans le domaine de la super-résolution d'images, en utilisant une approche basée sur des techniques d'apprentissage profond, en particulier le CNN. Nous avons fait le point sur nos principales contributions et nous avons indiqué nos futurs travaux de recherche, aussi bien que les perspectives ouvertes par notre approche.

Nous avons élaboré un système de super-résolution d'images qui rend possible la conversion d'une image à basse résolution en une image à haute résolution. Notre approche est basée sur l'utilisation d'un CNN, une technique d'apprentissage profond largement utilisée pour résoudre les problèmes de vision par ordinateur.

Nous envisageons plusieurs perspectives prometteuses pour les travaux de recherche à venir. Tout d'abord, nous prévoyons d'explorer l'utilisation des techniques GAN (generative adversarial network) pour améliorer encore la qualité des images générées par notre système. En effet, les GAN offrent des possibilités intéressantes pour affiner les détails et rendre les images plus réalistes.

En plus de ces travaux, notre approche devrait être étendue à d'autres types de données, comme les vidéos, afin de résoudre le problème de la super-résolution des vidéos. Cela nécessitera une adaptation de notre modèle pour prendre en compte la temporalité et les spécificités des séquences d'images.

Nous comptons en dernier lieu poursuivre nos recherches sur l'amélioration des mécanismes d'attention dans notre modèle. Les mécanismes d'attention permettent à l'apprentissage de se concentrer sur les parties les plus importantes de l'image, ce qui peut conduire à de meilleurs résultats de super-résolution.

Bibliographie

- [Abdel-Hamid et al., 2014] Abdel-Hamid, O., Mohamed, A.-r., Jiang, H., Deng, L., Penn, G., and Yu, D. (2014). Convolutional neural networks for speech recognition. *IEEE/ACM Transactions on audio, speech, and language processing*, 22(10) :1533–1545.
- [Andersen, 2022] Andersen, E. (2022). Imagedata : A python library to handle medical image data in numpy array subclass series. *Journal of Open Source Software*, 7(73) :4133.
- [Bengio et al., 2006] Bengio, Y., Lamblin, P., Popovici, D., and Larochelle, H. (2006). Greedy layer-wise training of deep networks. *Advances in neural information processing systems*, 19.
- [Bishop and Nasrabadi, 2006] Bishop, C. M. and Nasrabadi, N. M. (2006). *Pattern recognition and machine learning*, volume 4. Springer.
- [Carneiro et al., 2018] Carneiro, T., Da Nóbrega, R. V. M., Nepomuceno, T., Bian, G.-B., De Albuquerque, V. H. C., and Reboucas Filho, P. P. (2018). Performance analysis of google colab as a tool for accelerating deep learning applications. *IEEE Access*, 6 :61677–61685.
- [Chenebuah et al., 2023] Chenebuah, E. T., Nganbe, M., and Tchagang, A. B. (2023). A fourier-transformed feature engineering design for predicting ternary perovskite properties by coupling a two-dimensional convolutional neural network with a support vector machine (conv2d-svm). *Materials Research Express*, 10(2) :026301.
- [Chinnasamy et al., 2023] Chinnasamy, P., Sathya, K. S., Jebamani, B. J. A., Nithyasri, A., and Fowjiya, S. (2023). Deep learning : Algorithms, techniques, and applications—a systematic survey. In *Deep Learning Research Applications for Natural Language Processing*, pages 1–17. IGI Global.
- [Cho et al., 2014] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv :1406.1078*.
- [Chollet, 2021] Chollet, F. (2021). *Deep learning with Python*. Simon and Schuster.
- [Dong et al., 2015] Dong, C., Loy, C. C., He, K., and Tang, X. (2015). Image super-resolution using deep convolutional networks.

- [Girod, 1993] Girod, B. (1993). What's wrong with mean-squared error? *Digital images and human vision*, pages 207–220.
- [Glorot et al., 2011] Glorot, X., Bordes, A., and Bengio, Y. (2011). Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 315–323. JMLR Workshop and Conference Proceedings.
- [Goller and Kuchler, 1996] Goller, C. and Kuchler, A. (1996). Learning task-dependent distributed representations by backpropagation through structure. In *Proceedings of International Conference on Neural Networks (ICNN'96)*, volume 1, pages 347–352. IEEE.
- [Goodfellow et al., 2016] Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.
- [Goodfellow et al., 2014] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets advances in neural information processing systems. *arXiv preprint arXiv :1406.2661*.
- [Graves and Graves, 2012] Graves, A. and Graves, A. (2012). *Supervised sequence labelling*. Springer.
- [He et al., 2015] He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition.
- [Hinton et al., 2012] Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N., et al. (2012). Deep neural networks for acoustic modeling in speech recognition : The shared views of four research groups. *IEEE Signal processing magazine*, 29(6) :82–97.
- [Hinton, 2009] Hinton, G. E. (2009). Deep belief networks. *Scholarpedia*, 4(5) :5947.
- [Hinton and Salakhutdinov, 2012] Hinton, G. E. and Salakhutdinov, R. R. (2012). A better way to pretrain deep boltzmann machines. *Advances in Neural Information Processing Systems*, 25.
- [Howse and Minichino, 2020] Howse, J. and Minichino, J. (2020). *Learning OpenCV 4 Computer Vision with Python 3 : Get to grips with tools, techniques, and algorithms for computer vision and machine learning*. Packt Publishing Ltd.
- [Huang et al., 2017] Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2261–2269.
- [Intel, 2000] Intel (2000). Opencv - open computer vision library. <http://OpenCV.org>. Accessed on DEC, 2022.
- [Jais et al., 2019] Jais, I. K. M., Ismail, A. R., and Nisa, S. Q. (2019). Adam optimization algorithm for wide and deep neural network. *Knowledge Engineering and Data Science*, 2(1) :41–46.
- [J.M Field,] J.M Field, year = 2023, n. . A. h. u. Jm.field marketing company.
- [Jo et al., 2018] Jo, Y., Oh, S. W., Kang, J., and Kim, S. J. (2018). Deep video super-resolution network using dynamic upsampling filters without explicit motion

- compensation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3224–3232.
- [Kim et al., 2022] Kim, S., Wimmer, H., and Kim, J. (2022). Analysis of deep learning libraries : Keras, pytorch, and mxnet. In *2022 IEEE/ACIS 20th International Conference on Software Engineering Research, Management and Applications (SERA)*, pages 54–62. IEEE.
- [Kingma and Ba, 2014] Kingma, D. P. and Ba, J. (2014). Adam : A method for stochastic optimization. *arXiv preprint arXiv :1412.6980*.
- [Kingma and Welling, 2013] Kingma, D. P. and Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv :1312.6114*.
- [Krizhevsky et al., 2012] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60 :84 – 90.
- [Krizhevsky et al., 2017] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6) :84–90.
- [Kuwahara et al., 2023] Kuwahara, M., Fujima, J., Takahashi, K., and Takahashi, L. (2023). Improving scientific image processing accessibility through development of graphical user interfaces for scikit-image. *Digital Discovery*.
- [Learned-Miller et al., 2016] Learned-Miller, E., Huang, G. B., RoyChowdhury, A., Li, H., and Hua, G. (2016). Labeled faces in the wild : A survey. *Advances in face detection and facial image analysis*, pages 189–248.
- [LeCun et al., 1995] LeCun, Y., Bengio, Y., et al. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10) :1995.
- [LeCun et al., 2015] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553) :436–444.
- [LeCun et al., 1998a] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998a). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11) :2278–2324.
- [LeCun et al., 1998b] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998b). Gradient-based learning applied to document recognition. *Proceedings of the Institute of Radio Engineers*, 86(11) :2278–2323.
- [Ledig et al., 2017] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690.
- [Li and Wu, 2015] Li, X. and Wu, X. (2015). Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition. In *2015 IEEE international conference on acoustics, speech and signal processing (icassp)*, pages 4520–4524. IEEE.

- [Luo et al., 2022] Luo, Z., Huang, Y., Li, S., Wang, L., and Tan, T. (2022). Learning the degradation distribution for blind image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6063–6072.
- [Mahesh, 2020] Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], 9 :381–386.
- [Mao et al., 2016] Mao, X.-J., Shen, C., and Yang, Y.-B. (2016). Image restoration using convolutional auto-encoders with symmetric skip connections. *arXiv preprint arXiv :1606.08921*.
- [Milanfar, 2017] Milanfar, P. (2017). *Super-resolution imaging*. CRC press.
- [Mishra et al., 2018] Mishra, V., AGARWAL, S., and PURI, N. (2018). Comprehensive and comparative analysis of neural network. *INTERNATIONAL JOURNAL OF COMPUTER APPLICATION*, 2.
- [Nelli, 2015] Nelli, F. (2015). *Python data analytics : Data analysis and science using PANDAs, Matplotlib and the Python Programming Language*. Apress.
- [Pouyanfar et al., 2018] Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M. P., Shyu, M.-L., Chen, S.-C., and Iyengar, S. S. (2018). A survey on deep learning : Algorithms, techniques, and applications. *ACM Computing Surveys (CSUR)*, 51(5) :1–36.
- [Prasad Lokulwar, 2022] Prasad Lokulwar, Basant Verma, N. T. K. K. M. B. D. S. (2022). *Machine learning methods for engineering application development*. Bentham Science Publishers. ISBN : 9815079190,9789815079197.
- [Radford et al., 2015] Radford, A., Metz, L., and Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv :1511.06434*.
- [Rumelhart et al., 1986] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088) :533–536.
- [Shi et al., 2016] Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A. P., Bishop, R., Rueckert, D., and Wang, Z. (2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1874–1883.
- [Shocher et al., 2017] Shocher, A., Cohen, N., and Irani, M. (2017). "zero-shot" super-resolution using deep internal learning.
- [Simonyan and Zisserman, 2015] Simonyan, K. and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition.
- [Socher et al., 2011] Socher, R., Lin, C. C., Manning, C., and Ng, A. Y. (2011). Parsing natural scenes and natural language with recursive neural networks. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 129–136.
- [Stewart, 2019] Stewart, M. (2019). Comprehensive introduction to autoencoders. <https://towardsdatascience.com/>. Towards Data Science.

- [Sutton and Barto, 2018] Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning : An introduction*. MIT press.
- [Szegedy et al., 2015a] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015a). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9.
- [Szegedy et al., 2015b] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015b). Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–9.
- [Szeliski, 2022] Szeliski, R. (2022). *Computer vision : algorithms and applications*. Springer Nature.
- [Tai et al., 2017] Tai, Y., Yang, J., Liu, X., and Xu, C. (2017). Memnet : A persistent memory network for image restoration.
- [Van der Walt et al., 2014] Van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., and Yu, T. (2014). scikit-image : image processing in python. *PeerJ*, 2 :e453.
- [Vint et al., 2021] Vint, D., Anderson, M., Yang, Y., Ilioudis, C., Di Caterina, G., and Clemente, C. (2021). Automatic target recognition for low resolution foliage penetrating sar images using cnns and gans. *Remote Sensing*, 13 :596.
- [Wang et al., 2018] Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., and Change Loy, C. (2018). Esrgan : Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops*, pages 0–0.
- [Wang et al., 2002] Wang, Z., Bovik, A. C., and Lu, L. (2002). Why is image quality assessment so difficult? In *2002 IEEE International conference on acoustics, speech, and signal processing*, volume 4, pages IV–3313. IEEE.
- [Wang et al., 2004] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment : from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4) :600–612.
- [Yue et al., 2022] Yue, Z., Zhao, Q., Xie, J., Zhang, L., Meng, D., and Wong, K.-Y. K. (2022). Blind image super-resolution with elaborate degradation modeling on noise and kernel. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2128–2138.
- [Zhang et al., 2017] Zhang, K., Zuo, W., Chen, Y., Meng, D., and Zhang, L. (2017). Beyond a gaussian denoiser : Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, 26(7) :3142–3155.
- [Zhang, 2010] Zhang, Y. (2010). Application of machine learning. In-TehOlajnica 19/2, 32000 Vukovar, Croatia. ISBN 978-953-307-035-3.
- [Zhang et al., 2018] Zhang, Y., Tian, Y., Kong, Y., Zhong, B., and Fu, Y. (2018). Residual dense network for image super-resolution.
- [Zhang and Wallace, 2015] Zhang, Y. and Wallace, B. (2015). A sensitivity analysis of (and practitioners’ guide to) convolutional neural networks for sentence classification. *arXiv preprint arXiv :1510.03820*.