

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne démocratique et populaire

وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement supérieur et de la recherche scientifique

جامعة سعد دحلب بليدة
Université SAAD DAHLAB de BLIDA

كلية التكنولوجيا
Faculté de Technologie

قسم الإلكترونيك
Département d'Électronique



Mémoire de Master

Mention Électronique

Systèmes de Télécommunications (ST)

Présenté par

Tetbirt abdessamed

&

Khelifi douniazad

Fusion d'images Basée sur l'apprentissage profond

Proposé par :

Dr. Bendoumia rèdha

M. Talbi farid

Dr. Alim-Fatiha

Année Universitaire 2020-2021.

ملخص:

مشروعنا هو تقديم وتقييم تقنيات دمج الصور المبنية على التعلم العميق في تطبيقات الاستشعار عن بعد. يتم تعريف هذا الاندماج بالاسم pansharpening الذي يهدف لاستخدام صورة بانوكروماتية ذات دقة مكانية عالية وصورة متعددة الاطيف بدقة طيفية عالية ودقة مكانية منخفضة للحصول على صورة متعددة الاطيف ذات جودة مكانية وطيفية عالية. يركز هذا العمل على تنفيذ تقنية PanColorGAN التي تعتمد على التعلم العميق باستخدام النظام بايثون والنظام ما تلاب. اظهرت النتائج التي تما الحصول عليها ان تقنيات دمج الصور المبنية على التعلم العميق تعطي تحسنا في جودة وكفاءة الاندماج معى اداء مستقر مقارنة معى تقنيات دمج الصور الكلاسيكية.

Résumé :

Notre projet consiste à présenter et évalué, les techniques de fusion d'images basées sur apprentissage profond dans les applications de télédétection

Cette fusion est défini par le nom pansharpening qui vise à utiliser image panchromatique à haute résolution spatiale et multispectrale à une résolution spectrale élevée et résolution spatiale réduit, pour fusionner Image multispectrale haute résolution spatiale et spectrale.

Ce travail port sur l'implémentation de technique PanColorGAN qui est basée sur l'apprentissage profond, en utilisant l'environnement de Matlab et python.

Les résultats obtenus montrent que les techniques de fusion a la base de l'apprentissage profond donne une amélioration de qualité et l'efficacité de fusion avec performance stable par rapport les diverses méthodes classiques de fusion.

Abstract:

Our project is to present and evaluate, image fusion techniques based on deep learning in remote sensing applications

This fusion is defined by the name pansharpening which aims to use panchromatic image with high spatial resolution and multispectral at high spectral resolution and reduced spatial resolution, to merge multispectral image with high spatial and spectral resolution.

This work focuses on the implementation of the PanColorGAN technique which is based on deep learning, using the intoxication of Matlab and python.

The results obtained show that fusion techniques based on deep learning give an improvement in fusion quality and efficiency with stable performance compared to the various classical fusion methods.

Dédicace

Je dédie ce travail

A mes très chers parents dieu les gardes

A mon frère unique

A tous les membres de la famille

*A mes camarades, qui ont partagé avec moi mes joies et les peines des années
durant A tous ceux qui m'ont aidé de proche ou de loin*

Dédicace

*A mon père Taher, décédé, son rêve était de me voir terminer mes études et porter
mon diplôme entre mes mains, mais le destin n'a pas voulu*

*Aucun mot ni aucune description ne peut exprimer l'amour, l'appréciation, le
dévouement et le respect que j'avais pour toi, Père, pour les encouragements et les
sacrifices que tu as faits pour mon éducation.*

Merci papa et merci papa et merci papa

*A ma chère maman Zahra, tu es une source de soutien qui n'a jamais cessé de
m'encourager et de prier pour moi. Tu as fait plus qu'une mère peut faire pour que
ses enfants suivent le bon chemin dans leur vie et leurs études. Je vous dédie ce
travail en témoignage de mon amour profond. Que Dieu tout puissant vous
protège et vous accorde santé, longue vie et bonheur*

*A mes chers frères Zakria, Ishak et Adel qui sont toujours présents à tous mes
examens avec un soutien moral, je vous souhaite un avenir plein de joie, de
bonheur, de réussite et de sérénité.*

*À mes amis proches Ramadan, souhaib, Ishak et Azzedin, merci de m'avoir
toujours fait sentir que j'ai tissé une source de force dans ma vie*

Remerciements

Nous remercions tout d'abord, (الله سبحانه و تعالی) qui nous a donné la force et le courage afin de parvenir à élaborer ce modeste travail.

Nous tenons à exprimer toute nous reconnaissance à monsieur talbi farid et Bendoumia rédha. Et madame Alime nous les remercions pour leurs conseils, leur aide et leurs conseils.

J'adresse mes sincères remerciements à tous les professeurs, et toutes les personnes qui par leurs paroles, leurs écrits, leurs conseils

Nous remercions nos très chers parents, qui ont toujours été là pour nous.

À tous ces intervenants, je présente mes remerciements, mon respect et ma gratitude.

Table de matières

Introduction Générale	8
Chapitre01 : Notions fondamentale de la télédétection.....	14
1 Chapitre01 : Notions fondamentale de la télédétection :.....	15
1.1 Introduction	15
1.2 Télédétection	15
1.2.1 Différentes étapes de la télédétection.....	15
1.3 Rayonnement électromagnétique.....	16
1.3.1 Le spectre électromagnétique	18
1.4 Système d'acquisition.....	21
1.4.1 Les capteurs	21
1.4.2 Les plateformes	22
1.4.3 Caractéristiques des satellites	23
1.5 Application de la télédétection.....	26
1.5.1 La foresterie	26
1.5.2 La géologie	27
1.6 Conclusion :.....	28
Chapitre02 : Notions fondamentale de la télédétection.....	29
2 Chapitre 02 : Traitement d'images satellitaires	30
2.1 Introduction	30
2.2 Image satellitaire	30
2.3 La formation d'image satellitaire.....	30
2.4 Les résolutions des images satellitaires	32
2.5 Définition de la fusion d'image	32
2.6 Application de fusion image	33
2.6.1 L'aide à la navigation :	33
2.6.2 La télédétection :	33
2.7 Classification spectrale	33
2.8 Classification Spatiale	35
2.9 Les Méthodes de fusion d'image	36
2.9.1 Méthode de la transformée en ondelettes	36
2.9.2 Méthode ARSIS	36

2.10	Prétraitement	37
2.10.1	Enregistrement d'image.....	37
2.10.2	Suréchantillonnage et interpolation d'image	37
2.10.3	Étirement et correspondance d'histogramme.....	37
2.11	Méthodes de pansharpning.....	38
2.11.1	Méthodes à substitution de composantes.....	38
2.11.2	Méthodes d'analyse multi résolution.....	40
2.11.3	Méthodes variationnelles	41
2.11.4	Méthodes fondées sur l'apprentissage profond	41
2.12	Conclusion.....	42
3	Chapitre 03 : L'apprentissage automatique (Machine Learning) et L'apprentissage profond (Deep Learning).....	44
3.1	Introduction	44
3.2	Développement de l'IA	45
3.3	L'apprentissage automatique (Machine Learning)	46
3.3.1	Approches d'apprentissage automatique :	46
3.4	Apprentissage profond (deep Learning).....	52
3.4.1	Fonctionnement d'un réseau de neurones.....	54
3.5	Deep Learning en traitement d'images.....	59
3.6	Deep Learning et la télédétection	59
3.6.1	Fusion base sur Convolutional Neural Network (CNN).....	61
3.6.2	le Generative adversarial network (GAN).....	63
3.7	Conclusion.....	66
4	Chapitre4 : Implémentation et résultat.....	68
4.1	Introduction :	68
4.2	Ressources matérielles et logicielles	68
4.2.1	Ressources matérielles	68
4.2.2	Ressources logicielles	68
4.2.3	Bibliothèques utilisées dans l'implémentation :.....	71
4.3	Pansharpning avec la colorisation guidée (PANCOLORGAN):	73
4.3.1	La méthode PANCOLORGAN	73
4.4	Expériences et résultats	81
4.4.1	Ensemble de données propose pour le pansharpning.....	82
4.4.2	Code d'application	84

4.4.3	Résultat d'exécution de code :.....	85
4.5	Évaluation de la qualité des méthodes de fusion	87
4.6	Evaluation et Comparaison de résultat	90
	Résultat de comparaison	93
4.7	Conclusion :.....	99
5	Conclusion générale :.....	100

Liste des figures

Figure 1 .1 : Les étapes de télédétection.

Figure 1.2 : Nature et propagation d'une onde électromagnétique.

Figure 1.3: Le spectre électromagnétique.

Figure 1.4: Propriété des ondes électromagnétique.

Figure 1.5: Capteur passif et actif.

Figure1.6 : Plateforme (avion, satellite).

Figure 1.7 : Satellite Ikonos.

Figure 1.8: Satellite QuickBird.

Figure 1.9: Satellite WorldView-2.

Figure 1.10: Satellite WorldView-4.

Figure 1.11: Satellite Gaofen-1 .

Figure. 2.1: Formation d'une image satellitaire numérique multi-spectrale.

Figure 2.2: 1 Image panchromatique.

Figure 2.3: Affichage d'une image Pléiades en composition colorée.

Figure 2.4 : La compression des images hyperspectrales.

Figure 2.5 : Exemple d'étirement de l'histogramme.

Figure 2.6 : schéma représente les méthodes de Pansharpening .

Figure 2.7 : Schéma de la méthode IHS.

Figure 2.8: Schéma du principe de fusion par l'analyse en composantes principales.

Figure 3.1 : La relation entre l'intelligence artificielle, le ML, et le deep learning.

Figure 3.2 : Schéma de fonctionnement d'un neurone biologique et celui d'un neurone artificiel.

Figure 3.3 : L'apprentissage par renforcement.

Figure 3.4 : Apprentissage non supervisé.

Figure 3.5 : La régression linéaire uni variée.

Figure 3.6: La descente de gradient.

Figure 3.7 : La Régression polynomiale.

Figure 3.8 : Classification binaire.

Figure 3.9 : Classification multi-classe.

Figure 3.10 : Apprentissage supervisé.

Figure 3.11 : Architecteur de différents Méthodes de ML.

Figure 3.12: Réseaux de neurone artificiel.

Figure 3.13: Architecteur de Réseaux de neurone.

Figure 3.14 : Types de fonctions d'activation.

Figure 3.15: Apprentissage de Réseaux de neurone.

Figure 3.16 : Pansharpening ou fusion d'images spectrales et spatiales

Figure 3.17: Architecture de RSCNN pour la fusion d'images de télédétection.

Figure 3.18 : Le générateur essaie de générer de fausses images en prenant du bruit aléatoire en entrée et le discriminateur essaie de le classer comme réel ou faux.

Figure 3.19: Architecture de pansharpening à base de (GAN).

Figure 4.1: Schéma de formation proposé pour le modèle PanColorGAN.

Figure 4.2: Schéma de formation proposé pour le modèle PanColorGAN.

Figure 4.3: Schéma d'inférence (test) à pleine résolution.

Figure 4.4: Échantillons de l'ensemble de données proposé

Figure 4.5: image PAN.

Figure 4.6: image MS.

Figure 4.7: Résultat de la fusion PanColorGAN..

Figure 4.8: 2eme image PAN.

Figure 4.9: 2eme image MS.

Figure 4.10: Résultat de la fusion à chaque méthode (REDUCED RESOLUTION).

Figure 4.11: Résultat de la fusion à chaque méthode (FULL-RESOLUTION)

Liste des Tableaux

Tableau 4-1 : Les Détails de base des données.

Tableau 4-2 : LES MÉTHODES DE PANARPENING DANS LES EXPÉRIENCES

Tableau 4-3 : les performances (ERGAS, Qm,SAM,Scs-GT,Qavg,Time) calculées pour de données à Résolution complet (Rr)

Tableau 4-4 : les performances ($D\mu$, Ds , QNR) calculées pour de données à Résolution complet (Fr)

Liste des acronymes et abréviations

API : Application Programming Interface
A-PNN: Pansharpening by deep network
A-PNN-FT: Pansharpening by deep network-flight tuning
ATWT: A trous wavelet transform
ATWT-M3: A trous wavelet transform with injection model 3
AWL: additive wavelet luminance
AWLP: additive wavelet luminance prportional
ACP : l'Analyse en Composantes Principales
ARSIS : l'Amélioration de la Résolution Spatiale par Injection de Structures
ARM: Analyses multiples résolutions
BDS: Band-dependent spatial-detail with local parameter estimation
C-BDS: non-local extension of BDS
CS: Component substitution
CNN : Convolutional Neural Network(réseaux de neurones convolutifs)
CAE: Convolutional Autoencoder
DL: Deep Learning (apprentissage profond)
DBN : Réseau de croyances profondes
DSN : Réseau d'empilement profond
DCN : Deep Convex Network
DWT: Discrete Wavelet Transform
ERGAS : ERREUR RELATIVE GLOBALE ADIMENSIONNELLE DE SYNTHÈSE
Fr: Full resolution
HR MS : Haute Résolution Multispectrale
HPM: haigh pass modulation
IA : Intelligence artificielle
IHS : Intensity Hue Saturation (intensité-teinte-saturation)
GS: Gram-Schmidt
GAN: Generative adversarial network
GIHS : Generalized Intensity Hue Saturation
GP : La pyramide Gaussienne
GLP : Général laplacien pyramide
GRU: Gated Recurrent Uni
GPU : graphique processing unit
LSTM : mémoire longue et courte durée
LP : La pyramide laplacienne
LR MS : Low Resolution Multispectrale
ML : Machine Learning (apprentissage automatique)
MLP : Le perceptron multicouche
MTF-GLP: Generalized Laplacian pyramid algorithm
MTF-GLP-HPM: Generalized Laplacian pyramid algorithm- haigh pass modulation
MRA : Multi résolutions analysais
PANCOLORGAN: Pansharpening with guided colorization using GANS.

PNN: Pansharpener by convolutional neural networks

PP : post processing

RNN : Réseaux de neurones récurrents

RBM : Restricted Boltzmann Machine

PNN : pansharpener neural network

ReLU : Unités Rectifiées linéaires

RaGAN : relativistic average GAN

Rr: Reduced resolution

SFIM : Smoothing Filter-based Intensity Modulation

VO : méthodes Variationnelles

Introduction générale

Introduction générale :

La télédétection est une technologie qui permet l'utilisation de capteurs pour observer et enregistrer le rayonnement électromagnétique émis ou réfléchi par n'importe quelle cible sans la toucher directement. Le traitement et l'analyse des informations véhiculées par le rayonnement enregistré permettent d'accéder à certains attributs de la cible :

- Géométriques (position, forme et dimensions),
- Optiques (réflexion, transmission, absorption, etc.)

Dans ce mémoire nous nous intéresserons au développement de technique de fusion d'images satellitaire est cela en intégrant les informations contenues dans plusieurs images de la même scène en une seule image composite, la fusion d'images au niveau du pixel est reconnue comme ayant une grande importance dans divers domaines, notamment la télédétection, la photographie numérique, l'imagerie médicale, la vidéosurveillance, ect.

Les techniques de fusion d'images par télédétection visent à obtenir une image ayant simultanément une haute résolution spectrale et spatiale. Un exemple typique de fusion d'images de télédétection est connu sous le nom de pansharpening.

Le Pansharpening, qui est l'acronyme de panchromatic sharpening, est une méthodologie de fusion d'une image multi-spectrales(MS) et panchromatiques (PAN) pour obtenir un résultat final ayant la même résolution spatiale que l'image PAN et la même résolution spectrale que l'image MS. Le Panachage s'est avéré utile à la fois pour l'interprétation visuelle et comme étape préliminaire à un traitement de plus haut niveau.

Notre travail proposé par l'équipe de recherche AC2 (Architectures pour la classification et Cryptographie) de la division ASM du centre de Développement des Technologies Avancées (CDTA) consiste à d'implémenter et comparer des méthodes classique telque ; Multi resolution analysis (MRA), Component substitution (CS) et Optimisation variationnelle VO ; avec des méthodes à base de Deep learning (DL) parmi ces méthodes ; la technique Generative adversarial network (Gan).

Afin d'atteindre notre objectif, nous avons suivi le plan de travail suivant :

Le premier chapitre est axé sur une étude bibliographique sur la Télédétection c'est étape à suivre et c'est système System d'acquisition tell que les que le capteur et les satellites.

Le Second chapitre concerne le traitement d'images satellitaires et pour cela nous allons commencer par donner la définition de la fusion d'image et c'est domaine d'application. Pour ensuite pouvoir voir les méthodes utilisées et plus précisément les méthodes Pansharpening tell que Deep Learning

Le troisième chapitre est consacré à l'intelligence artificiel nous allons voir sont historique, donné une comparaison entre le Machine Learning et le Deep Learning pour ensuite pouvoir montrer la déférente application de ce dernier dans le traitement d'image.

Le quatrième chapitre est basé sur la méthode de pansharpening (PanColorGAN) qui est une technique basée sur DL pour la fusion des images Panchromatique et des images multispectrales dans les applications de télédétection, est positionné sur cette technique avec deux nouveaux développements, nommément l'injection de couleur et le hasard.

Enfin, nous donnerons une conclusion et quelques perspectives sur l'amélioration des travaux que nous avons abordé tout au long de ce mémoire.

*Chapitre01 : Notions
fondamentale de la
télédétection*

1 Chapitre01 : Notions fondamentale de la télédétection :

1.1 Introduction

Dans le monde, les citoyens sont informés quotidiennement sur la présence et les effets des phénomènes naturels tels que les séismes, les tsunamis, les éruptions volcaniques, les incendies de forêt ou les inondations. En plus, les humains ont toujours eu un intérêt particulier dans la contemplation de la nature et dans l'observation des caractéristiques du territoire.

La télédétection est définie comme la science et la technologie de collecte d'informations sur les surfaces terrestres, les surfaces marines, les nuages, l'atmosphère terrestre, ...etc., par des avions et des satellites sans être en contact avec la terre. C'est l'obtention, à distance, d'informations relatives à un objet situé sur la surface de la terre ou dans l'atmosphère. Ces informations sont généralement disponibles sous forme d'images.

Dans ce chapitre nous allons définir la télédétection et ces applications.

1.2 Télédétection

La télédétection est une technique d'acquisition à distance d'informations sur la surface terrestre, qui englobe tout le processus qui consiste à capter et enregistrer l'énergie d'un rayonnement électromagnétique émis ou réfléchi, à traiter et analyser l'information qu'il représente, pour ensuite mettre en application cette information [1] .

1.2.1 Différentes étapes de la télédétection

Le processus de la télédétection au moyen de systèmes imageurs comporte les sept étapes suivantes figure 1.1 [2] :

A) Source d'énergie ou d'illumination (A)

À l'origine de tout processus de télédétection se trouve nécessairement une source d'énergie pour illuminer la cible.

B) Rayonnement et atmosphère (B)

Durant son parcours entre la source d'énergie et la cible, le rayonnement interagit avec l'atmosphère. Une seconde interaction se produit lors du trajet entre la cible et le capteur.

C) Interaction avec la cible (C)

Une fois parvenue à la cible, l'énergie interagit avec la surface de celle-ci. La nature de cette interaction dépend des caractéristiques du rayonnement et des propriétés de la surface.

D) Enregistrement de l'énergie par le capteur (D)

Une fois l'énergie diffusée ou émise par la cible, elle doit être captée à distance (par un capteur qui n'est pas en contact avec la cible) pour être enfin enregistrée.

E) Transmission, réception et traitement (E)

L'énergie enregistrée par le capteur est transmise, souvent par des moyens électroniques, à une station de réception où l'information est transformée en images (numériques ou photographiques).

F) Interprétation et analyse (F)

Une interprétation visuelle et numérique de l'image traitée est ensuite nécessaire pour extraire l'information que l'on désire obtenir sur la cible.

G) Application (G)

La dernière étape du processus consiste à utiliser l'information extraite de l'image pour mieux comprendre la cible.

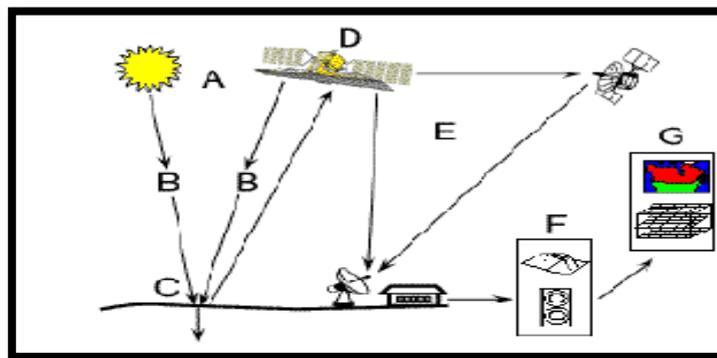


Figure 1.1 : Les étapes de télédétection [2].

1.3 Rayonnement électromagnétique

Une onde électromagnétique contient à la fois un champ électrique et un champ magnétique oscillant à la même fréquence. Ces deux champs, perpendiculaires l'un par rapport à l'autre se propagent dans un milieu selon une direction orthogonale figure 1.2.

La propagation de ces ondes s'effectue à une vitesse qui dépend du milieu considéré. Dans le vide, la vitesse de propagation est égale à 3.10^8 m.s^{-1} [3]

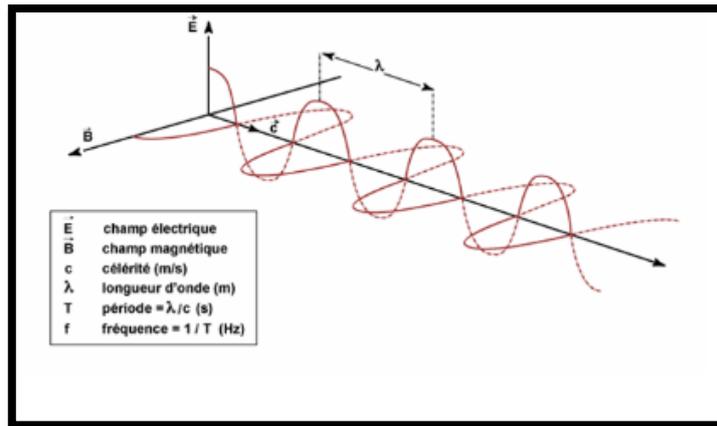


Figure 1.2 : Nature et propagation d'une onde électromagnétique [3]

Une onde électromagnétique est caractérisée par plusieurs grandeurs physiques [3] :

- **La longueur d'onde (λ) :** Elle exprime le caractère oscillatoire périodique de l'onde dans l'espace. C'est la longueur d'un cycle d'une onde, la distance séparant deux crêtes successives. Elle est mesurée en mètre ou en l'un de ses sous-multiples, les ondes électromagnétiques utilisées en télédétection spatiale ayant des longueurs d'onde relativement courtes :

le nanomètre $\Rightarrow 1 \text{ nm} = 10^{-9}$ mètre

le micromètre $\Rightarrow 1 \mu\text{m} = 10^{-6}$ mètre

le centimètre $\Rightarrow 1 \text{ cm} = 10^{-2}$ mètre.

- **La période (T) :** Elle représente le temps nécessaire pour que l'onde effectue un cycle. L'unité est la seconde.
- **La fréquence (ν) :** Inverse de la période, elle traduit le nombre de cycles par unité de temps. Elle s'exprime en Hertz (Hz), un Hz équivaut à une oscillation par seconde, ou en multiples du Hertz, les ondes électromagnétiques utilisées en télédétection spatiale ayant des fréquences très élevées :

le kilohertz $\Rightarrow 1 \text{ kHz} = 10^3 \text{ Hz}$ / le mégahertz $\Rightarrow 1 \text{ MHz} = 10^6 \text{ Hz}$ / le gigahertz $\Rightarrow 1 \text{ GHz} = 10^9 \text{ Hz}$

La longueur d'onde et la fréquence sont inversement proportionnelles et unies par la relation suivante :

$$\lambda = \frac{c}{\nu}$$

Où :

λ : longueur d'onde de l'onde électromagnétique

c : vitesse de la lumière ($3 \cdot 10^8 \text{ m} \cdot \text{s}^{-1}$)

ν : la fréquence de l'onde

Par conséquent, plus la longueur d'onde est petite, plus la fréquence est élevée.

1.3.1 Le spectre électromagnétique

La répartition des ondes électromagnétiques est représentée par le spectre électromagnétique en fonction de leur longueur d'onde, de leur fréquence ou bien encore de leur énergie

Figure 1.3 [4].

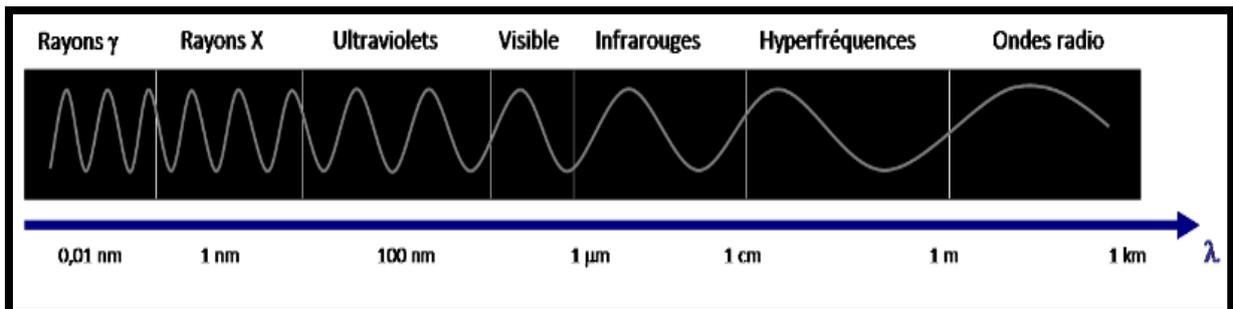


Figure 1.3: Le spectre électromagnétique [1]

En partant des ondes les plus énergétiques, on distingue successivement [4] :

Les rayons gamma (γ) : ils sont dus aux radiations émises par les éléments radioactifs. Très énergétiques, ils traversent facilement la matière et sont très dangereux pour les cellules vivantes.

Leurs longueurs d'onde s'étendent d'un centième de milliardième (10^{-14} m) à un milliardième (10^{-12} m) de millimètre.

Les rayons X : rayonnements très énergétiques traversant plus ou moins facilement les corps matériels et un peu moins nocifs que les rayons gamma, ils sont utilisés notamment en médecine pour les radiographies, dans l'industrie et dans la recherche pour l'étude de la matière. Les

rayons X ont des longueurs d'onde comprises entre un milliardième (10^{-12} m) et un cent millième (10^{-8} m) de millimètre.

Les ultraviolets : rayonnements qui restent assez énergétiques, ils sont nocifs pour la peau. Une grande part des ultraviolets est stoppée par l'ozone atmosphérique qui sert de bouclier protecteur des cellules. Leurs longueurs d'onde s'échelonnent d'un cent millième (10^{-8} m) à quatre dixièmes de millième ($4 \cdot 10^{-7}$ m) de millimètre.

Le domaine visible : correspond à la partie très étroite du spectre électromagnétique perceptible par notre œil. C'est dans le domaine visible que le rayonnement solaire atteint son maximum ($0,5 \mu\text{m}$) et c'est également dans cette portion du spectre que l'on peut distinguer l'ensemble des couleurs de l'arc en ciel, du bleu au rouge. Il s'étend de quatre dixièmes de millième ($4 \cdot 10^{-7}$ m), lumière bleue, à huit dixièmes de millième ($8 \cdot 10^{-7}$ m) de millimètre, lumière rouge.

L'infrarouge : rayonnement émis par tous les corps dont la température est supérieure au zéro absolu (-273°C). En télédétection, on utilise certaines bandes spectrales de l'infrarouge pour mesurer la température des surfaces terrestres et océaniques, ainsi que celle des nuages. La gamme des infrarouges couvre les longueurs d'onde allant de huit dixièmes de millième de millimètre ($8 \cdot 10^{-7}$ m) à un millimètre (10^{-3} m).

Les ondes radar ou hyperfréquences : Cette région du spectre est utilisée pour mesurer le rayonnement émis par la surface terrestre et s'apparente dans ce cas à la télédétection dans l'infrarouge thermique, mais également par les capteurs actifs comme les systèmes radar. Le domaine des hyperfréquences s'étend des longueurs d'onde de l'ordre du centimètre jusqu'au mètre.

Les ondes radio : Ce domaine de longueurs d'onde est le plus vaste du spectre électromagnétique et concerne les ondes qui ont les plus basses fréquences. Il s'étend des longueurs d'onde de quelques cm à plusieurs km. Relativement faciles à émettre et à recevoir, les ondes radio sont utilisées pour la transmission de l'information (radio, télévision et téléphone). La bande FM des postes de radio correspond à des longueurs d'onde de l'ordre du mètre. Celles utilisées pour les téléphones cellulaires sont de l'ordre de 10 cm environ.

Contrairement à l'œil humain qui n'est capable de capter le rayonnement que dans une fenêtre très étroite du spectre électromagnétique, celle correspondant au domaine du visible (longueurs d'onde comprises entre $0,4 \mu\text{m}$ et $0,7 \mu\text{m}$), les capteurs satellitaires utilisent une fraction beaucoup plus étendue du spectre.

Trois fenêtres spectrales sont principalement utilisées en télédétection spatiale [4]

- Le domaine du visible
- Le domaine des infrarouges (proche IR, IR moyen et IR thermique)
- Le domaine des microondes ou hyperfréquences

1.3.2 Propriété des ondes électromagnétiques

Le rayonnement doit traverser une certaine épaisseur d'atmosphère avant qu'il soit utilisé pour la télédétection qui n'atteigne la surface de la terre. Les particules et les gaz dans l'atmosphère peuvent dévier ou broyer le rayonnement incident. Ces effets sont causés par les mécanismes de diffusion et d'absorption figure 1.4 [5].

A- La diffusion : Elle se produit lors de l'interaction entre le rayonnement incident et les particules présentes dans l'atmosphère. La diffusion a pour effet une augmentation du rayonnement dans toutes les directions et une atténuation dans la direction de propagation de l'onde.

B- L'émission : Le corps émetteur est appelé source, et les ondes émises se déplacent en ligne droite tant que l'indice de réfraction du milieu reste constant.

C- L'absorption : Chaque corps dont la surface reçoit un rayonnement électromagnétique peut absorber une partie qui sert à modifier l'énergie interne de ce corps. L'absorptance α est définie comme le rapport entre l'énergie absorbée et l'énergie reçue (incidente) [8].

$$\alpha = \frac{\text{énergie absorbée}}{\text{énergie reçue}}, \alpha \in]0; 1[\dots\dots\dots(1.1)$$

D- La réflexion : Tous corps qui reçoivent une certaine quantité d'énergie rayonnante d'une source extérieure peuvent en réfléchir une partie. Le rapport entre l'énergie réfléchie et l'énergie incidente, est appelé la réflectance ρ .

$$\rho = \frac{\text{énergie réfléchi}}{\text{énergie reçue}}, \rho \in]0; 1[\dots\dots\dots(1.2)$$

E- Transmission : Tout corps qui reçoit une certaine quantité d'énergie rayonnante peut en transmettre une partie. La transmission τ est le rapport entre l'énergie transmise et l'énergie incidente.

$$\tau = \frac{\text{énergie transmise}}{\text{énergie reçue}}, \tau \in]0; 1[\dots\dots\dots (1.3)$$

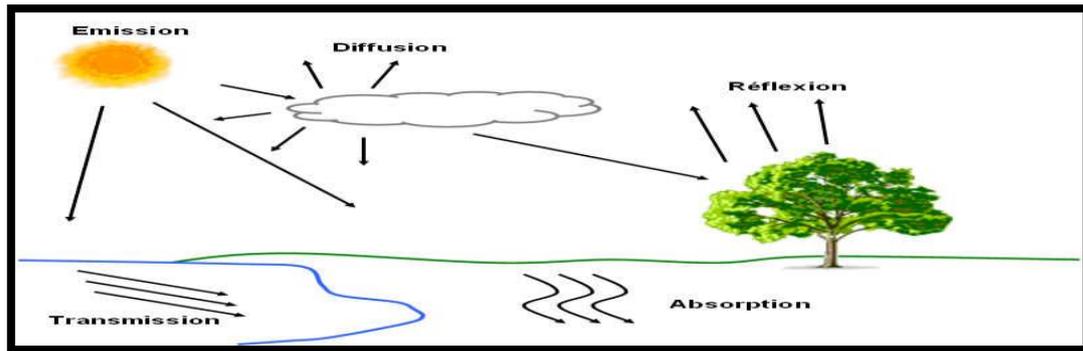


Figure 1.4: Propriété des ondes électromagnétique [5]

1.4 Système d'acquisition

Le système d'acquisition est composé de deux éléments principaux, le premier est le capteur, et le deuxième c'est la plateforme [9].

1.4.1 Les capteurs

Un capteur est un dispositif électronique qui permet de recueillir ou de transformer de l'énergie. En télédétection, il existe deux types de capteurs figure 1.5 [6] :

Capteurs actifs : Ils fournissent leurs propres sources de rayonnement électromagnétique.

Le laser fluoromètre et le radar à synthèse d'ouverture (RSO) sont des exemples de capteurs actifs.

Capteurs passifs : Les dispositifs de télédétection qui mesurent l'énergie disponible naturellement sont des capteurs passifs. Ils peuvent seulement percevoir l'énergie réfléchie lorsque le soleil illumine la terre. Ils n'illuminent pas la cible, ils peuvent seulement recevoir l'énergie émise ou dispersée par la cible.

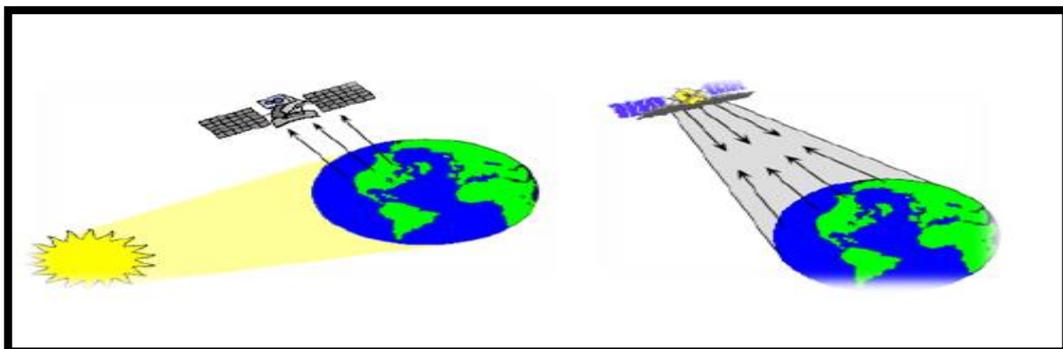


Figure 1.5: Capteur passif et actif [6]

1.4.2 Les plateformes

Pour enregistrer adéquatement l'énergie réfléchié ou émise par une surface ou une cible donnée, on doit installer un capteur sur une plateforme distante de la surface ou de la cible observée. Ces plateformes peuvent être situées près de la surface terrestre, comme par exemple au sol, dans un avion ou un ballon ; ou à l'extérieur de l'atmosphère terrestre, comme par exemple un satellite. Parmi les plateformes les plus utilisées, on a les avions et les satellites figure 1.6 [2] :

Les avions : Les avions volent à une altitude relativement faible (quelque kilomètre seulement) ils ne peuvent par conséquent prendre en photo (du point de vue de la télédétection) des portions de territoire limitées avec leurs détails. [7]

Les satellites : Ils se déplacent en suivant une orbite, généralement, l'orbite a une forme quasi-circulaire. Les satellites sont des plates-formes utilisées communément en télédétection ; ils véhiculent des capteurs extrêmement variés et souvent spécialisés dans l'observation de la météo, des paysages ou des catastrophes naturelles, etc. [7].

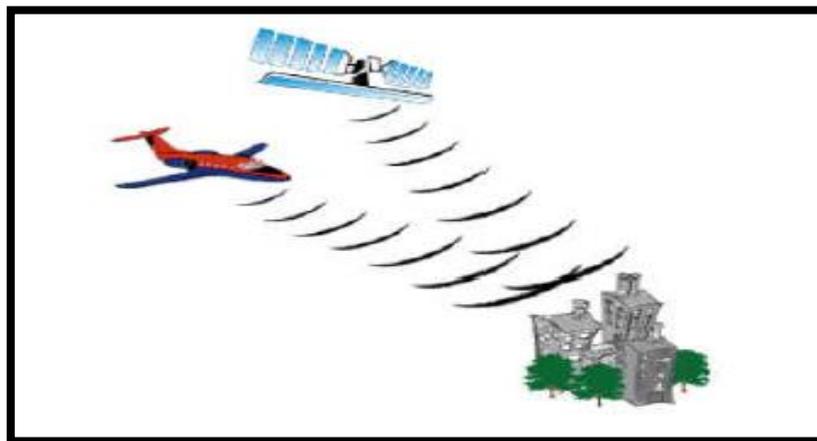


Figure 1.6: Plateforme (avion, satellite)[7]

Suivant les différentes missions des satellites nous pouvons citer [8] :

- Les satellites d'astronomie.
- Les satellites de navigation.
- Les satellites météorologiques.
- Les satellites militaires.
- Les satellites de communication.

Selon la trajectoire du satellite on peut distinguer deux types de satellites [7]:

a. Satellites à défilement : Ils permettent une observation cyclique de la terre. Ils ont une altitude comprise entre 800Km et 1500Km et décrivent une orbite polaire.

b. Satellites géostationnaires : Ils ont une position fixe par rapport à la terre, ainsi ils peuvent assurer une surveillance continue d'une partie du globe.

1.4.3 Caractéristiques des satellites

Parmi les satellites les plus connus, on cite :

a) Satellite Ikonos

Le satellite IKONOS a été lancé le 24 septembre 1999 depuis la base militaire de Vandenberg en Californie. Ses capteurs panchromatique et multispectral à très haute résolution peuvent différencier les objets avec une précision supérieure à un mètre carré au sol, soit suffisamment pour distinguer une voiture d'un camion.

Une telle performance réalisée à partir d'une altitude orbitale de 680 km représente un accroissement significatif de la résolution d'image par rapport à tous les systèmes commerciaux précédents de télédétection par satellite [9].

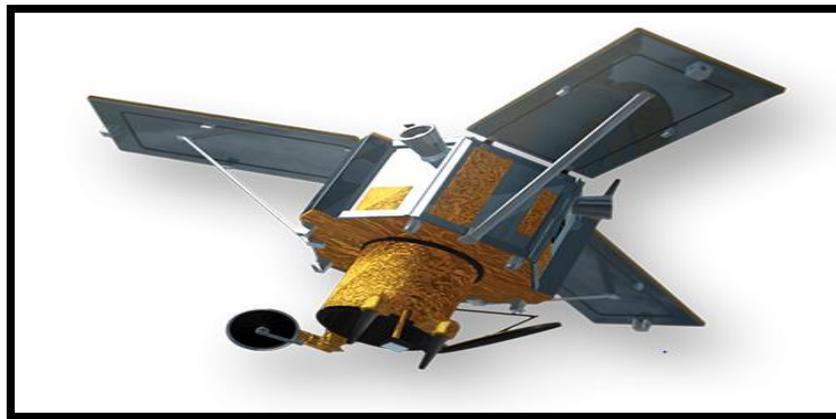


Figure 1.7 : Satellite Ikonos [9].

b) Satellite QuickBird

QuickBird est un satellite commercial de la société DigitalGlobe, il bénéficiait de la plus haute résolution parmi tous les satellites commerciaux en fonctionnement. Certains satellites, comme les Worldview exploités par DigitalGlobe, disposent actuellement de capteurs optiques. QuickBird a été mis au point pour couvrir de vastes étendues avec efficacité et précision. Tandis que QuickBird fait le tour du globe à une altitude de 600 km au-dessus de la Terre, son capteur (une caméra haute résolution) recueille des images de la surface de la Terre durant les heures du jour. Le système recueille des données panchromatiques de 61cm et des données stéréoscopiques multispectrales de 2,5 m [10].



Figure 1.8: Satellite QuickBird [10]

c) Satellite WorldView-2

Ce satellite est capable de collecter jusqu'à 975 000 km² d'images par jour et a effectivement doublé la capacité de collecte de DigitalGlobe .

Le système WorldView-2 offre une précision, une agilité, une capacité et une diversité spectrale inégalées. WorldView-2 se distingue par le fait qu'il capture des images haute résolution avec huit bandes spectrales, par rapport à la gamme habituelle des satellites à quatre bandes. Avec les quatre bandes spectrales supplémentaires, WorldView-2 offre des opportunités uniques pour l'analyse par télédétection de la végétation, des environnements côtiers, de l'agriculture, de la géologie et bien d'autres [11].



Figure 1.9: Satellite WorldView-2 [11].

d) Satellite WorldView-3

WorldView-3 est un satellite d'imagerie et de surveillance de l'environnement de DigitalGlobe des États-Unis qui a été lancé le 13 août 2014 et reste opérationnel. Il est très similaire à WorldView-2 mais positionné sur une orbite plus basse. Le capteur satellite WorldView-3 a été

autorisé par la National Oceanic and Atmospheric Administration (NOAA) à collecter, en plus des bandes panchromatiques et multispectrales standard, des images infrarouges à ondes courtes (SWIR) .WorldView-3 est le premier capteur satellite commercial multi-charges, super-spectral et haute résolution fonctionnant à une altitude de 617 km [12].

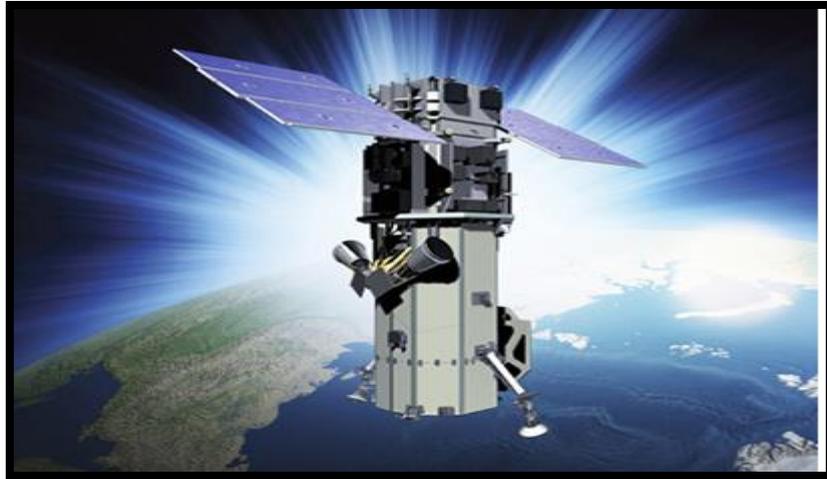


Figure 1.10: Satellite WorldView-3

e) Satellite WorldView-4

WorldView-4 était un satellite d'imagerie et de surveillance de l'environnement de DigitalGlobe avec des options d'imagerie à très haute résolution. Fonctionnant en coordination avec WorldView-3 et les autres satellites de la constellation WorldView, il offrait des images haute résolution à 31 cm de résolution en mode panchromatique et 1,23 m en mode multispectral [13].

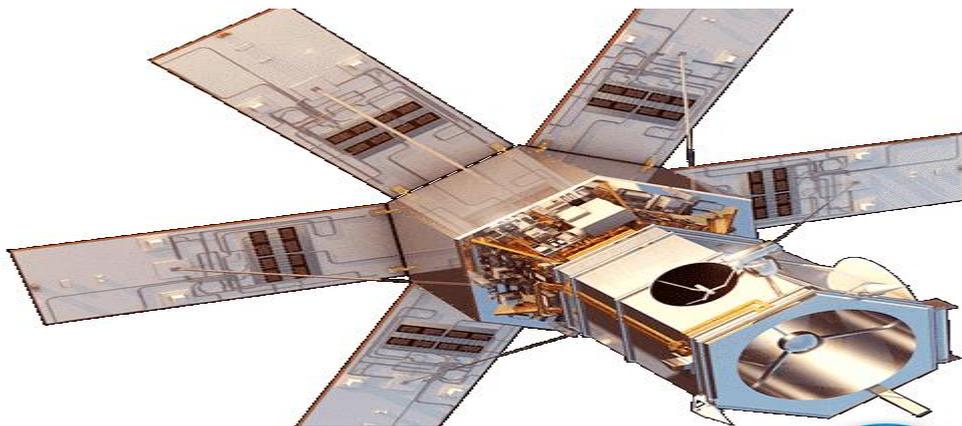


Figure 1.10: Satellite WorldView-4 [13]

f) Satellite Gaofen-1

Gaofen-1 (gao fen = haute résolution) est le premier d'une série de satellites optiques d'observation de la Terre à haute résolution de la CNSA (China National Space Administration), Pékin, Chine. L'objectif principal de la série HDEOS(High-Definition Earth Observation

Satellite) est de fournir des observations NRT (Near-Real-Time) pour la prévention des catastrophes et les secours, la surveillance du changement climatique, la cartographie géographique, l'étude de l'environnement et des ressources, ainsi que pour le soutien à l'agriculture de précision. Les principaux utilisateurs des données d'observation seront le ministère des Terres et des Ressources, le ministère de la Protection de l'environnement et le ministère de l'Agriculture [14].



Figure 1.11: Satellite Gaofen-1 [14].

1.5 Application de la télédétection

La télédétection couvre un grand champ d'applications qui satisfont de nombreux besoins, aussi bien pour les utilisateurs institutionnels (ministères, agences gouvernementales, organisations internationales) que pour des acteurs privés (sociétés, agriculteurs, organisations non gouvernementales, etc.) [15].

1.5.1 La foresterie

La foresterie est la science de gestion, de conservation, d'aménagement, et d'utilisation des forêts, parmi les applications de la télédétection en foresterie se trouvent également [2] :

1) La cartographie de reconnaissance

Les agences nationales forestières ou environnementales poursuivent des objectifs ; tels que la mise à jour du couvert forestier, la surveillance de la diminution des ressources et la mesure des propriétés biophysiques des forêts :

- Discrimination du type de couvert forestier.
- Cartographie agroforestière.

2) La foresterie commerciale

Les compagnies forestières ainsi que les agences de gestion des ressources accordent beaucoup d'importance à l'inventaire et aux applications cartographiques telles que la collecte

d'information sur les récoltes, la mise à jour des inventaires pour la demande de coupe, les types de forêts, la densité de végétation et les mesures de biomasse :

- Cartographie et évaluation de la régénération.
- Déboisement par le feu.
- Cartographie d'infrastructure.
- Inventaire des forêts.
- Estimation de la biomasse.

3) La Surveillance environnementale

Les gestionnaires de conservation sont préoccupées par la surveillance de la qualité, de la santé et de la diversité des forêts de la Terre :

- Déforestation.
- Protection des bassins déversant.
- Protection des côtes.
- Santé et vigueur des forêts.

1.5.2 La géologie

La géologie est une science qui consiste à étudier les parties de la Terre accessibles à l'observation et à élaborer des hypothèses qui permettent de reconstituer leur histoire et expliquer leur agencement. La liste des principales applications de la télédétection pour la géologie [2] :

- La cartographie des dépôts de surface.
- La cartographie lithologique.
- La cartographie structurale.
- L'exploration et exploitation des agrégats (sable et gravier).
- L'exploration minière.
- L'exploration pétrolière.
- La géologie environnementale.
- La géobotanique.
- La cartographie et surveillance des taux de sédimentation.
- La cartographie et surveillance des phénomènes naturels.
- La cartographie des risques géologiques.
- La cartographie planétaire.

1.6 Conclusion :

Dans ce chapitre nous avons présenté des notions fondamentales de la télédétection, nous nous intéressons dans le chapitre suivant aux méthodes de fusion d'images satellitaires (panchromatiques, multispectrales et hyperspectrales). Nous commençons par donner un panorama des méthodes de pansharpening proposées dans la littérature, qui consistent à fusionner une image hyperspectrale ou multispectrale avec une image panchromatique dans le but de produire une image avec des résolutions spatiale et spectrale élevées.

*Chapitre02 : Notions
fondamentale de la
télédétection*

2 Chapitre 02 : Traitement d'images satellitaires

2.1 Introduction

L'imagerie satellite désigne la prise d'images de la Terre ou d'autres planètes à partir de satellites artificiels. Compte tenu des méthodes utilisées (télédétection), il s'agit bien d'imagerie.

Les satellites d'observation de la Terre ont permis à la télédétection spatiale de devenir un outil incontournable pour la compréhension et la connaissance de notre environnement.

En effet, les satellites d'observation de la Terre fournissent différentes données issues de divers capteurs passifs ou actifs qui peuvent présenter des caractéristiques très différentes, en particulier en termes de résolutions spatiale, spectrale, temporelle et radiométrique.

La procédure la plus connue qui permet d'améliorer la résolution spatiale des images est la fusion (appelée aussi sharpening), dans laquelle une image de résolution spectrale élevée est généralement fusionnée à une image de résolution spatiale élevée pour obtenir une image de haute résolution à la fois spectrale et spatiale. La plupart des techniques de fusion pour l'amélioration de la résolution spatiale ont été développées dans le but d'améliorer la résolution spatiale d'une image multispectrale ou hyperspectrale en utilisant une image panchromatique de résolution spatiale plus élevée. Cette technique est appelée "Pansharpening".

2.2 Image satellitaire

Une image satellitaire est une représentation graphique, constituée de nombreux carrés appelés pixels. Le pixel représente la plus petite unité figurant sur une image satellitaire, quelle que soit la longueur d'onde ou le dispositif de télédétection qui ont été utilisés pour capter et enregistrer l'énergie électromagnétique[5]

2.3 La formation d'image satellitaire

Les images satellitaires sont un arrangement de données numériques, les pixels, disposées de telle façon que, visuellement, elles nous donnent l'information de ce qui se trouve sur une certaine scène. Nous pouvons observer sur la figure 2.2, les éléments qui interviennent pendant la formation des images numériques multi-spectrale. Ces images sont formées d'une façon

générale par un système de capteur qui donne une représentation de la scène de la façon suivante [16] :

La source d'énergie : Elle envoie la radiation à la scène observée dont la nature peut être de type ondulatoire (électromagnétique ou acoustique) ou corpusculaire (neutrons, positrons, électrons).

La scène : Elle est composée d'éléments basiques nommés champ instantanés et champs instantanés visuels. Chacun de ces champs interagit avec la source de radiation ce qui résulte en une émission d'énergie vers le capteur.

Le capteur : Le capteur suit une certaine trajectoire pour enregistrer le signal émis par la source d'énergie qui est réfléchié par la scène.

Les filtres : C'est ici que la séparation des différentes bandes d'énergie est faite.

Les Convertisseurs : La nature des signaux est continue, c'est pour cela qu'une discrétisation est faite dans cette étape.

Une image Numérique Multi-spectrale : Il s'agit d'une simplification de la scène. Par la suite, nous notons l'image comme un ensemble S , une grille, composée de sites $s \in S$, la valeur qui est prise dans chaque site est $x_{s \in S}^{(n)} \in G$, ou G est l'ensemble des niveaux de gris, dans la plupart des cas $[0,255]$. Dans la suite de ce manuscrit, nous allons montrer comment ces valeurs sont acquises.

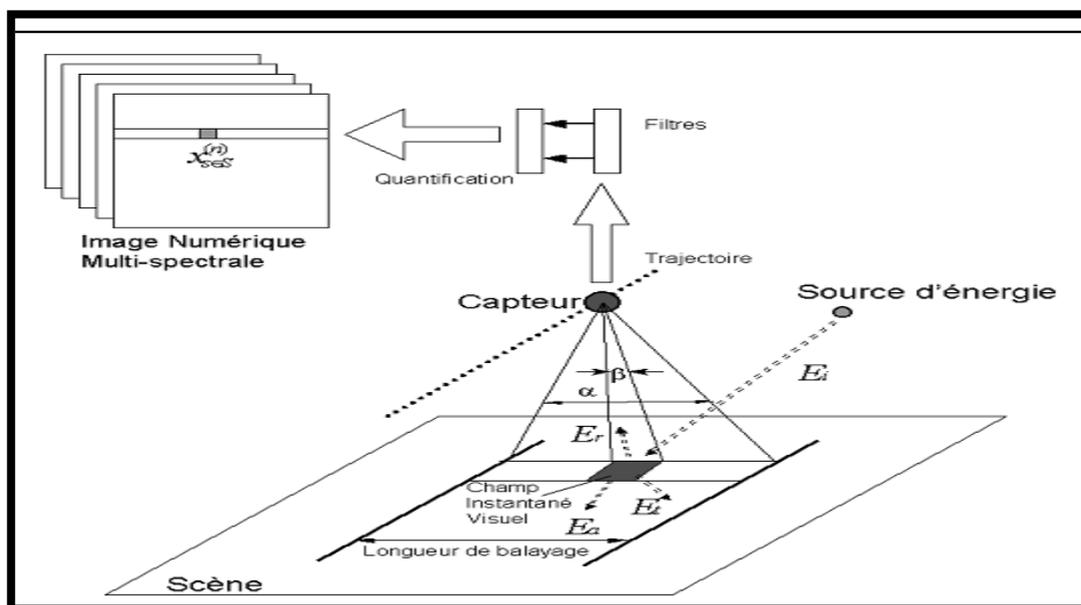


Figure. 2.1: Formation d'une image satellitaire numérique multi-spectrale [16].

2.4 Les résolutions des images satellitaires

La résolution est un terme souvent utilisé pour déterminer, soit le nombre de pixels que l'on peut afficher ou alors la surface du sol qu'un pixel représente dans un fichier image. Quand il s'agit de décrire des données d'images satellitaires. [17].

A). Valeurs des pixels : Chaque pixel d'une image a une valeur. Cette valeur correspond à l'intensité du rayonnement réfléchi par l'objet observé dans la gamme de longueur d'ondes auxquelles le capteur est sensible, le nombre maximum de niveaux d'intensité disponibles dépend du nombre de bits utilisés pour représenter l'intensité enregistrée.[5]

B) Notions de résolution : On distingue, quatre sortes de résolutions :

La résolution spectrale, qui se rapporte aux intervalles spécifiques des longueurs d'ondes dans le spectre électromagnétique qu'un senseur puisse enregistrer. Plus l'intervalle est large dans le spectre électromagnétique, plus la résolution spectrale est petite et vice-versa. [17].

La résolution spatiale, qui représente la mesure des objets les plus petits discernés

Par le senseur ou alors la surface du sol représentée par chaque pixel. Chaque nombre ou valeur dans le fichier image est appelé pixel qui est la plus petite partie d'une image avec une seule valeur. Cette dernière représente la mesure de la brillance du pixel à une certaine longueur d'onde [17].

La résolution radiométrique, qui se réfère au rayon d'action dynamique ou le nombre possible de valeurs dans chaque bande, et ce, par le nombre de bits par lequel l'énergie enregistrée est répartie dans le fichier image. Par exemple des données 8 bits sont celles ayant 28 valeurs, soit 256 niveaux de brillances allant de 0 à 255[17].

La résolution temporelle, qui se rapporte à la fréquence d'un senseur à obtenir une image d'une région particulière. Par exemple, le satellite Landsat peut prendre des images de la même région tous les 16 jours, alors que SPOT peut le faire tous les jours, grâce à la technique de visée latérale [17].

2.5 Définition de la fusion d'image

La fusion d'images est la combinaison de deux ou de plusieurs images différentes pour former à l'aide d'un algorithme une nouvelle image. Le processus de fusion d'images est défini comme la collecte de toutes les informations importantes à partir de plusieurs images et leur inclusion dans moins d'images, généralement une seule. Cette image unique est plus informative et précise que n'importe quelle image source unique, et elle comprend toutes les informations nécessaires. Le but de la fusion d'images n'est pas seulement de réduire la quantité de données mais aussi de construire des images plus appropriées et compréhensibles pour la perception humaine et machine. En imagerie satellitaire, deux types d'images sont disponibles. L'image panchromatique acquise par les satellites est transmise avec la résolution maximale disponible et les données multispectrales sont transmises avec une résolution plus grossière. Ce sera généralement deux ou quatre fois inférieure. A la station réceptrice, l'image

panchromatique est fusionnée avec les données multispectrales pour transmettre plus d'informations [16].

2.6 Application de fusion image

La fusion des images vise à intégrer différentes informations dans le but d'extraire une information supplémentaire.

Ces informations peuvent provenir de plusieurs capteurs, ou peuvent être multi-temporels ou bien de résolutions spatiales différentes. Les objectifs de la fusion d'images dépendent principalement des sujets spécifiques de recherches et de la disponibilité des images sources. On retrouve parmi les domaines d'applications [16] :

2.6.1 L'aide à la navigation :

Permettre aux pilotes par exemple de voler dans de mauvaises conditions de visibilité (Telles que le brouillard ou la forte pluie). Les hélicoptères sont équipés de plusieurs capteurs d'imagerie, qui peuvent être consultés par le pilote. Une suite typique de capteurs inclut un capteur de lumière basse et un capteur infrarouge pour les images thermiques.

Une amélioration possible est de combiner les deux sources d'images dans une seule image fusionnée qui contient l'information appropriée des deux dispositifs imageurs [16].

2.6.2 La télédétection :

La télédétection est une application typique de la fusion d'image : Les modules à balayage spectraux modernes recueillent jusqu'à plusieurs centaines de bandes spectrales qui peuvent être visualisées et traitées individuellement, ou qui peuvent être fusionnées en une seule image, en fonction de la tâche d'analyse d'image., la fusion peut se pratiquer lors d'applications les plus diverses :

- Établir et mettre à jour une carte topographique.
- Étudier les occupations du sol.
- Aider à la décision dans l'agriculture et la foresterie.
- Surveiller les inondations, les neiges et les glaces.
- Améliorer les capacités d'interprétation et de lisibilité.

2.7 Classification spectrale

Les capteurs optiques capturent des images optiques de la surface de la Terre dans les bandes visibles, proche infrarouge et infrarouge à ondes courtes. Les longueurs d'onde des bandes dans un système de télédétection optique vont de **0,30 mm à 15.0 mm**. Les informations radiométriques, spectrales, texturales, géométriques et contextuelles, contenues dans une image optique, servent généralement à l'interprétation de l'image. En fonction du nombre de bandes spectrales utilisées, les images optiques sont classées en quatre grandes catégories ; panchromatique(Pan), multi-spectrale(MS), super-spectrale(SS) et hyper-spectrale(HS)[19].

A) Image panchromatique

Les images panchromatiques sont obtenues à partir de l'enregistrement du rayonnement dans un unique intervalle de longueur d'onde situé dans le domaine du visible, entre **0,4 m et 0,7 m**.

Comme les données ne sont acquises que dans un seul canal, seules des images en noir et blanc peuvent être obtenues (si l'image est codée sur 8 bits, elle sera visualisable en 255 niveaux de gris). Bien que moins riche du point de vue de la résolution spectrale, l'image panchromatique offre une résolution spatiale plus importante [20].

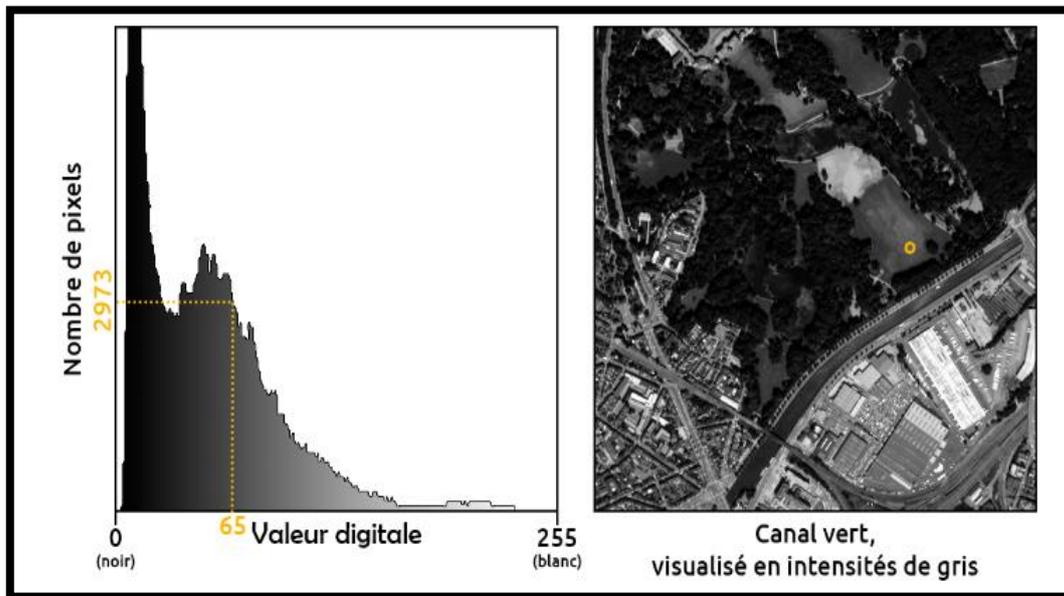


Figure 2.2: 1 Image panchromatique [20].

b).Image multi spectrale :

Les données multispectrales sont obtenues par des enregistrements simultanés dans un petit nombre de bandes spectrales (3 à 8 bandes), celles-ci n'étant pas nécessairement contiguës.

La représentation par combinaison de ces bandes d'information numérique en utilisant les trois couleurs primaires (rouge, vert, bleu) permet d'obtenir des images en couleurs. Les données de chaque bande sont représentées comme une couleur primaire et, selon la luminosité relative (c.-à-d. valeur numérique) de chaque pixel dans chaque bande, les couleurs primaires se combineront en proportions différentes pour produire des couleurs distinctes [20].

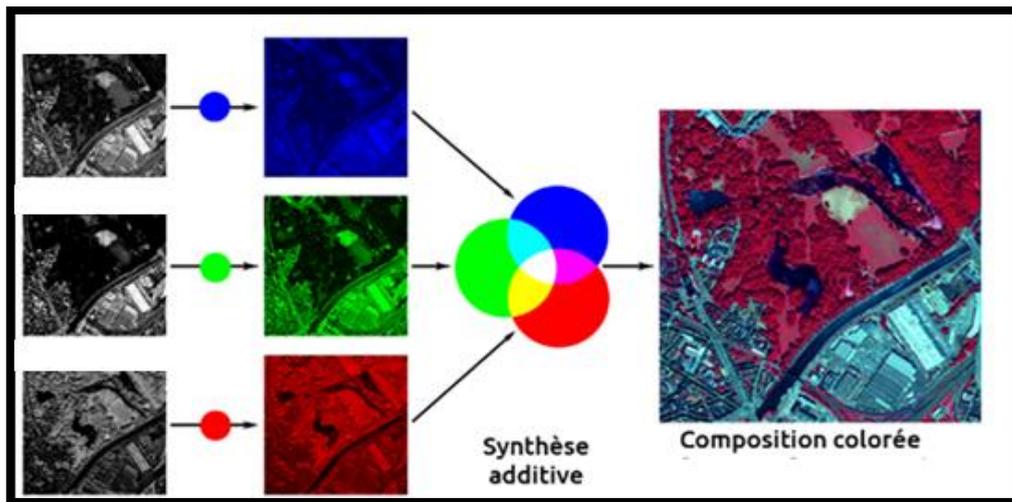


Figure 2.3: Affichage d'une image Pléiades en composition colorée [20]

2.8 Classification Spatiale

La résolution spatiale peut également être décrite comme la surface au sol qui forme un pixel dans l'image satellite. En termes de résolution spatiale, les systèmes d'imagerie par satellite peuvent être classés en :

- Systèmes à basse résolution,
- Systèmes à moyenne résolution,
- Systèmes à haute résolution,
- Et les systèmes à très haute résolution.

La résolution spatiale du capteur panchromatique Ikonos, par exemple, est de 1 m [19].

A). Basse résolution

Les satellites à basse résolution se caractérisent par une résolution spatiale d'environ 1 km offrant une couverture quotidienne régulière. Ces données ne sont obtenues qu'en mode multispectral incluant une partie visible et infrarouge du spectre optique. Ces données sont utiles pour la cartographie de grande surface, mondiale et continentale, les besoins de couverture régulière, y compris la surveillance de la température de surface de la mer, surveillance des conditions, études régionales de la vigueur de la végétation et de la sécheresse, surveillance des grandes catastrophes, surveillance de la couverture neigeuse et des glaciers [19].

B). Résolution moyenne

Les satellites à moyenne résolution offrent généralement une couverture hebdomadaire à mensuelle. Ces données ont des résolutions spatiales de 100 m à 1 km et des bandes spectrales allant du bleu clair à l'infrarouge à ondes courtes. Des zones relativement vastes sont couvertes par chaque satellite. Ces données conviennent à une variété d'utilisations de l'information

impliquant la cartographie, la surveillance et la détection de la couverture terrestre et des caractéristiques d'utilisation des terres.

Les données de télédétection à résolution spatiale moyenne peuvent fournir des informations utiles sur la production primaire brute, en particulier à l'échelle des zones urbaines [19].

C) Haute et très haute résolution

Les satellites à haute résolution spatiale sont caractérisés par une résolution de pixels supérieure à 1 m ; contrairement à une résolution spatiale très élevée où la résolution est inférieure à 1 m. Ces données sont obtenues uniquement en mode panchromatique, le plus souvent en combinaison de modes panchromatique et multi spectral. Les images peuvent être collectées sur une zone particulière tous les 1 à 3 jours, mais à des angles de vue très variables. La majorité des satellites embarquent les systèmes les plus modernes avec une grande flexibilité et capacité à obtenir des données selon les demandes très concrètes.

Les images à haute résolution spatiale trouvent de plus en plus d'applications dans tous les domaines.

Les applications typiques supplémentaires comprennent: les modèles de villes 3D, le contrôle des activités agricoles, l'inventaire des forêts, la surveillance d'un lac glaciaire, la surveillance des mines à ciel ouvert, la planification et la conception d'infrastructures linéaires, cartographie des infrastructures de transport, du secteur des assurances, cartographie des bâtiments détruits par un tremblement de terre [19].

2.9 Les Méthodes de fusion d'image

Les méthodes de fusion d'images existantes peuvent être classifiées dans plusieurs groupes [16].

2.9.1 Méthode de la transformée en ondelettes

Cette méthode utilise l'information spatiale de la bande Pan pour simuler des images MS. La simulation des images MS se fait grâce à l'utilisation de la transformée en ondelettes et de l'analyse multi résolution. Cette analyse a été introduite par Mallat (1989). Cet outil mathématique permet de calculer des approximations successives d'une même image à des résolutions spatiales de plus en plus grossières. Elle peut être représentée par une pyramide dont la base est l'image originale et dont les différentes approximations sont les différents étages de la pyramide [16].

2.9.2 Méthode ARSIS

ARSIS est l'acronyme de « Amélioration de la Résolution Spatiale par Injection de Structures ». Cette méthode permet, dans un ensemble d'images de résolutions spatiale et spectrale différentes, d'amener les images ayant une moins bonne résolution spatiale à la meilleure résolution spatiale disponible dans l'ensemble des données. Cependant, la qualité spectrale des images dont on améliore la résolution spatiale est préservée. Les images obtenues peuvent être utilisées à d'autres fins que la simple visualisation et ainsi améliorer, par exemple, la précision

des classifications réalisées sur les images multibandes originales. Cette méthode s'appuie sur la transformée en ondelettes et l'analyse multirésolution [16].

2.10 Prétraitement

De nombreuses opérations de prétraitement sont utilisées dans l'imagerie de télédétection. Certains d'entre eux tentent de diminuer les distorsions issues des capteurs et des plates-formes. Ces distorsions radiométriques et géométriques sont dues à :

Des variations dans l'éclairage de la scène, la géométrie de la vue, les conditions atmosphériques, la réponse du capteur et le bruit du capteur. Bien qu'appliquant des méthodes pour corriger les distorsions, des algorithmes d'amélioration de l'image sont nécessaires pour l'interprétation visuelle et la compréhension de l'imagerie. Dans ce chapitre, nous nous intéressons principalement aux algorithmes comprenant principalement les techniques d'enregistrement, d'interpolation et de correspondance d'histogrammes [19].

2.10.1 Enregistrement d'image

L'enregistrement d'image est l'une des opérations de traitement d'image les plus utilisées en télédétection. Il s'agit d'un processus par lequel la correspondance la plus précise est déterminée entre deux images qui ont été prises au même point de vue ou à des points de vue différents et par des capteurs différents. En règle générale, ce processus est requis en télédétection, en médecine, en cartographie, en vision par ordinateur [19].

2.10.2 Suréchantillonnage et interpolation d'image

Les images MS acquises ont une faible résolution spatiale par rapport aux images Pan. Afin d'appliquer des algorithmes de prétraitement panoramique, les images MS doivent être suréchantillonnées à la même résolution que les images Pan.

Le processus de suréchantillonnage peut impliquer une interpolation, généralement effectuée par convolution de l'image avec un noyau d'interpolation. Les méthodes d'interpolation linéaire couramment utilisées, telles que le plus proche voisin, l'interpolation bilinéaire et la convolution cubique, présentent des avantages en termes de simplicité et de mise en œuvre

Rapide. Le rééchantillonnage du voisin le plus proche utilise la valeur de pixel dans l'image d'origine qui est la plus proche du nouvel emplacement de pixel dans l'image rééchantillonnée [19].

2.10.3 Étirement et correspondance d'histogramme

L'histogramme d'une image représente la distribution statistique des valeurs de luminance comprises dans une image. C'est un outil utile pour l'amélioration du contraste.

Un exemple figure 2.8, une image Pan et son histogramme correspondant sont affichés avant et après l'étirement de l'histogramme dans les parties supérieure et inférieure de la figure, respectivement [19].

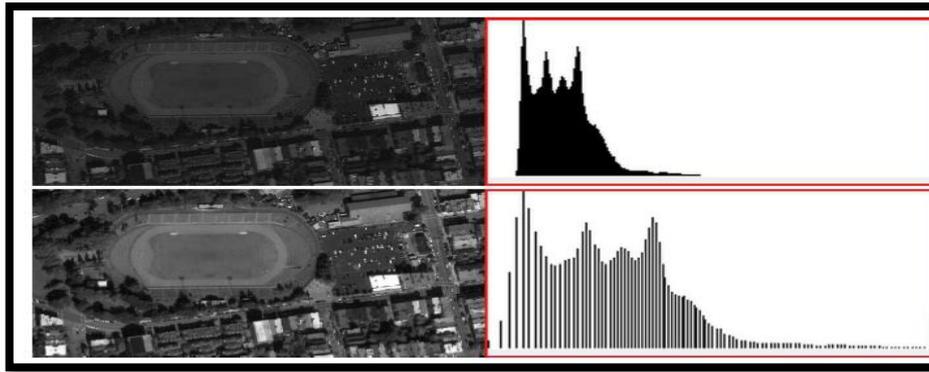


Figure 2.5 Exemple d'étirement de l'histogramme [19].

2.11 Méthodes de pansharpening

La fusion d'image est un domaine de recherche actif en traitement d'image, de nombreuses solutions dans pansharpening exploitent des méthodes a été proposé pour améliorer la résolution spatiale des images hyperspectrales ou multispectrales, en utilisant les informations spatiales provenant d'une image panchromatique. Nous présentons quatre solutions de fusion d'images :

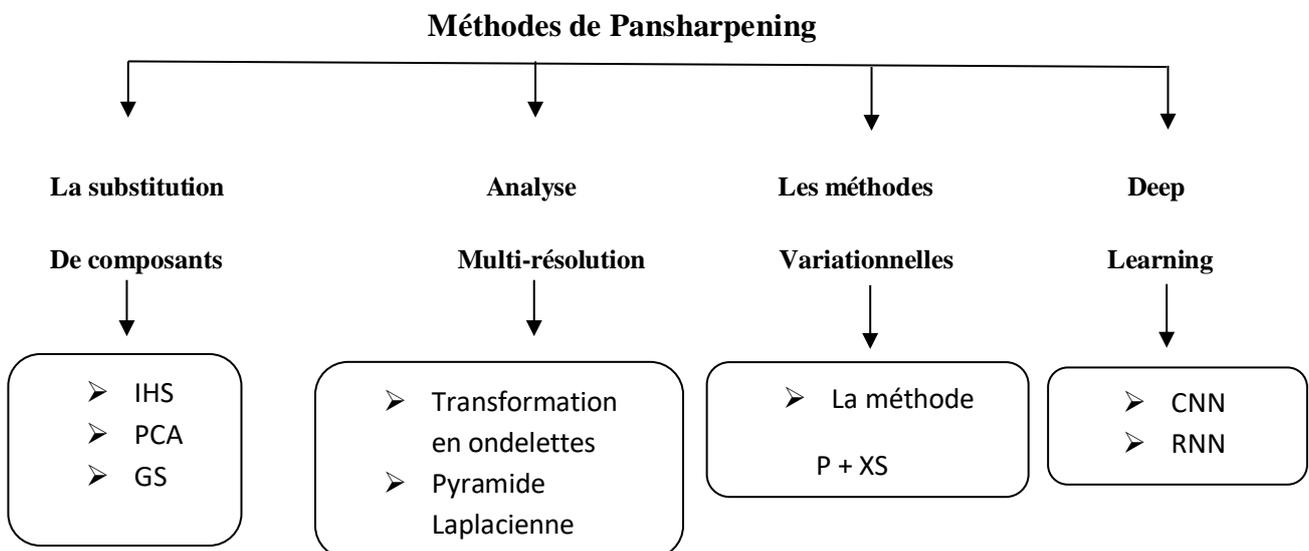


Figure 2.6 : schéma représente les méthodes de Pansharpening

2.11.1 Méthodes à substitution de composantes

Ces méthodes sont basées sur la transformation de l'image de résolution spectrale élevée d'un espace vers un autre, dans le but de séparer l'information spectrale de l'information spatiale.

Une fois la composante contenant l'information spatiale isolée, cette dernière est remplacée par l'image panchromatique qui possède une meilleure résolution spatiale.

Enfin, l'image fusionnée est obtenue par la transformation inverse pour le retour vers l'espace d'origine.

Cette catégorie de méthodes comprend de nombreuses approches populaires pour le pansharpening, parmi ces approches [21] :

la transformation Intensity Hue Saturation (IHS)

Cette méthode fait passer l'image multispectrale "sur-échantillonnée spatialement" (à la résolution spatiale de l'image panchromatique) de l'espace RVB (espace Rouge Vert et Bleu) à l'espace IHS (espace Intensité, Teinte et Saturation (ITS)), où la composante intensité est remplacée par l'image panchromatique. .

L'inconvénient de cette méthode est qu'elle se limite aux images multispectrales à trois bandes. [21].

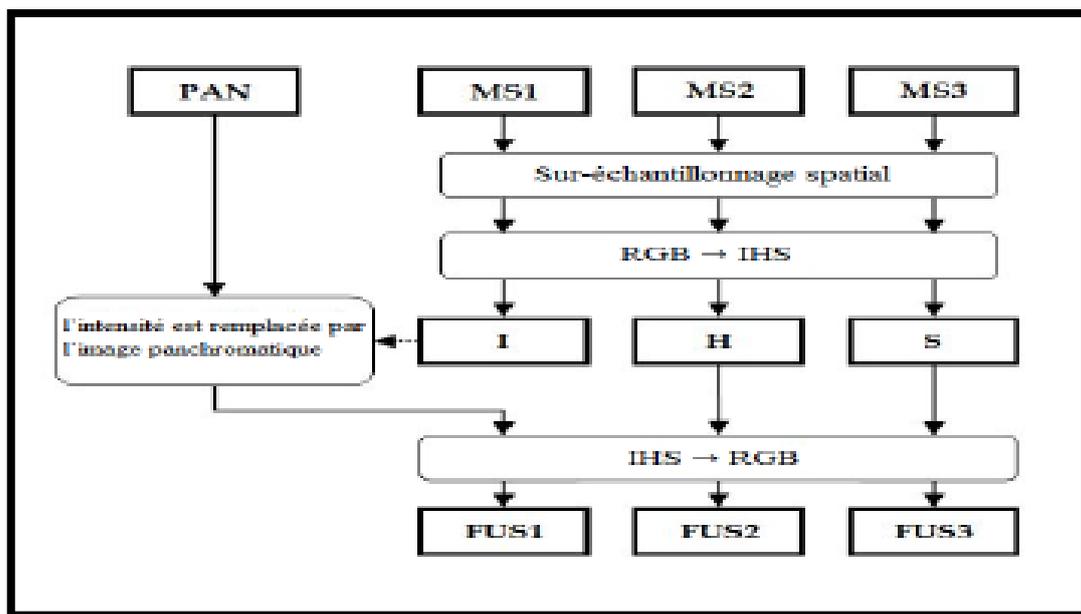


Figure 2.7 : Schéma de la méthode IHS [21]

L'Analyse en Composantes Principales (ACP)

Cette méthode présente l'avantage de pouvoir traiter un nombre de bandes spectrales supérieur à trois. Après avoir appliqué une ACP à l'image hyperspectrale/multispectrale sur-échantillonnée spatialement, afin de la ramener à la même résolution que l'image panchromatique, la première composante principale(PC1) est remplacée par la bande panchromatique. Par la suite, une transformation inverse est appliquée pour revenir à l'espace

de départ, une méthode de fusion fondée sur l'Analyse en Composantes Principales non linéaire est proposée. Celle-ci combine à la fois la réduction de dimensionnalité et la fusion pour améliorer la résolution spatiale des images hyperspectrales[21].

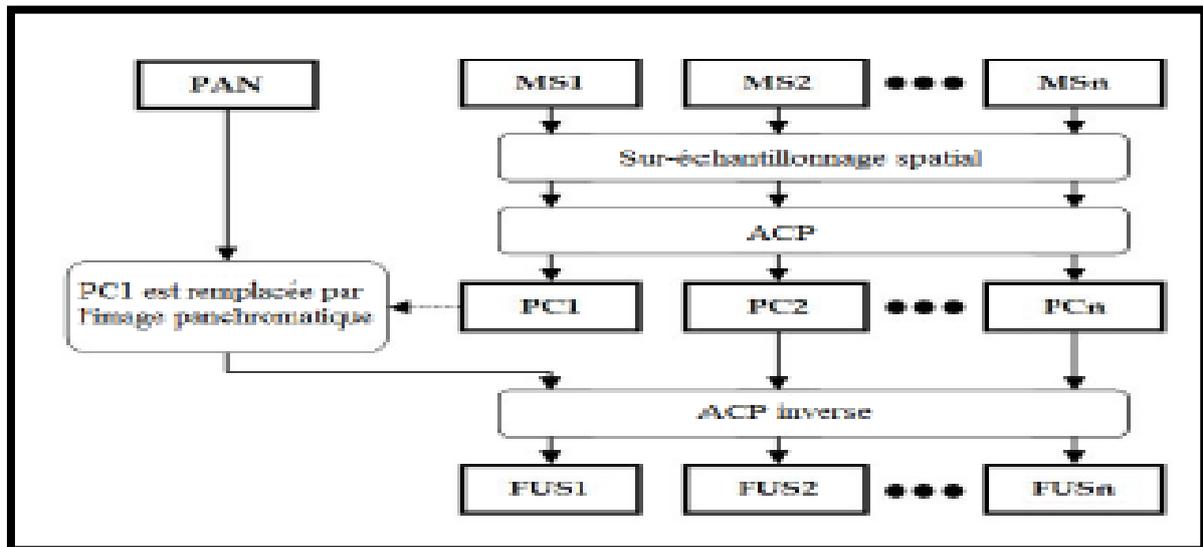


Figure 2.8: Schéma du principe de fusion par l'analyse en composantes principales [21]

Gram-Schmidt (GS)

Le processus de pansharpening commence par l'utilisation, comme première composante de l'entrée de la transformation GS, d'une image panchromatique synthétique à faible résolution (à la même résolution spatiale que l'image hyperspectrale ou multispectrale). Une décomposition orthogonale complète est ensuite effectuée, en commençant par cette composante. La procédure de pansharpening est complétée en remplaçant cette composante par l'image panchromatique à haute résolution, et en inversant la décomposition.

Une version améliorée est proposée, appelée GS Adaptative (GSA), dans laquelle l'image panchromatique à faible résolution est générée par la moyenne pondérée des bandes de l'image hyperspectrale/multispectrale avec des poids estimés par la minimisation de l'erreur quadratique moyenne entre la composante estimée et la version sous-échantillonnée de l'image panchromatique. Les méthodes de cette catégorie sont généralement rapides et faciles à mettre en œuvre. Cependant, elles souffrent de distorsions spectrales en raison des différences entre les caractéristiques des images panchromatiques et des images hyperspectrales ou multispectrales [21].

2.11.2 Méthodes d'analyse multi résolution

Les méthodes de pansharpening basées sur l'ARM sont fondées sur l'injection de la haute fréquence, qui est extraite par filtres tels que la transformée d'ondelette et Laplacian pyramides de l'image HR Pan dans le rééchantillon image LR MS pour obtenir l'image HR MS fusionnée [22].

Pyramide laplacienne

La pyramide laplacienne (Laplacian Pyramid (LP)) est une technologie développée par Burt et Adelson qui ont été influencés par la pyramide de Gauss (Pyramide de Gauss) [22].

Dans la GP, l'image originale est convoluée avec un filtre gaussien, produisant une version filtrée passe-bas de l'image originale. La différence entre l'image originale et la version passe-bas filtrée représente le laplacien[22].

Transformation en ondelettes

La fusion basée sur une transformée en ondelettes est une approche bien connue des méthodes de pansharpening multirésolution. Le principe de ces méthodes est de fournir un cadre de décomposition des images dans une structure hiérarchique à résolution réduite, permettant de séparer des informations spatiales détaillées entre niveaux successifs. En général, les méthodes de fusion par ondelettes ont produit des résultats encourageants [22].

2.11.3 Méthodes variationnelles

L'idée principale de la méthode variationnelles est de développer l'énergie fonctionnelle basé sur certains attributs de l'image, dans lesquels la valeur énergétique fonctionnelle est faible. Correspond à des images de haute qualité. Différents termes pour leur énergie fonctionnelle sont utilisés selon la méthode considérée. Les méthodes de cette catégorie comprennent Plusieurs méthodes, la plus connue est la méthode P+XS :

Cette méthode a été développée pour améliorer la résolution spatiale de l'image SPOT multispectral (XS), utilisant une image panchromatique associée (P). Dans cette méthode, chaque bande spectrale de l'image fusionnée est obtenue par une combinaison linéaire des bandes spectrales de l'image multispectrale [23].

2.11.4 Méthodes fondées sur l'apprentissage profond

Depuis quelques années, un nouveau lexique lié à l'émergence de l'intelligence artificielle dans notre société inonde les articles scientifiques. Pour commencer, revenons sur ces trois définitions essentielles [24] :

- L'intelligence artificielle : c'est un champ de recherche qui regroupe l'ensemble des techniques et méthodes qui tendent à comprendre et reproduire le fonctionnement d'un cerveau humain
- Le Machine Learning : c'est un ensemble de techniques donnant la capacité aux machines d'apprendre automatiquement un ensemble de règles à partir de données. Contrairement à la programmation qui consiste en l'exécution de règles prédéterminées.
- Le Deep Learning ou apprentissage profond : c'est une technique de machine learning reposant sur le modèle des réseaux neurones : des dizaines voire des centaines de couches de neurones sont empilées pour apporter une plus grande complexité à l'établissement des règles, qui seront traité dans le chapitre 3.
-

2.12 Conclusion

Parmi les méthodes Pansharpening qui existent la méthode Deep Learning est la meilleur méthode de fusionner de l'image HR MS (on utilise l'image PAN et l'image LR MS pour fusionner Image HR MS) et permet de donner des informations de plus en plus précises sur les objets ,les surface et les profondeurs les plus inexplorés de notre univers .Dans le chapitre suivant les méthodes Pansharpening par Deep Learning seront détaillées pour la fusion d'image , dans le but d'améliorer la résolution spatiale de celle-ci. Comment la technique de Deep learning fonctionne ?

Chapitre 3 :
L'apprentissage
automatique (Machine
Learning) et
L'apprentissage profond
(Deep Learning)

3 Chapitre 03 : L'apprentissage automatique (Machine Learning) et L'apprentissage profond (Deep Learning)

3.1 Introduction

L'intelligence artificielle (IA) est en plein développement ces dernières années. Comprendre les dernières avancées dans ce domaine revient à étudier deux concepts très populaires tels que l'apprentissage automatique ou Machine Learning (ML) et l'apprentissage profond ou Deep Learning (DL).

Dans ce chapitre, nous allons présenter l'apprentissage automatique et ses algorithmes ainsi que ses applications dans différents domaines.

Une comparaison entre le Machine Learning et le Deep Learning sera présentée tout en montrant que l'apprentissage profond gagne en popularité en raison de ses performances dans différentes tâches.

Présente les différentes applications du Deep Learning dans le contexte routier afin de montrer ses différents usages dans ce domaine., nous nous concentrons sur les applications des réseaux de neurones dans la télédétection et Pansharpening .

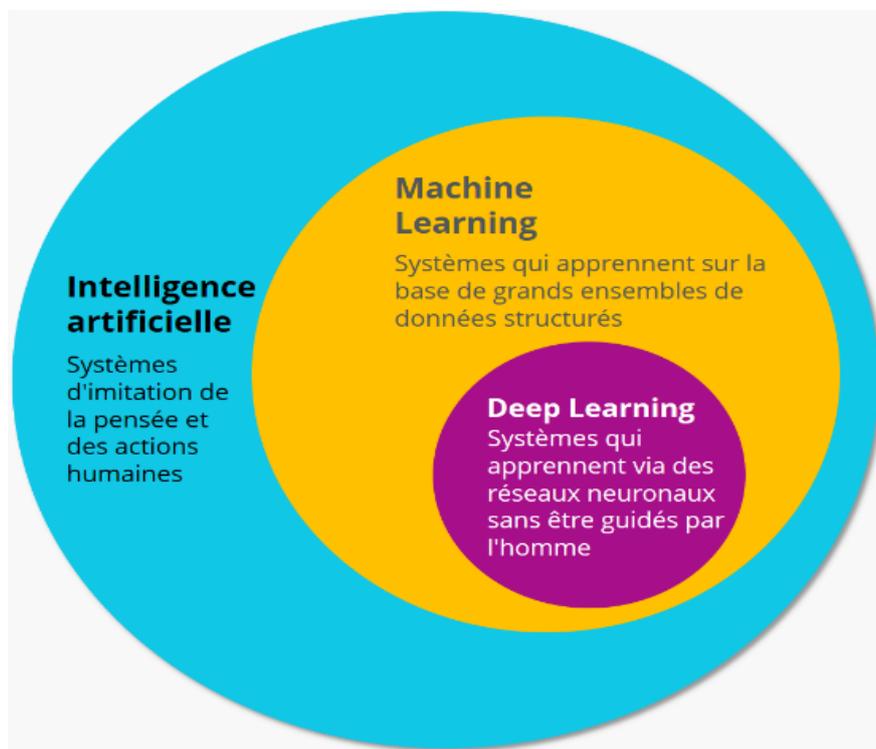


Figure 3.1 : La relation entre l'intelligence artificielle, le ML, et le deep learning[34]

3.2 Développement de l'IA

Dans les années 1940 et 1950, l'intelligence artificielle vise à simuler, sur des machines, les différentes facultés de l'intelligence. Cette discipline reposait sur l'idée selon laquelle toutes les fonctions cognitives (apprentissage, raisonnement, calcul, mémorisation...) pouvaient être décrites précisément par une notation logique de manière à être reproduites par des machines.

Les premières recherches sur les réseaux de neurones ont été menées dans ce contexte. Le premier modèle mathématique pour un neurone a été proposé en 1943 par Warren McCulloch et Walter Pitts. Ce neurone artificiel est un outil mathématique qui reçoit une ou plusieurs valeurs en entrées et qui produit une valeur en sortie. Il simule le fonctionnement d'un neurone biologique.

En 1957, Frank Rosenblatt invente le type de réseau de neurones le plus simple et la base des principaux systèmes utilisés aujourd'hui. Il s'agit de neurones interconnectés comme peuvent l'être ceux d'un cerveau humain : la sortie d'un neurone devenant l'une des entrées d'un autre neurone.

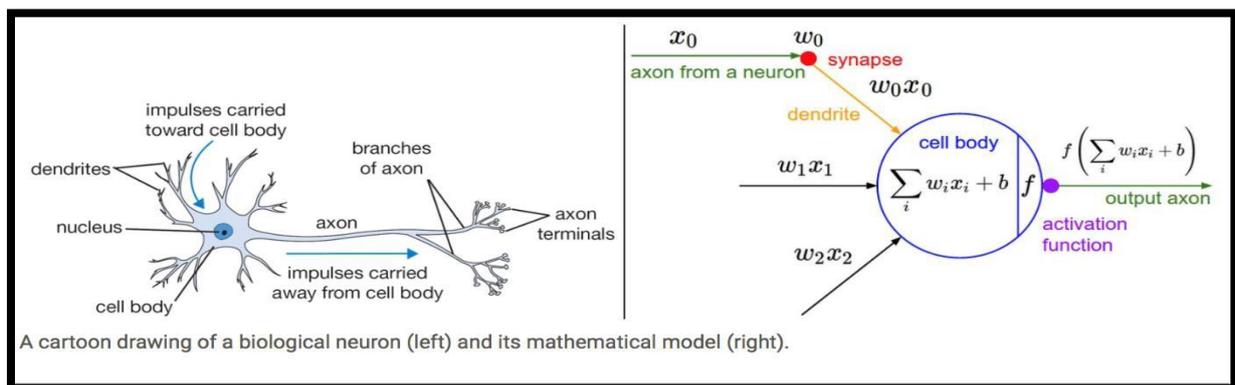


Figure 3.2 - Schéma de fonctionnement d'un neurone biologique et celui d'un neurone artificiel [26]

Un neurone artificiel est un petit modèle statistique, comme un neurone biologique, il doit synthétiser les informations fournies en entrée (l'équivalent des synapses et des dendrites du neurone biologique) et produire quelque chose de nouveau en sortie (l'équivalent de l'axone d'un neurone biologique), cette transformation a lieu grâce à une fonction d'activation qui caractérise le neurone.

Ces neurones peuvent être mis en réseau, ce qui permet d'obtenir des modèles de calcul plus complexe. Chacune de ces connexions reçoit un poids, noté w dans la figure 2.2 ci-dessus, permettant de déterminer l'impact de l'entrée x dans la fonction d'activation. L'apprentissage consiste à calculer les poids optimaux des différentes connexions, en utilisant des données sélectionnées avec soin. Les données disponibles conditionnent le type de réseau et ses caractéristiques (nombre de neurones dans la couche d'entrée et celle de sortie...)[25].

3.3 L'apprentissage automatique (Machine Learning)

Le Machine Learning (ML) est un sous-domaine de l'IA. Il a pour objectif de comprendre la structure des données et de les intégrer dans des modèles qui peuvent être compris et utilisés pour tester d'autres données. ML est né grâce aux technologies de reconnaissance de pattern et à la théorie selon laquelle les ordinateurs peuvent apprendre sans être programmés pour effectuer des tâches spécifiques (ex ; classification). Les chercheurs intéressés par l'intelligence artificielle souhaitent vérifier si les ordinateurs pouvaient apprendre à partir de données et s'adapter avec les nouvelles données.

L'idée générale est d'être capable de reproduire une relation grâce à une fonction (f). Cette fonction est identifiée par un algorithme d'apprentissage. Pour aboutir à la décision ou à la prédiction Y , l'algorithme utilise en entrée un jeu de données X afin de déterminer un modèle de classification [27]

Les principes de l'apprentissage :

- L'objectif est d'apprendre en produisant de nouvelles connaissances afin d'améliorer ses performances.
- Le manque de connaissances est comblé par l'abondance des données.
- Le machine Learning consiste à déduire des connaissances en utilisant des données dites d'entraînement.
- Machine learning développer un modèle en servant d'un algorithme d'optimisation pour minimiser l'erreur entre les modèles et nos données.

3.3.1 Approches d'apprentissage automatique :

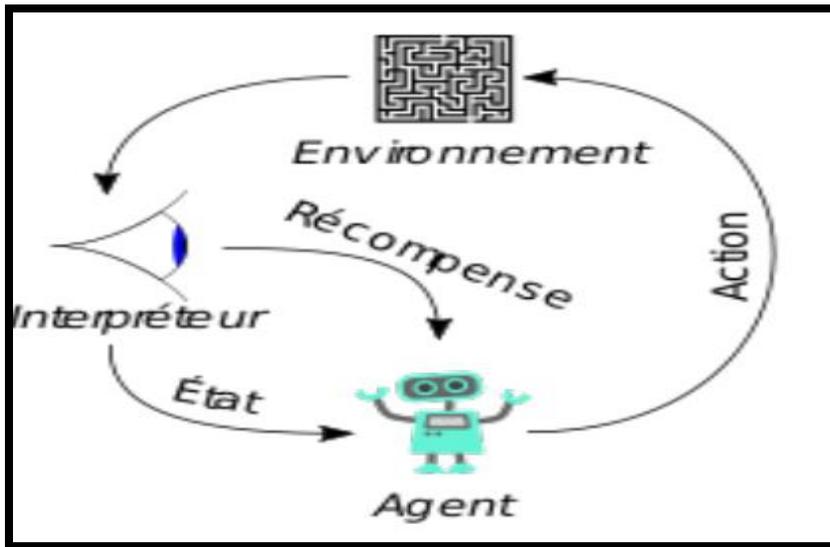
Il existe différents types d'algorithmes à utiliser chacun selon la tâche à accomplir et selon le type de données disponibles. Ces algorithmes se divisent en trois catégories : supervisés, non supervisés et par renforcement

3.3.1.1 L'apprentissage par renforcement :

Consiste, pour un agent autonome, à apprendre les actions à prendre, à partir d'expériences, de façon à optimiser une récompense au cours du temps. [29]

Dans l'apprentissage par renforcement, le réseau réagit aux données environnementales (appelées état) et contrôle les actions d'un agent pour tenter de maximiser une récompense.

La sortie du système est une séquence d'actions, la séquence d'actions proposée pour atteindre l'objectif fixé doit être le meilleur possible.



- Décisions en temps réel
- Navigation robotique
- Apprentissage de tâches
- Acquisition de compétences

Figure 3.3 : L'apprentissage par renforcement [49]

3.3.1.2 L'apprentissage non supervisé :

Les algorithmes d'apprentissage non supervisé utilisent un ensemble de données contenant de nombreuses caractéristiques, puis apprennent des propriétés utiles de la structure de cet ensemble de données. La majorité des algorithmes d'apprentissage non supervisés se basent sur la méthode de clustering (annexe 2) qui consiste à diviser la base de données en groupes d'exemples similaires, et regrouper des données similaires, mais sans annotation et la méthode réduction de dimension (annexe 3) [25].

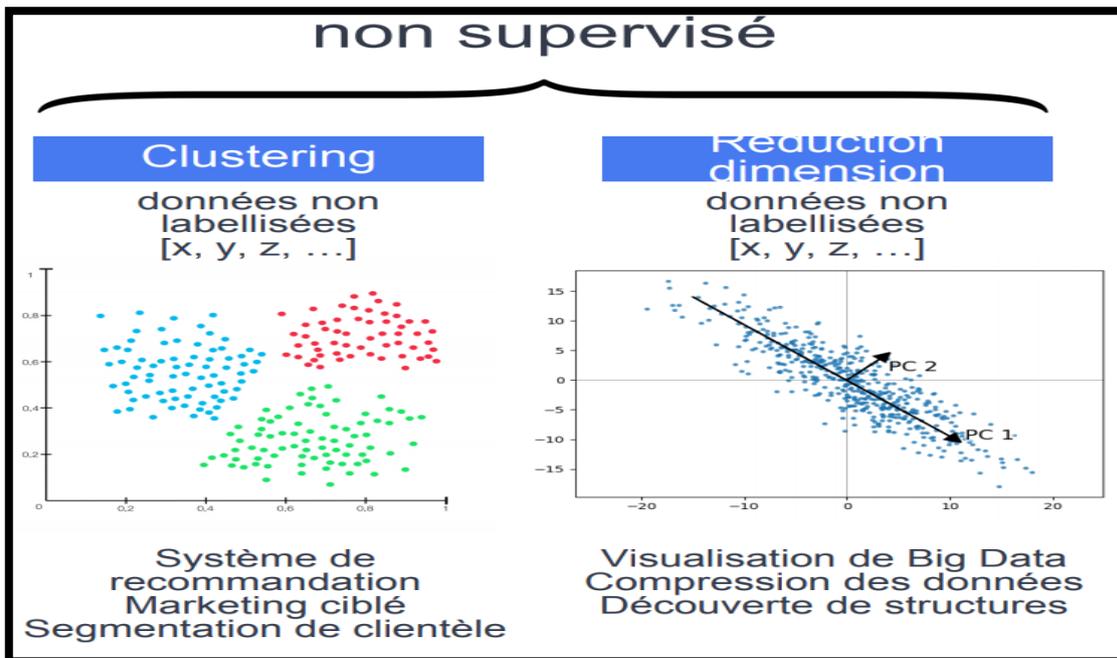


Figure 3.4 : Apprentissage non supervisé [28]

3.3.1.3 L'apprentissage supervisé :

Les algorithmes d'apprentissage supervisé utilisent la base de données contenant des caractéristiques, mais chaque exemple est également associé à une étiquette ou à une cible. Un algorithme d'apprentissage supervisé peut étudier la base de données et apprendre à classer en espèces différentes en fonction de leurs mesures à partir deux grandes tâches **classification** et **régression** [27].

3.3.1.3.1 Les algorithmes d'apprentissage supervisés pour la régression (prédiction de valeurs) :

1. La régression linéaire uni variée (linear regression) :

Cet algorithme cherche à établir, sous forme d'une droite, une relation entre une variable expliquée et une variable explicative. Par exemple, prédire une note à un examen (variable expliquée) en fonction du nombre d'heures de révisions (variable explicative). En d'autres termes, les données d'une série d'observations sont représentées sous forme d'un nuage de points et l'on cherche à trouver une droite passant au plus près de ces points. [32]

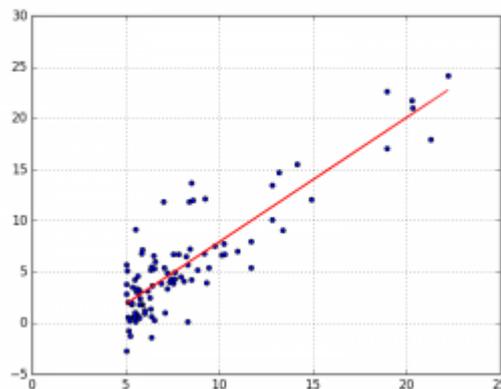


Figure 3.5 : La régression linéaire uni variée [56]

La méthode de descente de gradient

Cette notion est essentielle, car elle est appliquée dans divers algorithmes d'apprentissage du Machine Learning et du Deep Learning que nous verrons un peu plus loin dans cet ouvrage. Lorsqu'un système est en phase d'apprentissage, il commet des erreurs. Le taux d'erreur diminue au fur et à mesure de l'apprentissage, mais il se peut qu'à un moment donné l'erreur augmente pour à nouveau rediminuer et atteindre un niveau d'erreur plus bas que le précédent qui est le niveau optimal d'apprentissage [32]

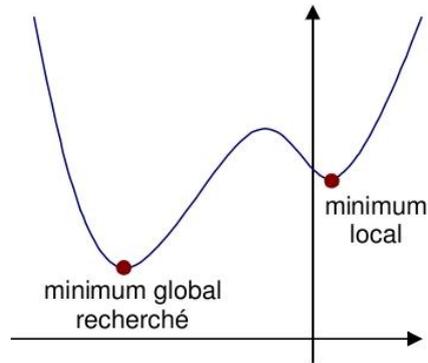


Figure 3.6 : La descente de gradient [56]

L'algorithme du gradient consiste donc à trouver par itérations successives le minimum global de la fonction de coût (erreur). Par analogie souvent reprise dans la littérature, imaginez-vous-en haut d'une montagne avec pour objectif d'atteindre la plaine en contre bas. À chaque pas, vous analysez votre situation et décidez d'avancer de quelques pas, quitte à remonter pour prendre le chemin qui mène au but. Le gradient correspondant à la pente du sol que vous êtes en train de parcourir. Le "pas" porte également le nom de taux d'apprentissage dont nous verrons la mise en pratique dans le chapitre Un neurone pour prédire.

Régression polynomiale (polynomial régression)

Il est parfois difficile de trouver une droite pouvant passer parmi les points de la série d'observations de façon optimale. Cependant, il est parfois possible de trouver un lien entre les variables à l'aide d'une courbe. C'est ce que permet la régression polynomiale en ajoutant des plis à la courbe à l'aide d'éléments appelés polynômes. [32]

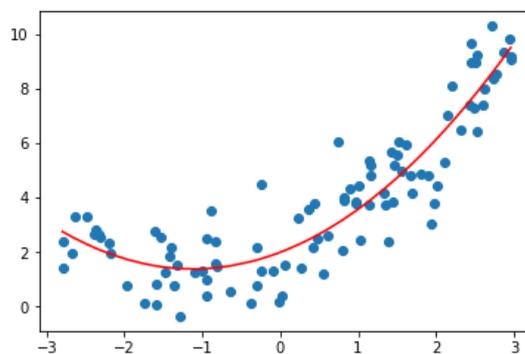


Figure 3.7 : La Régression polynomiale

3.3.1.3.2 Les algorithmes d'apprentissage supervisés pour la classification :

Classification binaire

Dans le cas où les étiquettes sont binaires, elles indiquent l'appartenance à une classe. On parle alors de classification binaire.

Définition (Classification binaire) : Un problème d'apprentissage supervisé dans lequel l'espace des étiquettes est binaire, autrement dit $Y = \{0, 1\}$ est appelé un problème de classification binaire.

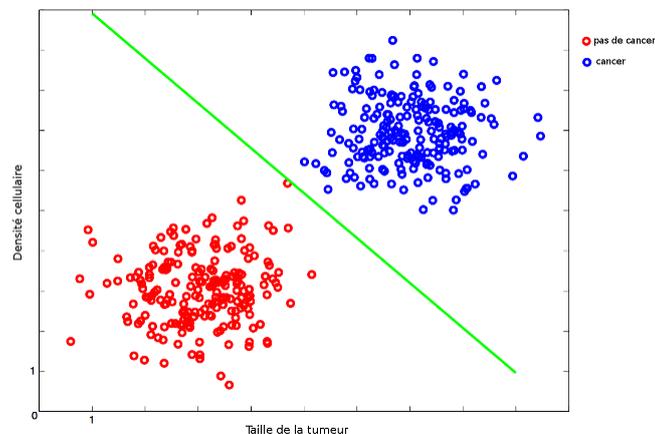


Figure 3.8 : Classification binaire [62]

Classification multi-classe

Dans le cas où les étiquettes sont discrètes, et correspondent donc à plusieurs classes, on parle de classification multi-classe.

Définition (Classification multi-classe) : Un problème d'apprentissage supervisé dans lequel l'espace des étiquettes est discret et fini, autrement dit $Y = \{1, 2, \dots, C\}$ est appelé un problème de classification multi-classe. C est le nombre de classes. [32]

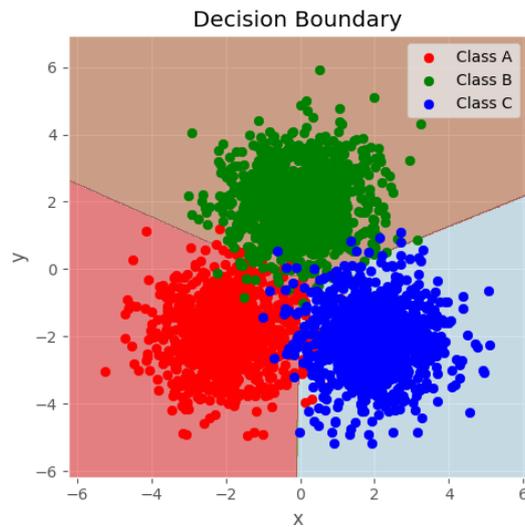


Figure 3.9 : Classification multi-classe [57]

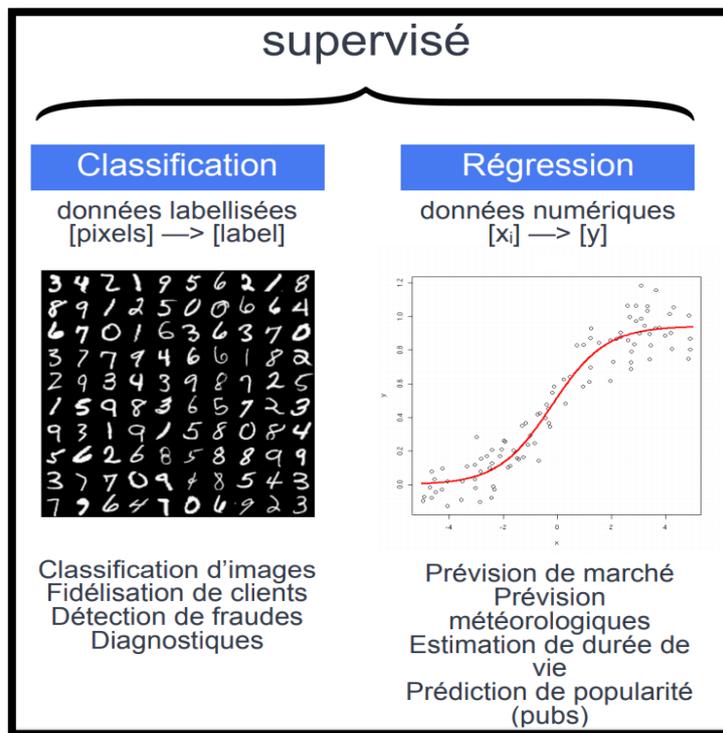


Figure 3.10 : Apprentissage supervisé [28]

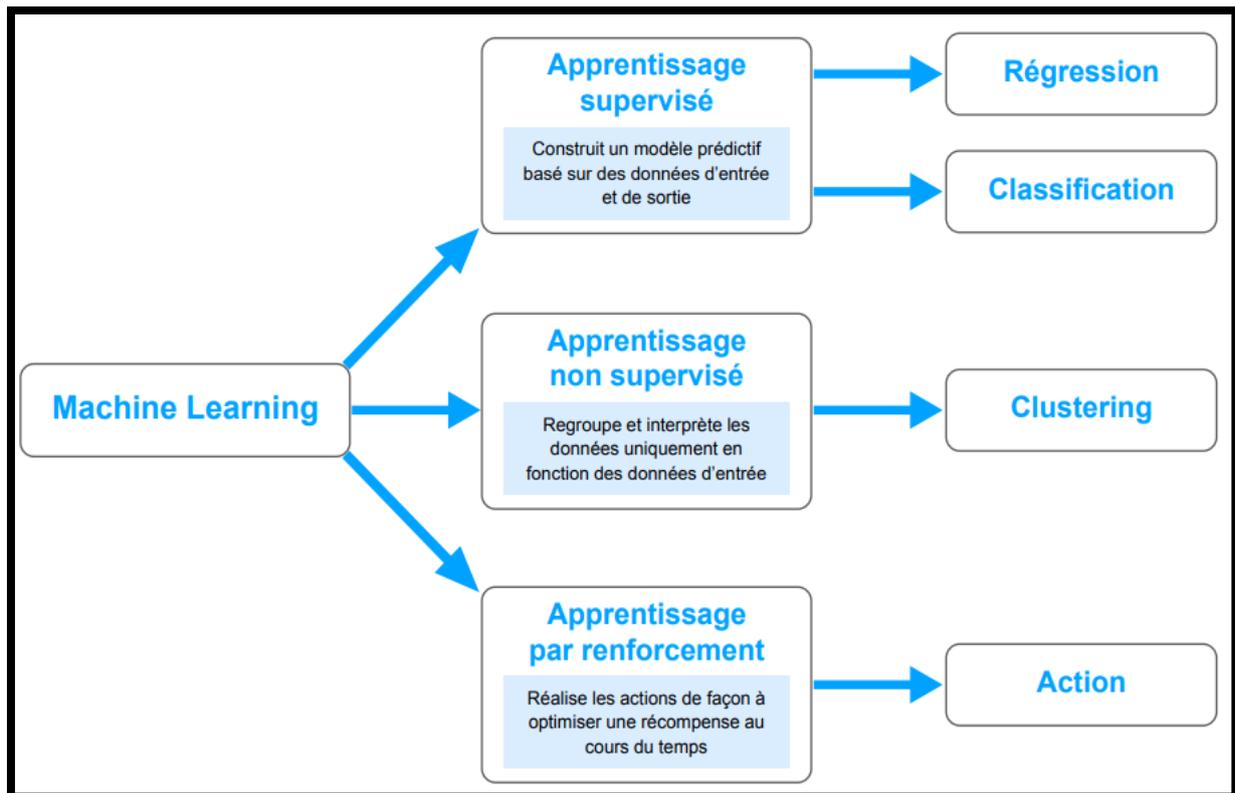


Figure 3.11 : Architecteur de différents Méthodes de ML [28]

3.4 Apprentissage profond (deep Learning)

Le deep learning ou apprentissage profond est un sous-domaine de l'intelligence artificielle (IA). Ce terme désigne l'ensemble des techniques d'apprentissage automatique (machine learning), autrement dit une forme d'apprentissage fondée sur des approches mathématiques, utilisées pour modéliser des données. Pour mieux comprendre ces techniques, il faut remonter aux origines de l'intelligence artificielle en 1950, année pendant laquelle Alan Turing s'intéresse aux machines capables de penser.

En 1958, F. Rosenblatt développe le modèle du Perceptron. C'est un réseau de neurones inspiré du système visuel. Il possède deux couches de neurones : une couche de perception et une couche liée à la prise de décision. C'est le premier système artificiel capable d'apprendre par expérience

Cette réflexion va donner naissance au machine learning, une machine qui communique et se comporte en fonction des informations stockées. Le deep learning est un système avancé basé sur le cerveau humain, qui comporte un vaste réseau de neurones artificiels.

Ces neurones sont interconnectés pour traiter et mémoriser des informations, comparer des problèmes ou situations quelconques avec des situations similaires passées, analyser les solutions et résoudre le problème de la meilleure façon possible.

Comme chez les êtres humains, le deep learning consiste à apprendre des expériences vécues ou, dans le cas des machines, des informations enregistrées [26].

Pour quoi le deep learning ?

Le deep learning est d'une grande utilité dans l'univers des technologies de l'information et de la communication.

Ces technologies sont aussi présentes dans les systèmes satellitaire (télédétection), Détermination et classification des caractéristiques des images satellites de divers terrains, plans d'eau et mouvements des nuages ...etc., dans divers domaines militaires, agricoles, navals et aériens ..ect.

Et en médecine pour établir un diagnostic à partir d'un examen d'imagerie (radio, IRM, scanner), en physique pour rechercher des particules et dans le domaine artistique pour reproduire une œuvre.

Principe de fonctionnement de deep learning ?

Le Deep learning (apprentissage profond) n'a pas besoin de données structurées. Le système fonctionne à partir de plusieurs couches de réseaux neuronaux, qui combinent différents algorithmes en s'inspirant du cerveau humain. Ainsi, le système est capable de travailler à partir de données non structurées.

Cette approche est particulièrement adaptée pour les tâches complexes, lorsque tous les aspects des objets à traiter ne peuvent pas être catégorisés en amont. Le système du Deep learning identifie lui-même les caractéristiques discriminantes. Dans chaque couche, il recherche un nouveau critère spécifique de l'objet, qui sert de base pour décider de la classification retenue pour l'objet à la fin du processus. [34]

Différence entre ML et DL ?

Une des grandes différences entre le Deep Learning et les algorithmes de ML traditionnelles c'est qu'il s'adapte bien, plus la quantité de données fournie est grande plus les performances d'un algorithme de Deep Learning sont meilleurs. Contrairement à plusieurs algorithmes de ML classiques qui possèdent une borne supérieure à la quantité de données qu'ils peuvent recevoir des fois appelée "plateau de performance", les modèles de Deep Learning n'ont pas de telles limitations (théoriquement) et ils sont même allés jusqu'à dépasser la performance humaine dans des domaines comme l'image processing.

Autre différence entre les algorithmes de ML traditionnelles et les algorithmes de Deep Learning c'est l'étape de l'extraction de caractéristiques. Dans les algorithmes de ML traditionnelles l'extraction de caractéristiques est faite manuellement, c'est une étape difficile et coûteuse en temps et requiert un spécialiste en la matière alors qu'en Deep Learning cette étape est exécutée automatiquement par l'algorithme. [33]

3.4.1 Fonctionnement d'un réseau de neurones

Un réseau de neurones artificiels est un système de technologie de l'information basé sur le fonctionnement du cerveau humain, dont sont équipés les ordinateurs dotés de fonctions d'intelligence artificielle. [34]

La conception des réseaux de neurones artificiels s'appuie sur la structure des neurones biologiques du cerveau humain, les réseaux de neurones artificiels peuvent être décrits comme des systèmes composés d'au moins deux couches de neurones

- une couche d'entrée et une couche de sortie
- et comprenant généralement des couches cachées (« hidden layers »).

Plus le problème à résoudre est complexe, plus le réseau de neurones artificiels doit comporter de couches. Chaque couche contient un grand nombre de neurones artificiels spécialisés. [28]

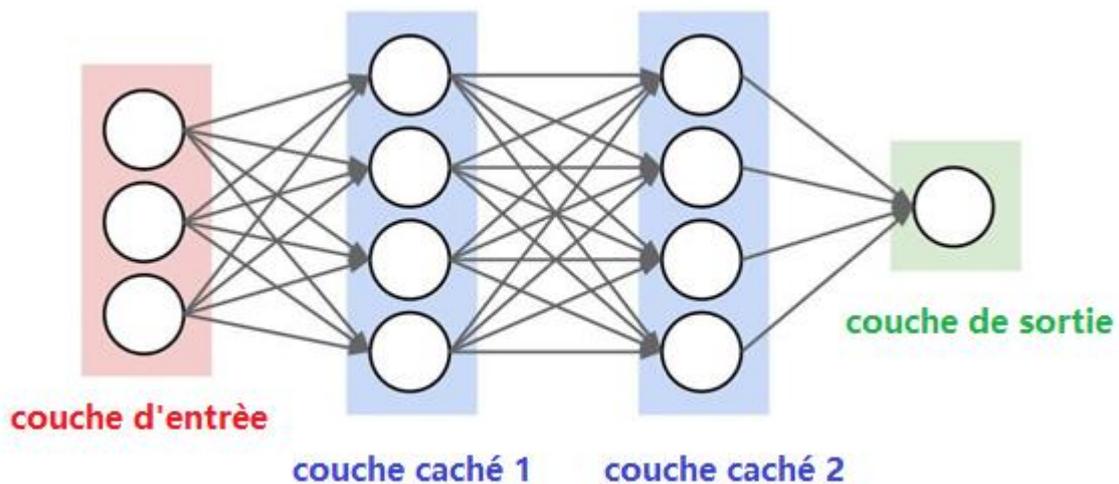


Figure 3.12: Réseaux de neurone artificiel [46]

- **Traitement de l'information au sein du réseau de neurones**

Au sein d'un réseau de neurones artificiels, le traitement de l'information suit toujours la même séquence : les informations sont transmises sous la forme de signaux aux neurones de la couche d'entrée, où elles sont traitées. À chaque neurone est attribué un « poids » particulier, et donc une importance différente. Associé à la fonction dite de transfert, le poids permet de déterminer quelles informations peuvent entrer dans le système.

À l'étape suivante, une fonction dite d'activation associée à une valeur seuil calculent et pondèrent la valeur de sortie du neurone. En fonction de cette valeur, un nombre plus ou moins grand de neurones sont connectés et activés.

Cette connexion et cette pondération dessinent un algorithme qui fait correspondre un résultat à chaque entrée. Chaque nouvelle itération permet d'ajuster la pondération et donc l'algorithme de façon à ce que le réseau donne à chaque fois un résultat plus précis et fiable. [32]

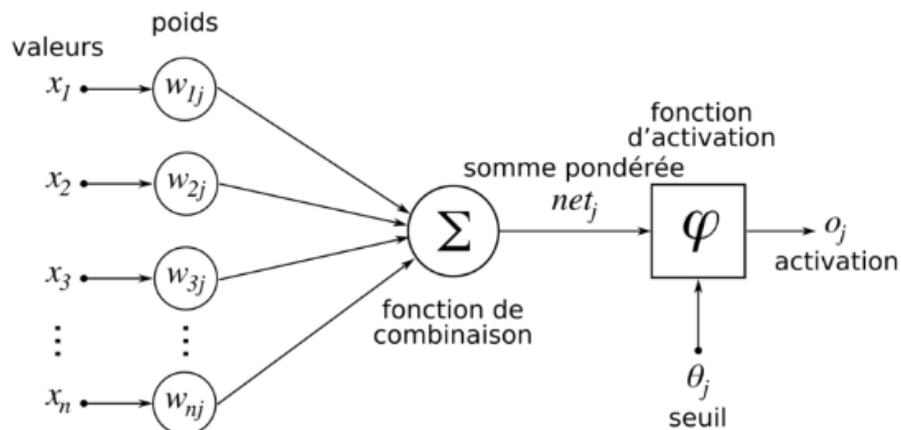


Figure 3.13: Architecteur de Réseaux de neurone [47]

Les principales composantes du réseau de neurones :

➤ Neurons (ensemble de fonctions) :

Ils prennent une donnée d'entrée et produisent une donnée de sortie. Un certain nombre de neurones sont groupés en couches (ou layer). Tous les neurones du même groupe remplissent un type de fonction similaire.

Les neurones d'entrée reçoivent des données d'entrée, les traitent et les transmettent aux neurones dans la couche suivante. Les neurones cachés prennent les données de sortie des précédents neurones en entrée, calculent de nouvelles données de sortie et les transmettent à des couches successives.

Dans un réseau de neurones (voir 3.14), les neurones de la dernière couche cachée (hidden layer) transmettent les données de sortie en entrée des neurones de la couche de sortie (output layer). A partir de cela les neurones de la couche de sortie produisent les données de sortie finales.

➤ Couches (groupement de neurones) :

Les couches (ou layers) contiennent des neurones et aident à faire circuler l'information. Il existe au moins deux couches dans un réseau de neurones : la couche d'entrée (input layer) et la couche de sortie (output layer).

Les couches, autres que les couches d'entrée et de sortie, sont appelées les couches cachées (ou hidden layers).

➤ Poids et biais (valeurs numériques) :

Les poids et biais sont des variables du modèle qui sont mises à jour pour améliorer la précision du réseau, un poids est appliqué à l'entrée de chacun des neurones pour calculer une donnée de sortie.

Les réseaux de neurones mettent à jour ces poids de manière continue. Il existe donc une boucle de rétro-action mise en œuvre dans la plupart des réseaux de neurones.

Les biais sont également des valeurs numériques qui sont ajoutées une fois que les poids sont appliqués aux valeurs d'entrée. Les poids et les biais sont donc en quelque sorte des valeurs d'auto-apprentissage de nos réseaux de neurones.

➤ Fonction d'activation (algorithmes mathématiques appliqués aux valeurs de sortie) :

Les fonctions d'activation lissent ou normalisent la donnée de sortie avant qu'elle ne soit transmise aux neurones suivants. Ces fonctions aident les réseaux de neurones à apprendre et à s'améliorer.

Comme son nom l'indique, la fonction d'activation est une formule mathématique (algorithme) activée dans certaines circonstances. Lorsque les neurones calculent la somme pondérée des valeurs d'entrée + le biais, elles sont transmises à la fonction d'activation, qui vérifie si la valeur calculée est supérieure au seuil requis.

Si la valeur calculée est supérieure au seuil requis, la fonction d'activation est activée et une valeur de sortie est calculée.

Si nous traçons les sorties par les fonctions d'activation, nous obtiendrons une courbure.

La pente de la courbe est utilisée pour calculer le gradient, et le gradient nous aide à comprendre le taux de changement et les relations entre les variables. [37]

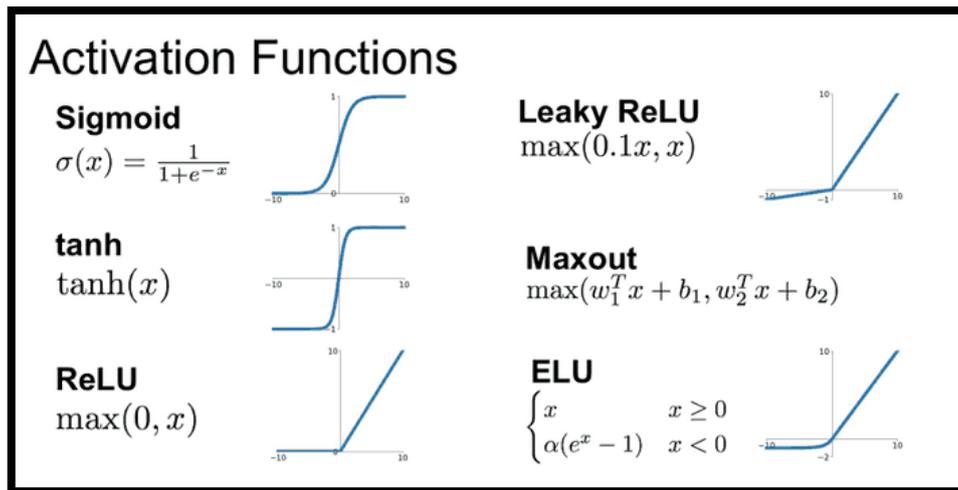


Figure 3.14: Types de fonctions d'activation [37]

- **Méthodes d'apprentissage**

Pour veiller à ce que les connexions au sein d'un réseau de neurones artificiels soient correctement établies, il faut au préalable procéder à son « entraînement ». On peut distinguer ici deux procédures de base : l'apprentissage supervisé et l'apprentissage non supervisé. [35]

L'apprentissage supervisé :

- Le réseau apprend à classifier des entrées en s'appuyant sur l'ensemble des exemples avec les réponses correctes associées (plus généralement il s'agit de la régression non-linéaire)

L'apprentissage non supervisé :

- Les réponses correctes ne sont pas données
- Le réseau classifie les entrées en s'appuyant sur la similarité entre différentes entrées (plus généralement il s'agit de clustering automatique)

	Type de réseau	Domaine d'application
Apprentissage supervisé	Réseaux de neurones artificiels	Régression et Classification
	Réseaux de neurones à convolution	Vision numérique
	Réseaux de neurones récurrents	Séries temporelles
Apprentissage non supervisé	Cartes auto-adaptatives	Détection de caractéristiques
	Machines de Boltzmann profondes	Systèmes de recommandation
	Auto-encodeurs	Systèmes de recommandation
	Réseaux antagonistes génératifs	Génération de contenus réalistes

Figure 3.15: Apprentissage de Réseaux de neurone [28]

Les architectures de réseaux neuronaux peuvent être divisées différentes familles :

- Réseaux de neurones Feed forwarded (propagation avant)
 - Réseaux de neurones récurrent (RNN)
 - Réseaux de neurones à résonance
 - Réseaux de neurones auto-organisés
 - Réseau de neurones de convolution – Convolution Neural Network (CNN)
 - Réseau de neurones générative (GANS)
 - Réseaux de neurones auto-encodeur (DAE)
-
- **Les domaines d'application des réseaux neuronaux sont souvent caractérisés par une relation entrée-sortie de la donnée d'information :**
 - La reconnaissance d'image
 - Les classifications de textes ou d'images
 - Identification d'objets
 - Prédiction de données
 - Filtrage d'un set de données

3.5 Deep Learning en traitement d'images

Le Deep Learning et le traitement de l'image sont deux domaines d'un grand intérêt pour les universitaires et les professionnels de l'industrie. Les domaines d'application de ces deux disciplines varient largement, englobant des domaines tels que la médecine, la robotique, la sécurité et la surveillance. Aucun apprentissage profond (Deep Learning) ne nuit le traitement d'image. Il faut d'énormes ensembles de données et de nombreuses ressources de calcul pour faire le Deep Learning. Il existe de nombreuses applications où il est souhaitable de pouvoir traiter des images avec moins de charge de calcul et des empreintes de mémoire plus petites et sans avoir accès à des bases de données volumineuses.

Le deep Learning est en réussite en ce moment car il existe des résultats très impressionnants à la classification. la classification est un problème parmi beaucoup d'autres traités par le traitement d'images, même s'il était vrai que le Deep Learning résoudrait tous les problèmes de classification, il y aurait beaucoup d'autres types de traitement d'image à faire. [37].

- Réduction du bruit,
- Enregistrement d'image
- Calcul de mouvement
- Morphing (procédé numérique consistant à transformer progressivement une image en une autre)
- Les modes de fusion
- Affinement (rendre une image plus nette)
- Calcul de géométries
- Estimation 3D
- Modèles de mouvement 3D temporel
- Vision stéréoscopique
- Compression et codage de données
- Segmentation d'images
- Suppression du flou de l'images)
- Stabilisation de mouvement,
- Infographie.

3.6 Deep Learning et la télédétection

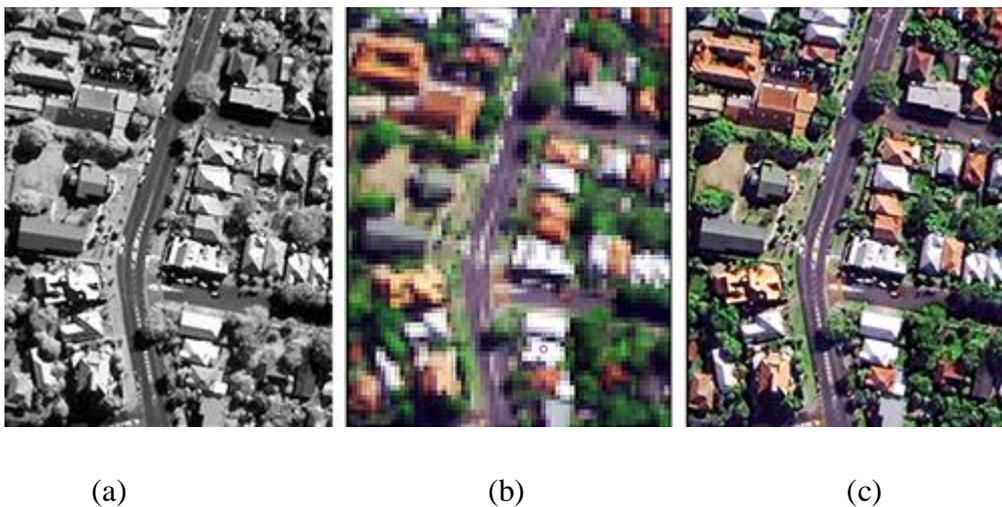
Dans cette partie on présente une méthode fusion d'images basée sur l'apprentissage profond pour la fusion des dans les applications de télédétection.

Pansharpening ou fusion d'images spectrales et spatiales attributs est couramment utilisé en télédétection. La fusion processus est effectué afin de remédier à la limitation de la bande

passante dans la capture de données d'images par des satellites à un certain rapport signal/bruit (SNR). Deux types de données sont saisies par satellites :

Données d'image PANchromatique (PAN) fournissant des données spatiales informations.
Données d'image MultiSpectral (MS) fournissant informations spectrales.

Après les étapes de prétraitement d'enregistrement et de réduction du bruit, le processus de fusion est appliqué à fusionner le contenu spectral des données d'image MS avec le contenu des données d'image PAN. L'efficacité de la fusion est souvent évaluée par le niveau de qualité dans le résultat fusionné.



((a) image PAN, (b) image MS base résolution, (c) image MS haut résolution)

Figure 3.16 : Pansharpening ou fusion d'images spectrales et spatiales [48]

En résumé, les principales raisons qui ont conduit à la popularité des méthodes basées sur DL pour la fusion d'images incluent les aspects suivants :

- 1) DL les modèles peuvent extraire automatiquement les caractéristiques les plus efficaces des données pour surmonter la difficulté de la conception manuelle.
- 2) Les modèles DL peuvent caractériser des relations très complexes entre les données d'entrée et de ciblage.
- 3) La communauté DL propose des approches potentielles de représentation d'images qui pourraient être utiles à l'étude de la fusion d'images.
- 4) La disponibilité de bibliothèques DL conviviales (par exemple, Caffe, TensorFlow, Theano et MatConvNet) et des ensembles de données d'images à grande échelle (par exemple, CIFAR, PASCAL VOC et ImageNet) assure des outils pratiques sur ce sujet. [41]

➤ Quelques modèles des méthodes basées sur DL pour la fusion d'images :

3.6.1 Fusion base sur Convolutional Neural Network (CNN)

CNN est l'un des réseaux les plus populaires développé ces dernières années en raison de ses excellentes performances en classification d'images détection de cibles, la reconnaissance faciale et la détection des piétons, le CNN a été appliqué avec succès au domaine de la télédétection pour résoudre les problèmes de classification d'images hyper spectrales.

CNN obtient des propriétés d'invariance de décalage, d'échelle et de distorsion en fusionnant trois idées architecturales de base, qui comprennent les champs récepteurs locaux, les poids partagés et le sous-échantillonnage.

Premièrement, les champs récepteurs locaux indiquent que chaque neurone à une certaine couche convolutive sera connecté uniquement à une région spatialement voisine de sa couche précédente, ce qui aide le réseau à extraire caractéristiques visuelles primaires. En d'autres termes, les neurones de la couche convolutive l prennent un sous-ensemble de neurones de la couche $l - 1$ en entrée.

Deuxièmement, les poids partagés signifient que le poids d'un noyau convolutif reste le même lorsqu'il est utilisé pour générer une carte de caractéristiques à une certaine couche. Par conséquent, le nombre de paramètres requis être formé au CNN diminuera considérablement par rapport avec celui des réseaux de neurones conventionnels [45].

Troisièmement, le sous-échantillonnage réduit la résolution spatiale d'une carte de caractéristiques, qui se combine avec l'opérateur de convolution pour obtenir une invariance de traduction.

Les CNN sont connectés localement via une opération convolutive et une stratégie de partage de poids, ce qui diminue considérablement la quantité d paramètres à apprendre [43]. Mathématiquement, la couche convolutive dans un 2D-CNN peut être exprimé comme :

$$Y^i = b^i + \sum k^{ij} * X^i \dots\dots\dots (1)$$

Où X^i est la i ème carte de caractéristiques en entrée et Y^i est la i ème carte de caractéristiques en sortie.

Les paramètres de réseau k^{ij} et b^i sont le noyau convolutif 2D et biais, respectivement. Le symbole * désigne la convolution 2D.

Pour une couche convolutive 3D, X^i et Y^i sont les deux cartes de caractéristiques 3D, k^{ij} est un noyau convolutif 3D, et le symbole * désigne la convolution 3D.

Les activations non linéaires largement utilisées dans les CNN incluent la fonction sigmoïde, les unités linéaires rectifiées (ReLU), etc.

Par exemple, une couche convolutive et une couche Relue peuvent être conjointement

Exprimé comme :

$$Y^i = \text{MAX} (0, b^i + \sum k^{ij} * X^i) \dots\dots\dots (2)$$

La plupart des méthodes actuelles de fusion d'images de télédétection comportent généralement deux étapes : l'extraction de caractéristiques et la fusion de caractéristiques, tout en mettant CNN pour traiter les problèmes de reconstruction de super-résolution d'image.

Plus précisément, les entrées et les étiquettes pour l'apprentissage du réseau sont respectivement des images à basse résolution et à haute résolution, pour utiliser CNN pour la fusion d'images de télédétection, nous adoptons les mêmes réflexions dans le domaine de la reconstruction d'images en super-résolution.

Fusion vise à générer une image MS avec une haute résolution spatiale, par conséquent, l'étiquette du réseau est une haute image MS à résolution spatiale et les entrées sont une image PAN et une image MS à faible résolution spatiale.

À l'heure actuelle, il existe peu d'études basées sur l'apprentissage en profondeur dans la fusion d'images de télédétection, qui est un réseau populaire pour la reconstruction d'images en super-résolution, pour mettre en œuvre la fusion d'images de télédétection.

Cependant, les premières deux méthodes n'utilisent que trois couches convolutives et ne peuvent exploiter adéquatement la profondeur du réseau pour extraire des profondeurs fonctionnelles. De plus, ces méthodes considèrent les images PAN comme une bande et le superposer sur les images MS pour former le réseau, ce qui ignore les caractéristiques distinctives de ces deux types d'images.

Alors que la troisième méthode utilise uniquement super résolution CNN pour faire de la super résolution d'images plutôt que de la fusion d'images. La procédure de fusion est toujours finie par la méthode traditionnelle de transformation de Gram-Schmidt [45].

Conception de réseau :

La figure 3.20 montre le programme de formation du RSCNN, où les images d'entrée MS et PAN générées par le satellite, la sortie du réseau doit être l'image MS avec la même résolution spatiale comme l'image PAN.

Cette image doit être aussi similaire que possible à l'image MS idéale (étiquette) qui a obtenu par un capteur avec la même résolution spatiale que l'image PAN.

Cependant, cette image MS idéale n'existe pas. Cela provoquera problèmes pour notre formation et évaluation de la performance s'il y a il n'existe pas d'étiquettes. Heureusement, ce problème peut être résolu en en utilisant le protocole de Wald.

Dans le protocole de Wald, le MS original les images sont considérées comme des étiquettes. Pour garder le même espace résolution que l'image MS d'origine (étiquettes), le PAN d'origine l'image doit être sous-échantillonnée en fonction du rapport entre la résolution de l'image MS et de l'image PAN. En même temps, l'image MS d'entrée à faible résolution spatiale est générée en sous-échantillonnant l'image MS d'origine et en interpolant le données sous-échantillonnées.

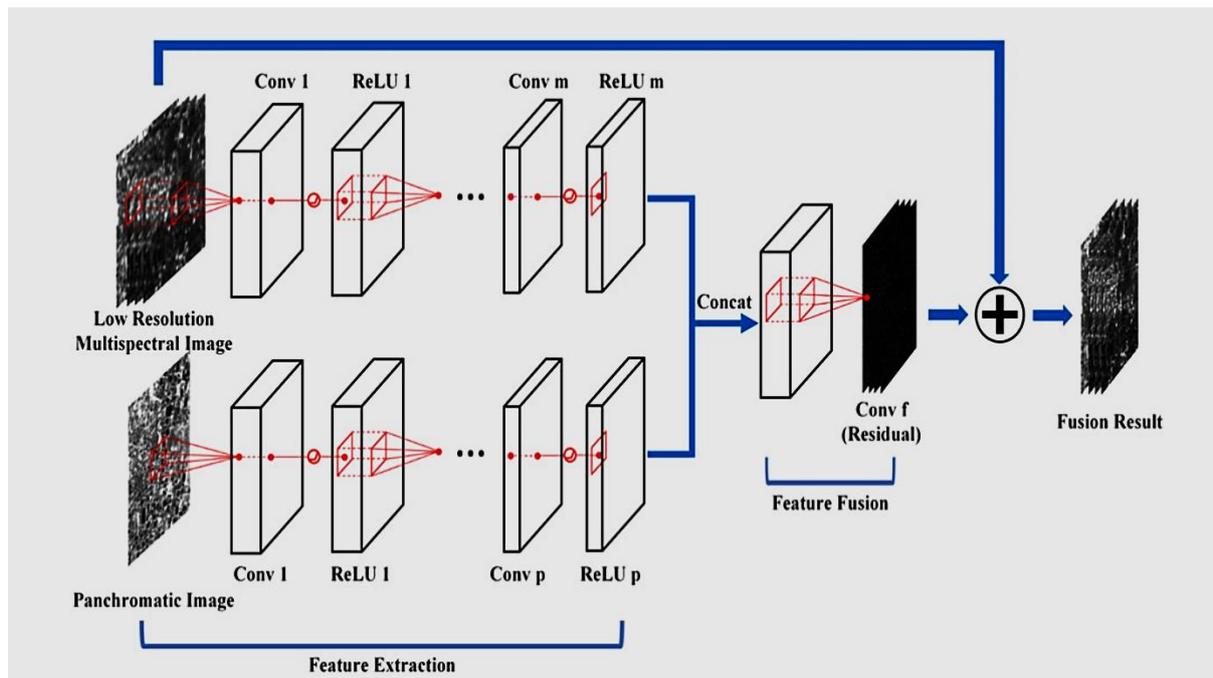


Figure 3.17: Architecture de RSCNN pour la fusion d'images de télédétection [45].

L'architecture est très flexible pour la rendre applicable à une variété de tâches de classification et de régression. De plus, la plupart des bibliothèques DL existantes comme Caffe, TensorFlow et MatConvNet développent le CNN lié fonctions et interfaces très bien, rendant les explorations basées sur CNN très pratique.

3.6.2 le Generative adversarial network (GAN)

Le GAN a d'abord été proposé par Goodfellow et al, qui est conçu à l'origine pour générer des images plus réalistes dans un environnement non supervisé chemin. Depuis lors, il a remporté de grands succès dans de nombreux domaines de la vision par ordinateur.

L'idée principale des GAN est de construire un jeu à deux joueurs minmax jeu entre l'apprentissage d'un générateur et d'un discriminateur.

Le générateur prend le bruit en entrée et essaie de générer différents échantillons pour tromper le discriminateur, tandis que le discriminateur vise à déterminer si un échantillon provient de la distribution du modèle ou de la distribution des données. [39]

La relation entre le générateur et le discriminateur se poursuit jusqu'à ce que les échantillons générés ne puissent pas être distingués par le discriminateur.

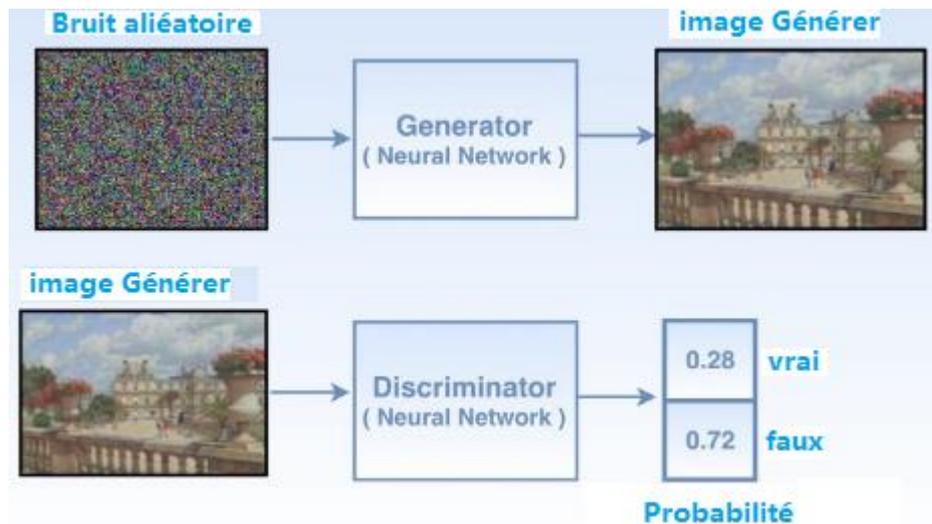


Figure 3.18 : Le générateur essaie de générer de fausses images en prenant du bruit aléatoire en entrée et le discriminateur essaie de le classer comme réel ou faux.

Le générateur peut alors capturer la distribution des données avec des échantillons d'images plus réalistes produits. Mathématiquement, un modèle génératif G vise à générer des échantillons, dont la distribution (P_G) essaie de se rapprocher la distribution (P_{data}) des données d'entraînement réelles, G et D jouent le minimax jeu à deux joueurs comme suit :

$$\min_G \max_D V_{GAN}(G, D) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))].$$

Les GAN d'origine peuvent bien fonctionner pour générer des images numériques. Cependant, il existe encore du bruit et de l'incompréhensibilité dans les résultats générés, en particulier pour les images haute résolution.

En outre pour améliorer la qualité des images générées, de nombreux travaux ont été proposés récemment. LAPGAN (a été utilisée La pyramide laplacienne) pour générer une image haute résolution supervisée par l'image basse résolution, mais cela ne fonctionne pas bien pour les images contenant des objets blancs [39].

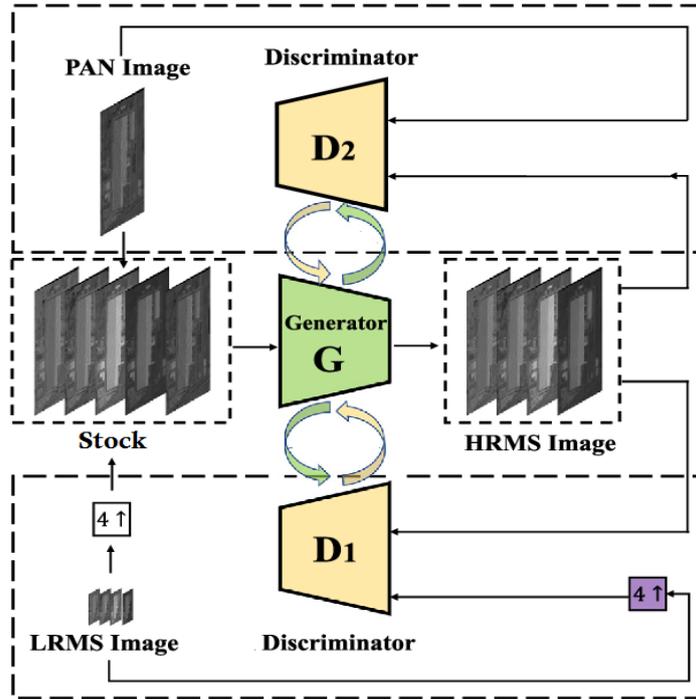


Figure 3.19 Architecture de pansharpening à base de (GAN) [39].

Dans il a réussi à générer des images de la nature, mais ne peut pas tirer parti les générateurs pour l'apprentissage supervisé. Aussi, pour améliorer la stabilité des GAN pendant le processus de formation, LAPGAN utilisée la fonction d'objective pour, minimiser la fonction perte, permet de minimiser la divergence Pearson χ^2 ,

LAPGAN a deux avantages par rapport aux GAN réguliers. Au d'une part, LAPGAN peut générer des images de meilleure qualité que la normale les GAN.

D'autre part, LAPGAN fonctionne de manière plus stable que la normale GAN pendant le processus d'apprentissage. Les fonctions objectives des LAPGAN sont définis comme suit [40] :

$$\min_D V_{\text{LSGAN}}(D) = \frac{1}{2} \mathbb{E}_{x \sim p_{\text{data}}(x)} [(D(x) - b)^2] + \frac{1}{2} \mathbb{E}_{z \sim p_z(z)} [(D(G(z)) - a)^2] \quad (1)$$

$$\min_G V_{\text{LSGAN}}(G) = \frac{1}{2} \mathbb{E}_{z \sim p_z(z)} [(D(G(z)) - c)^2] \quad (2)$$

3.7 Conclusion

Dans ce chapitre, nous avons parlé du Deep Learning et le Machin Learning en général, et sont relations avec Intelligence artificiel.

Puis, nous avons présenté les réseaux de neurones avec leur architecture détaillée.et leur utilisation dans traitement d'images et la télédétection généralement et la fusion d'image spécifiquement.

Et à la fin nous prestons quelques techniques de deep Learning dans le domaine de pansharpening.

Le prochain chapitre, traite les détails de la conception, ainsi que la méthode et les outils utilisés pour la réalisation de notre application.

Chapitre 4 :

Implémentation et

résultats

4 Chapitre4 : Implémentation et résultat

4.1 Introduction :

Ces dernières années les approches basées sur Deep Learning ont montré des résultats prometteurs dans le pansharpening des images satellites. Cependant, ils présentent des limites dans la production d'une fusion panoramique de haute qualité.

Après l'état de l'art que nous avons étudiés sur les méthodes de fusion d'images basées sur deep Learning, nous avons trouvé plusieurs algorithmes qui donnent de bons résultats comme, CNN, GAN, CDA, PNN, A-PNN, A-PNN-FT.

Après avoir implémenté ces méthodes, et d'après les résultats de tests de ces méthodes la méthode GAN donne les meilleurs résultats et avec des performances stables et fluides.

Comme nous traitons le pansharpening comme un problème de colorisation, qui apporte une perspective et une solution entièrement nouvelles au problème par rapport aux méthodes existantes qui fondent leur solution uniquement sur la production d'une version super-résolution (SR) de l'image multispectrale.

Dans ce chapitre, nous avons aussi implémenté la méthode PanColorisation ou (PanColorGAN), qui aide à surmonter les problèmes de perte de détails spatiaux et du flou qui sont observé dans le pansharpening basé sur CNN. Cette approche surpasse les précédentes méthodes.

4.2 Ressources matérielles et logicielles

4.2.1 Ressources matérielles

La configuration du matériel utilisé dans notre ordinateur est :

- Un PC portable HP i7, CPU 2.9 GHZ
- Carte graphique Nvidia GeForce GT525M
- RAM de taille 16 GO
- Disque dur de taille 500 GO
- Système d'exploitation : Windows 10, 64bit

4.2.2 Ressources logicielles

Python :

Le langage de programmation Python a été créé en 1989 par Guido van Rossum, aux Pays-Bas. Le nom Python vient d'un hommage à la série télévisée Monty Python's Flying Circus dont G. van Rossum est fan. La première version publique de ce langage a été publiée en 1991

Est un langage de programmation, un des langages de niveau supérieur, caractérisé par une écriture et une lecture simple, facile à apprendre, et programmation ouverte,

Python est un langage utilisé dans de nombreux domaines, tels que la création de programmes autonomes utilisant des interfaces graphiques connues et l'exécution de programmes Web, et son utilisation comme langage de script pour contrôler les performances des programmes les plus populaires.

En général, Python peut être utilisé pour programmer des programmes simples pour les débutants et pour réaliser de grands projets en même temps que tout autre langage de programmation. Il est souvent conseillé aux débutants en programmation d'apprendre cette langue car c'est l'une des langues de logiciel d'apprentissage les plus rapides [51].

Ce langage de programmation présente de nombreuses caractéristiques intéressantes :

- Il est multiplateforme. C'est-à-dire qu'il fonctionne sur de nombreux systèmes d'exploitation : Windows, Mac OS X, Linux, Android, iOS, depuis les mini-ordinateurs Raspberry Pi jusqu'aux supercalculateurs.
- Il est gratuit. Vous pouvez l'installer sur autant d'ordinateurs que vous voulez (même sur votre téléphone !).
- C'est un langage de haut niveau. Il demande relativement peu de connaissance sur le fonctionnement d'un ordinateur pour être utilisé.
- C'est un langage interprété. Un script Python n'a pas besoin d'être compilé pour être exécuté, contrairement à des langages comme le C ou le C++.
- Il est orienté objet. C'est-à-dire qu'il est possible de concevoir en Python des entités qui miment celles du monde réel (une cellule, une protéine, un atome, etc.) avec un certain nombre de règles de fonctionnement et d'interactions.
- Enfin, il est très utilisé en bio-informatique et plus généralement en analyse de données.
- Python est le langage de programmation le plus utilisé dans le domaine du Machine Learning, du Big Data et de la Data Science.

Toutes ces caractéristiques font que Python est désormais enseigné dans de nombreuses formations, depuis l'enseignement secondaire jusqu'à l'enseignement supérieur.

Python top des meilleures bibliothèques et packages :

- Pandas
- Agate
- Bokeh
- Scikit-learn
- TensorFlow

Anaconda :

Anaconda est une plateforme de gestion de paquets en Open source, d'environnement et de distribution des langages de programmation R et Python.

appliqué au développement d'applications dédiées à la science des données et à l'apprentissage automatique (traitement de données à grande échelle, analyse prédictive, calcul scientifique) adaptés pour Windows, Linux et MacOS.

Vous pouvez trouver le logiciel Anaconda à partir de son site officiel: <https://www.anaconda.com/>. Anaconda

Propose son propre gestionnaire de paquets appelé « conda » [52].

Les applications suivantes sont disponibles par défaut dans le navigateur :

- JupyterLab
- Jupyter Notebook
- Spyder
- PyCharm
- VSCode
- Glueviz
- Orange 3 App
- RStudio
- Anaconda Prompt (Windows only)
- Anaconda PowerShell (Windows only)

Google colab :

Google Colab ou Colaboratory est un service cloud, offert par Google (gratuit), basé sur Jupyter Notebook et destiné à la formation et à la recherche dans l'apprentissage automatique. Cette plateforme permet d'entraîner des modèles de Machine Learning directement dans le cloud. Sans donc avoir besoin d'installer quoi que ce soit sur notre ordinateur à l'exception d'un navigateur.

Est un service hébergé de notebooks Jupyter qui ne nécessite aucune configuration et permet d'accéder gratuitement à des ressources informatiques, dont des GPU [53].

Colab permet :

- D'améliorer vos compétences de codage en langage de programmation Python.
- De développer des applications en Deep Learning en utilisant des bibliothèques Python populaires telles que Keras, TensorFlow, PyTorch et OpenCV.
- D'utiliser un environnement de développement (Jupyter Notebook) qui ne nécessite aucune configuration.
- Développement et entraînement de réseaux de neurones
- Expérimentation avec les TP

Principales fonctionnalités :

- Plusieurs types de GPU
- Quels que soient vos besoins en performances et votre budget, les GPU NVIDIA K80, P100, P4, T4, V100 et A100 offrent une variété d'options de calcul adaptées à votre charge de travail.
- Des performances flexibles
- Profitez d'un parfait équilibre entre le processeur, la mémoire, le disque hautes performances et jusqu'à 12GPU par instance en fonction de votre charge de travail

individuelle. En outre, avec la facturation à la seconde, vous ne payez que ce que vous utilisez.

- Tous les avantages de Google Cloud
- Exécutez vos charges de travail GPU sur Google Cloud Platform afin de profiter des technologies de stockage, de mise en réseau et d'analyse de données les plus performantes du secteur.

Matlab :

Matlab est un logiciel de calcul numérique commercialisé par la société MathWorks¹. Il a été initialement développé à la fin des années 70 par Cleve Moler, professeur de mathématique à l'université du Nouveau-Mexique puis à Stanford, pour permettre aux étudiants de travailler à partir d'un outil de programmation de haut niveau et sans apprendre le Fortran ou le C. Matlab signifie Matrix laboratory.

Il est un langage pour le calcul scientifique, l'analyse de données, leur visualisation, le développement d'algorithmes. Son interface propose, d'une part, une fenêtre interactive type console pour l'exécution de commandes, et d'autre part, un environnement de développement intégré (IDE) pour la programmation d'applications [54].

Matlab trouve ses applications dans de nombreuses disciplines. Il constitue un outil numérique puissant pour la modélisation de systèmes physiques, la simulation de modèles mathématiques, la conception et la validation (tests en simulation et expérimentation) d'applications. Le logiciel de base peut être complété par de multiples toolboxes, c'est-à-dire des boîtes à outils. Celles-ci sont des bibliothèques de fonctions dédiées à des domaines particuliers. Nous pouvons citer par exemple : l'Automatique, le traitement du signal, l'analyse statistique, l'optimisation...

4.2.3 Bibliothèques utilisées dans l'implémentation :

Tensorflow :

TensorFlow est une librairie logicielle, open source depuis 2015, de calcul numérique développée par Google. La particularité de TensorFlow est son utilisation des graphes de flux de données (data flow graphs).

TensorFlow a été développé par des chercheurs et des ingénieurs de Google afin de mener à bien des projets de recherche en apprentissage automatisé (machine learning) ainsi que sur le sujet des réseaux de neurones profonds (deep neural networks). Le système est néanmoins assez généraliste pour être applicable dans une large gamme de domaines d'application. TensorFlow permet notamment d'entraîner des modèles nécessitant de gros volumes de données (par exemple des banques d'images) de façon optimale.

L'architecture flexible de TensorFlow permet de déployer le calcul sur un ou plusieurs CPU/GPU sur un ordinateur personnel, un serveur *etc.* sans avoir à réécrire de code.

Google a également développé des Tensor Processing Units (ou TPU) construits spécialement pour l'apprentissage automatisé et pour être utilisés avec TensorFlow. Les TPU sont destinés à utiliser et tester des modèles plutôt qu'à les entraîner. Depuis Février 2018, les TPU sont disponibles sur la version beta de Google Cloud Platform[59].

Pourquoi TensorFlow ? :

- Basé sur python
- Supporté par Google
- Le plus important dans apprentissage profond (Deep Learning)
- Support toutes les opérations dans machine Learning
- Facile à utilisation et très efficace
- Support du multi GPU (CPU, GPU, TPU)

Keras :

Keras est une API de réseaux de neurones de haut niveau, écrite en Python et ineffaçable avec TensorFlow, CNTK et Theano. Elle a été développée avec pour objectif de permettre des expérimentations rapides. Être capable d'aller de l'idée au résultat avec le plus faible délai possible étant la clef d'une recherche efficace [58].

Nous vous recommandons Keras si vous êtes à la recherche d'une librairie qui :

- Permet le prototypage rapide et facile (de par sa convivialité, sa modularité et son extensibilité).
- Supporte à la fois les réseaux convolutifs et les réseaux récurrents ainsi que la combinaison des deux.
- Fonctionne de façon transparente sur CPU et GPU.

Les moteurs supportés sont :

- TensorFlow
- MicrosoftCNTK
- Theano
- Amazon travaille actuellement à l'intégration de MXNet avec Keras.

Vos modèles Keras peuvent être entraînés sur de nombreuses plateformes :

- GPUs NVIDIA
- TPUs GOOGLE, avec le moteur TensorFlow et Google Cloud
- Les GPU supportant OpenCL, tels que ceux d'Amazon, avec le moteur Keras PlaidML

PyTorch :

PyTorch est une bibliothèque d'apprentissage automatique pour Python utilisée principalement pour le traitement du langage naturel. Le logiciel open source a été développé par les équipes d'intelligence artificielle de Facebook Inc. en 2016. PyTorch offre deux fonctionnalités importantes, notamment le calcul tensoriel, ainsi que des réseaux de neurones profonds fonctionnels

PyTorch utilise un module Autograd pour calculer la différenciation automatique. Bref, un enregistreur détaille les opérations effectuées puis les rejoue pour synthétiser les gradients. Cela

permet de gagner du temps dans le développement des réseaux de neurones car la différenciation des données est effectuée rapidement lors de la passe avant. Le package optim de PyTorch permet à un utilisateur de définir un optimiseur qui mettra à jour les poids automatiquement. Cependant, lorsque les utilisateurs souhaitent créer leur propre modèle personnalisé, ils peuvent tirer parti du module nn.module de PyTorch. Compte tenu des différents modules, PyTorch vous permet d'implémenter différents types de couches telles que des couches convolutives, des couches récurrentes et des couches linéaires, entre autres [60].

Theano :

Theano est une bibliothèque Python qui vous permet de définir, d'optimiser et d'évaluer efficacement des expressions mathématiques impliquant des tableaux multidimensionnels. Il est construit sur NumPy. Caractéristiques de Théano [61] :

- Intégration étroite avec NumPy : une interface similaire à celle de NumPy. numpy.ndarrays sont également utilisés en interne dans les fonctions compilées par Theano.
- Utilisation transparente d'un GPU : effectuez des calculs gourmands en données jusqu'à 140 fois plus rapidement que sur un CPU (prise en charge de float32 uniquement).
- Différenciation symbolique efficace : Theano peut calculer des dérivées pour des fonctions d'une ou plusieurs entrées.
- Optimisations de vitesse et de stabilité : évitez les bogues désagréables lors du calcul d'expressions telles que $\log(1 + \exp(x))$ pour les grandes valeurs de x .
- Tests unitaires étendus et auto-vérification : comprend des outils pour détecter et diagnostiquer les bogues et/ou les problèmes potentiels

4.3 Pansharpening avec la colorisation guidée (PANCOLORGAN):

Pour remédier au inconvénient des approches des CNN, une nouvelle méthode de pansharpening qui préserve fidèlement les détails spatiaux donnés par l'entrée image panchromatique au l'inférence.

Ceci est atteint en concevant une procédure d'apprentissage auto-supervisé basée sur la tâche de colorisation, qui exige au modèle pendant la phase d'apprentissage, de fournir l'image multispectral au niveaux de gris, dont les détails spatiaux s'accordent parfaitement avec ceux de l'image multispectrale originale.

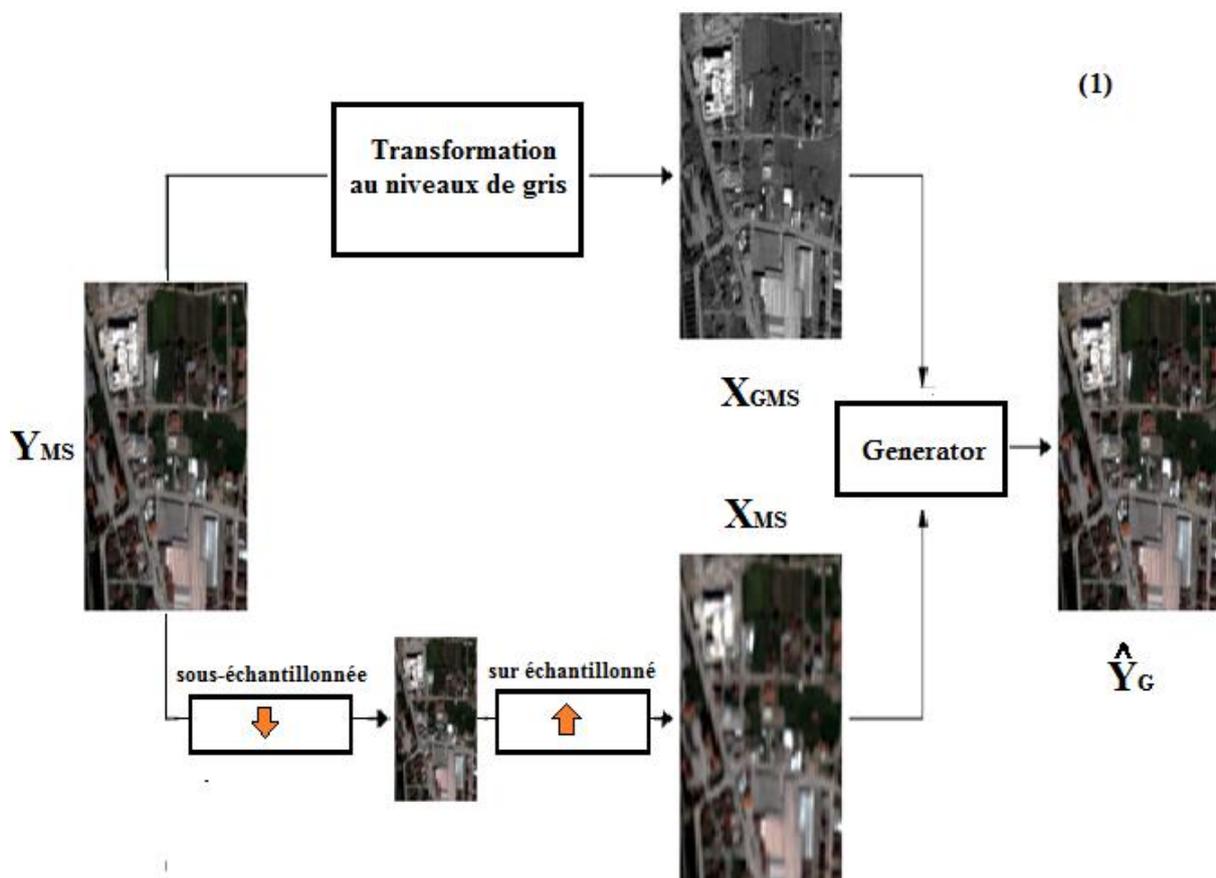
4.3.1 La méthode PANCOLORGAN

La méthode PanColorGAN apprend à générer une image multispectrale originale en prenant sa résolution réduite d'image multispectrale et l'image multispectrale en niveaux de gris correspondante comme entrée, ce qui est interprété comme une colorisation.

Nous pouvons également interpréter cela d'une manière que la modèle apprend à séparer les composantes spectrales et spatiales de l'image multispectrale pendant l'entraînement. Ensuite, dans l'étape des tests, le modèle fournit l'image panchromatique correspondante au lieu de l'image multispectrale au niveau de gris [50].

Le modèle d'apprentissage PanColorGAN Pansharpning est illustré à la Fig (4.1) :

- Reconstruction Loss(L1)
- la sortie colorisée l'entrée \hat{Y}_G et l'entrée Y_{MS} ,
- perte qui évalue la qualité de génération de \hat{Y}_G Les G générés à partir de X_{GMS} et X_{MS} sont utilisés pour entraîner le PanColorGAN.



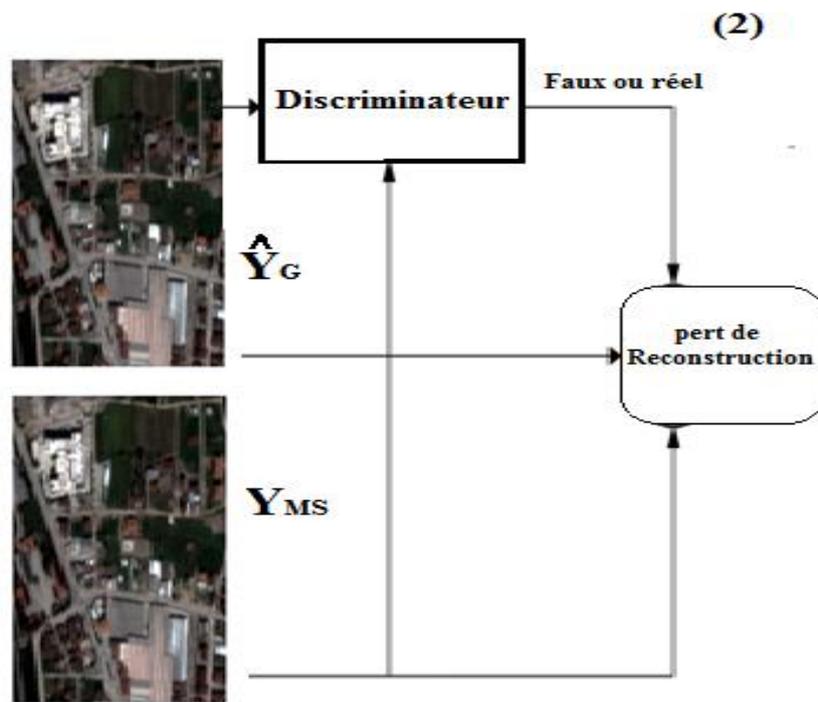


Figure 4.1: Schéma de formation proposé pour le modèle PanColorGAN[50].

Implémenté le PanColorGAN avec un ratio de down/up sampling fixe. Supposer que l'image multispectrale d'entrée YMS est d'abord sous-échantillonnée par $k = 4$ puis sur échantillonné par $k = 4$ pour obtenir XMS.

YMS est également transformé en niveaux de gris en prenant une moyenne de canaux pour construire un XGMS d'entrée en niveaux de gris. Plus tard, XGMS et XMS sont fournis en entrée du réseau de générateurs G et $\hat{Y}_G = G(XGMS, XMS)$ est obtenu en sortie.

La méthode PanColorGAN, comme dans les GAN traditionnels, est un réseau discriminatoire supplémentaire D est également construit pour fournir une Perte contradictoire, qui est calculée pour le générateur \hat{Y}_G , pour augmenter la capacité de représentation du réseau de générateurs en fournissant un retour d'information sur la qualité ou la crédibilité de sa production générée.

Les GAN appartiennent à la classe des réseaux génératifs qui apprennent synthétiser les images avec une distribution cible par la concurrence de typiquement de deux réseaux, dont l'un est le générateur et l'autre est le discriminatoire.

Le générateur G apprend à transformer un bruit aléatoire à la distribution de l'image cible.

Discriminateur D vise à classer correctement la sortie de G avec un « généré » étiquette versus étiquette « réelle », le « réel » renvoie à une étiquette des données d'entraînement. D effectue également la même opération sur les images générées par le modèle. C'est la base de la perte accusatoire dans le GAN, qui est utilisé dans la mise à jour à la fois G et D [50].

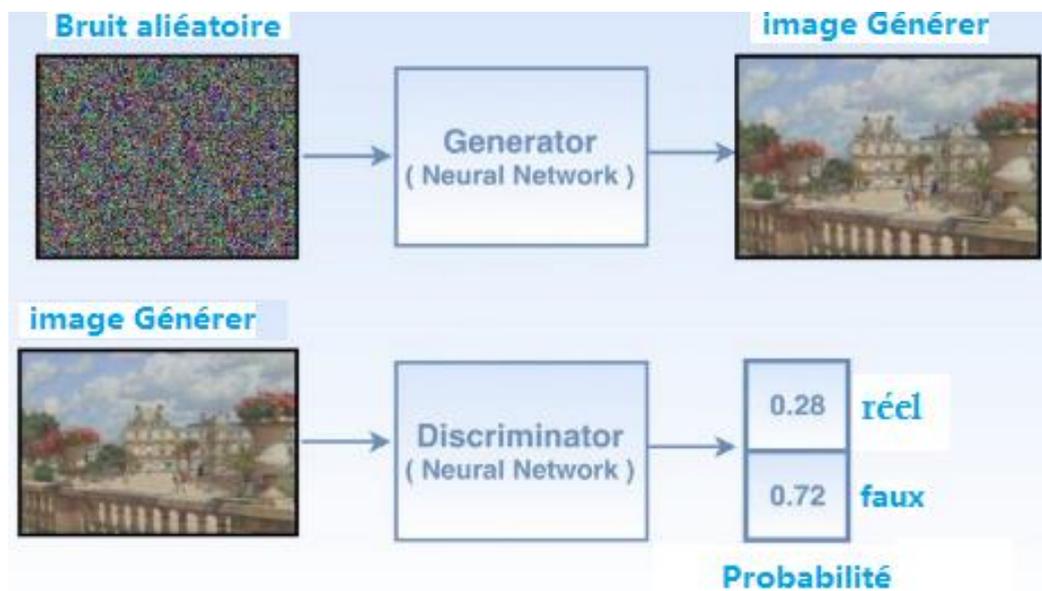


Figure 3.21 : Principe de GAN pour la classification on réel ou faux

- **Adversarial Loss (pertes) (GAN - moyenne relativist GAN (RaGAN))**

D effectue également la même opération sur les images générées par le modèle. C'est la base de la perte accusatoire dans le GAN, qui est utilisé dans mettre à jour à la fois G et D

$$L_{GAN}^D = -\mathbb{E}_{x_r \sim \mathbb{P}} [\log(D(x_r))] - \mathbb{E}_{x_f \sim \mathbb{Q}} [\log(1 - D(x_f))] \quad (1)$$

$$L_{GAN}^G = -\mathbb{E}_{x_f \sim \mathbb{Q}} [\log(D(x_f))]. \quad (2)$$

Ici, $D(x) = \sigma(C(x))$, où $C(x)$ fait référence à la sortie du réseau discriminateur après quoi l'activation la fonction est appliqué,

(x_r) fait référence aux données réelles qui sont échantillonnées de l'ensemble de données

(x_f) fait référence aux données générées avec générateur G.

Dautre part : Un cadre GAN plus récent, RaGAN, utilise à la place les pertes suivantes :

$$L_{RaGN} = -\mathbb{E}_{x_1} [\log(D(x_1, x_2))] - \mathbb{E}_{x_2} [\log(1 - D(x_2, x_1))]$$

$$L_{RaGN}^D = L_{RaGN}(X_f, X_r),$$

$$L_{RaGN}^G = L_{RaGN}(X_r, X_f),$$

Où $D(x_1, x_2) \triangleq \sigma(C(x_1) - E[C(x_2)])$.

Alors qu'un discriminateur du GAN prédit le réalisme d'une image, le discriminateur relativiste évalue la réalité du réel et de fausses images relativement. Comme il a été démontré que l'utilisation la perte RaGAN fournit des détails plus nets tout en ayant plus formation stable.

Le perte RaGAN est également intégrée dans PANCOLORGAN [50].

- **PANColorization GAN (PanColorGAN) Model**

Les détails de l'architecture de ses réseaux générateurs et discriminateurs sont représentés. Deux modes existent pour le réseau de générateurs : dans le phase de formation, XGMS est fourni avec XMS pour générer \hat{Y}^G . Dans la phase de test, XPAN est fourni avec XMS pour générer \hat{Y}^p .

De plus, pendant la phase d'apprentissage, le réseau discriminateur obtient deux types de modèles différents.

Un modèle réel consiste en un ensemble concaténé de XGMS, XMS et YMS, et un faux modèle consiste en un ensemble concaténé de XGMS, XMS et \hat{Y}^G comme indiqué en figure 4.2, qui illustre les détails de l'architecture PanColorGAN. Son générateur G est une version modifiée et il a des raccourcis de concaténation à travers les couches

afin de fournir une optimisation améliorée dans les termes de réduction du problème de perte. G a quatre parties principales qui servent comme des objectifs spécifiques :

- Les détails spatiaux d'extraction ;
- L'injection de couleur ;
- La transformation des caractéristiques ;
- La synthèse d'image pansharpened.

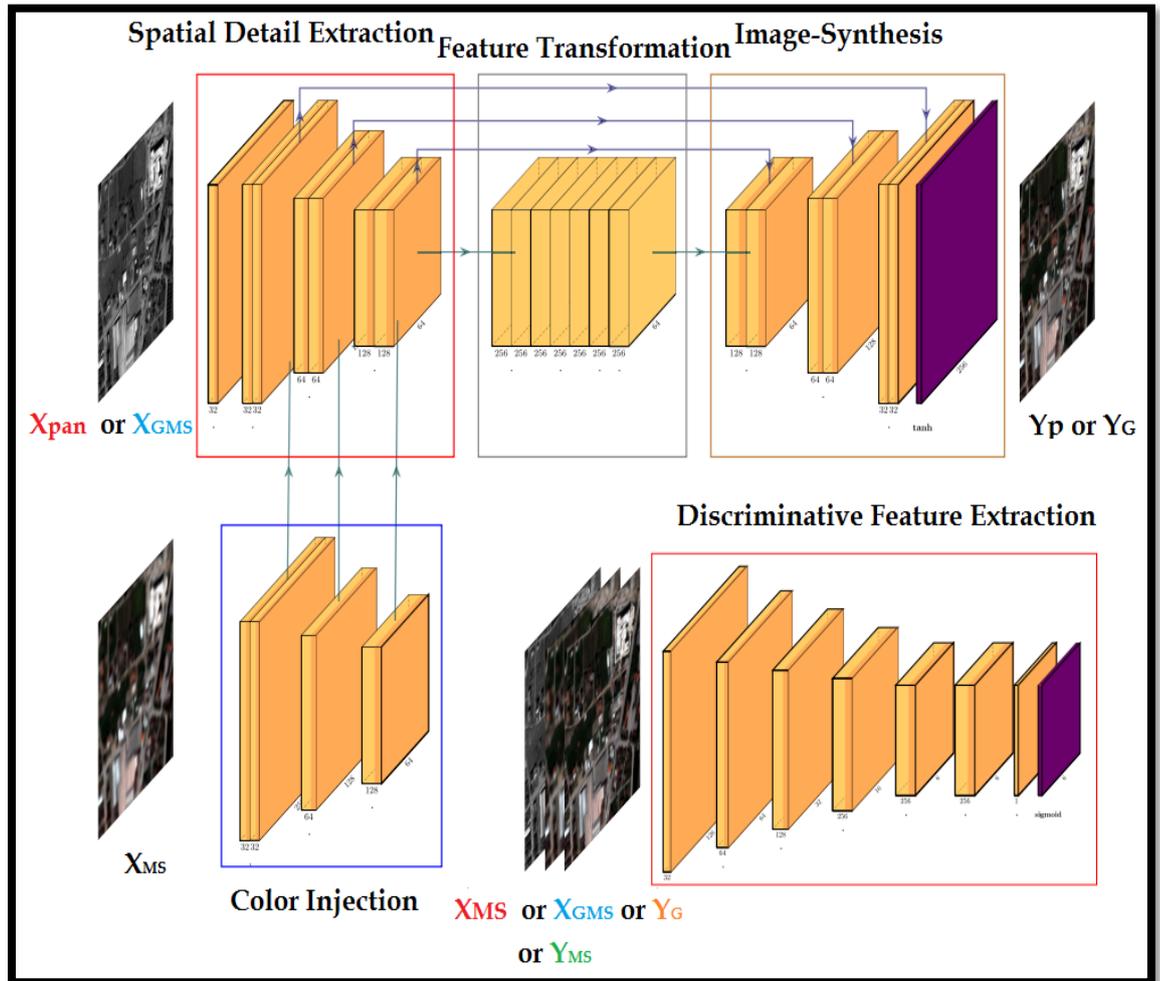


Figure 4.2: Schéma de formation proposé pour le modèle PanColorGAN [50].

L'extraction des détails spatiaux partie prend une image en niveaux de gris (XGMS) comme entrée et applique (3×3) convolutions tout en obtenant des caractéristiques de couleur à partir de la partie d'injection de couleur.

La partie injection de couleur est une architecture entièrement convolutive qui applique quatre fois une convolution (3×3) et injecte des caractéristiques de couleur extraites du multispectral (XMS), aux couches d'extraction de détails spatiaux du réseau après chaque opération de convolution sauf la première.

Il y a un bloc résiduel au milieu du réseau qui transforme les caractéristiques spatiales et spectrales concaténées et les prépare à une synthèse de l'image affinée.

Enfin, le réseau augmente lentement la hauteur et la largeur, et diminue la profondeur des caractéristiques en appliquant un suréchantillonnage et convolutions (3×3) tout en obtenant des caractéristiques du détail partie extraction, Grouper la normalisation et l'activation LeakyReLU sont insérées après chaque opération de convolution.

Après avoir obtenu des fonctionnalités en tant que même dimension que l'image multispectrale, l'activation tanh est appliqué pour mapper les intensités de l'image à l'intervalle $[-1, 1]$.

L'utilisation de tanh permet un entraînement plus rapide et plus stable du GAN.

Cela produit la sortie \hat{Y}_G du générateur réseau.

Le discriminateur D de PanColorGAN a un patchGAN conditionnel l'architecture, qui opère sur les patches d'image, et donne une sortie pour chaque champ récepteur qu'il voit. D'où, la sortie indique si ces champs récepteurs vus par D avoir l'air réaliste ou non, ensuite ces sorties sont agrégées dans une perte patchGAN pour le réseau discriminateur D. la perte de reconstruction pour le générateur n'est pas calculée sur patches, mais calculé au pixel près pour l'ensemble de l'image.

Dans un cadre GAN conditionnel, classiquement, le réseau D prend la sortie du générateur ou de la vérité terrain avec les contributions.

Le pansharpening peut être considéré dans le cadre de l'idée de traduction image-à-image, avec les réseaux contradictoires conditionnels, pour le discriminateur réseau D, les entrées du générateur sont prises comme conditions dans sa décision de « réel » ou de « faux », pour cette raison , les faux modèles sont constitués d'images en niveaux de gris XGMS, images multispectrales réduites XMS, et sorties de réseau \hat{Y}_G , les modelé réels comprennent XGMS, XMS et le images multispectrales originales YMS.

Le réseau de générateurs synthétise des images à partir des variables latentes échantillonnées au hasard et le discriminateur ne reçoit que les images générées et images réelles à son entrée, fournir tous les intrants connexes avec les images générées et réelles garantissent que le discriminateur réseau comprend les relations visuelles entre l'entrée et images de sortie.

D applique circonvolutions (4×4) par 2 pas et appliquant cinq fois, et réduit la hauteur et la largeur tout en augmentant la profondeur.

Ensuite, une dernière convolution réduit la profondeur à 1. La normalisation des modelés et l'activation LeakyReLU sont exécutées après chaque couche de convolution. L'opération sigmoïde est appliquée dans afin de réduire l'intervalle à $[0, 1]$. Ainsi, à la sortie de D, les indicateurs de la réalité des champs récepteurs dans l'image donnée sont obtenus.

Le modèle PanColorGAN utilise les pertes suivantes pour l'apprentissage les poids des réseaux G et D :

$$\begin{aligned} L_D &= L_{RaGN}(\mathbf{Y}_{MS}, \hat{\mathbf{Y}}_G), \\ L_G &= L_{Rec} + \alpha L_{RaGN}(\hat{\mathbf{Y}}_G, \mathbf{Y}_{MS}), \\ L_{Rec} &= \|\mathbf{Y}_{MS} - \hat{\mathbf{Y}}_G\| \end{aligned}$$

La perte de reconstruction est conçue comme la moyenne absolue erreur (perte L1), alors que la perte contradictoire est conçue comme la perte RaGAN. Alors que la perte de reconstruction augmente la similitude au niveau des pixels entre les images générées et les images multispectrales correspondantes, la perte contradictoire rapproche la distribution des images générées des images multispectrales et offre une netteté dans les détails. Dans PanColorGAN, LRec mesure la distance entre Y_{MS} et \hat{Y}_G plutôt que \hat{Y}_p , car ce dernier dirigerait la formation réseau pour biaiser la distribution spatiale du pansharpened image vers le domaine de l'image multispectrale,

- **Sous-échantillonnage aléatoire d'images multispectrales (Random Downsampling of Multispectral Images)**

La formation du pansharpening réseau avec l'échelle de sous-échantillonnage 4×4 réduit la capacité de représentation du modèle, en particulier pour le scénario de pansharpening pleine résolution.

Par conséquent, nous substituons l'opération de sous-échantillonnage $4 \times$ avec un sous-échantillonnage aléatoire opération dans un modèle amélioré, que nous appelons PanColorGAN + RD (Random Downsampling).

Dans la phase de test, aléatoire le sous-échantillonnage n'est pas utilisé, cette modification offre un moyen à PanColorGAN pour améliorer son apprentissage comme suit :

Quand seulement le sous-échantillonnage $4 \times$ est utilisé dans la phase d'apprentissage, le réseau apprend à interpoler l'image panchromatique réduite et l'image multispectrale réduites à l'échelle donnée.

Et ne pas apprendre correctement la tâche de colorisation.

- **Inférence via les modèles PanColorGAN proposés**

Une fois la phase d'entraînement terminée, pendant la phase de test à résolution réduite, l'image Y_{PAN} d'origine est réduite à la même taille que l'image multispectrale pour obtenir X_{PAN} .

Les images X_{PAN} et X_{MS} sont fournies aux Réseau générateur PanColorGAN (G), et $\hat{Y} = G(X_{PAN}, X_{MS})$ est obtenu en sortie, pour l'inférence à résolution réduite.

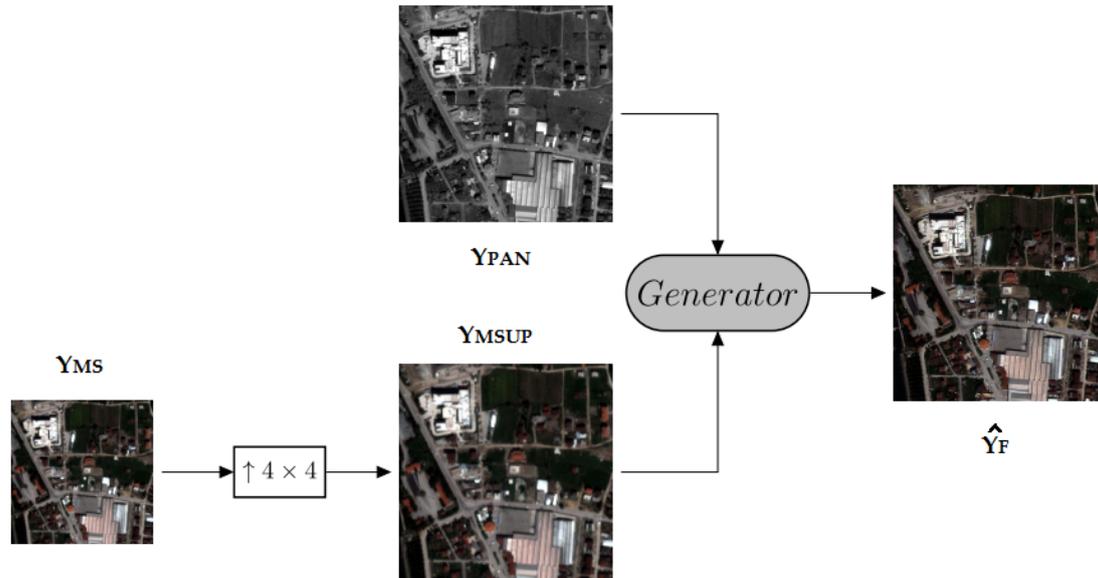


Figure 4.3: Schéma d'inférence (test) à pleine résolution

La figure 4.3 illustre comment exécuter la pleine résolution, c'est-à-dire le scénario réel dans le pansharpening. YPAN d'origine et $4\times$ une version suréchantillonnée de YMS est fournie au réseau de générateur PanColorGAN formé G, et $\hat{Y}_G = G(YPAN, YMSUP)$ est obtenu en tant que sortie d'image pansharpened pleine résolution

4.4 Expériences et résultats

Cette partie contient le code nécessaire pour exécuter la méthode de pansharpening (PanColorGAN) basée sur l'apprentissage en profondeur pour la fusion de panchromatique et des images multispectrales dans les applications de télédétection.

Nous avons implémenté PanColorGAN dans Pytorch pour tester sur un GPU.

Nous avons utilisé l'optimiseur Adam avec un taux d'apprentissage initial 0,0002, β_1 égal à 0,5 et β_2 égal à 0,999. Nous n'avons pas utilisé de poids car il diminue les performances de la synthèse d'images. La perte de poids antagoniste α a été fixée à 0,005 dans :

$$L_G = L_{Rec} + \alpha L_{RaGN}(\hat{Y}_G, YMS)$$

Une activation leakyReLU avec une pente de 0,2 est utilisée dans toutes les activations couches, lors de la formation du modèle PanColorGAN + RD, pour chaque image d'un modèle donné, un sous-échantillonnage aléatoire la taille est échantillonnée uniformément sous forme d'entier intervalle, l'échelle de suréchantillonnage est alors automatiquement réglée sur redimensionner l'image sous-échantillonnée à 256. Les deux suréchantillonnage et le sous-échantillonnage sont effectués avec une interpolation bicubique schème

4.4.1 Ensemble de données propose pour le pansharpening

Pour faire progresser l'état de l'art en matière de pansharpening, un ensemble de données à grande échelle, il se compose de 2 270 paires de HR Pan et LR images MS. Pour la construction de l'ensemble de données, un repérage géométrique a été effectué et les images HR Pan et LR MS ont été recadrées pour une utilisation pratique. De plus, toutes les images de l'ensemble de données ont été réorganisées et regroupées en fonction de types de satellites de télédétection et caractéristiques de surface thématiques, respectivement. Des exemples de l'ensemble de données proposé sont présentés dans La figure 4.4 et les détails sont affichés dans le tableau 4-1[22].

ENSEMBLE DE DONNÉES GROUPEES PAR SATELLITES DE TÉLÉDÉTECTION

L'ensemble de données proposé a été regroupé par types de satellites de télédétection, consistant en 200 paires d'images Ikonos Pan/MS, 500 paires d'images QuickBird Pan/MS, 410 paires de GF-1 Images Pan/MS, 500 paires d'images WorldView-4 Pan/MS, 500 paires d'images Pan/MS WorldView-2 et 160 paires d'images Images panoramiques/MS WorldView-3. Les images Pan ont une dimension de (1024 * 1 024), et les images MS ont une dimension spatiale de (256 * 256). Dans la dimension spectrale, le Images MS d'Ikonos, QuickBird, GF-1 et WorldView-4 ont quatre bandes, et le WorldView-2 et le WorldView-3 MS les images en ont 8 bandes.

L'ensemble de données proposé a été regroupé par caractéristiques de surface matic typiques. Ils incluent la thématique urbaine ensemble de données avec 510 paires d'images Pan/MS, ensemble de données thématiques sur la végétation verte avec 258 paires et ensemble de données thématiques sur les scénarios d'eau avec 318 paires. De plus, un ensemble de données non étiquetées à grande échelle contenant 1 184 paires d'images Pan/MS avec des caractéristiques de surface mixtes a été construit.

1) Ensemble de données urbaines : les images de cet ensemble de données sont principalement situées dans les régions urbaines, et les bâtiments sont au centre de l'attention. Dans De plus, des arbres, des routes, etc. peuvent exister, ce qui est inévitable pour les images de télédétection des zones urbaines. Les Les images HR Pan ont une dimension de 1024 * 1 024 * 1, et les images LR MS ont une dimension de 256 * 256 * 4 (Ikonos, QuickBird, GF-1 et WorldView-4) et 256 * 256 * 8 (WorldView-2 et WorldView-3).

2) Ensemble de données sur la végétation verte : les images de cet ensemble de données prennent végétation verte comme partie principale. Ils sont généralement situés dans les banlieues, les zones montagneuses et rurales emplacements, tels que les terres

agricoles. Les images HR Pan ont une dimension de 1024 * 1 024 * 1, et les images LR MS ont une dimension de 256 # 256 # 4 (Ikonos, QuickBird, GF-1 et WorldView4) et 256 * 256 *8 (World View-2 et World View-3)

3) Ensemble de données sur les scénarios d'eau : cet ensemble de données est principalement construit avec les caractéristiques de la surface de l'eau comme corps principal. Pour ces images, bateaux et autres éléments au sol peuvent exister. Comme les ensembles de données sur la végétation urbaine et verte, le HR Les images panoramiques ont une dimension de 1024*1024*1 et les images LR MS ont une dimension de 256*256* 4 (Ikonos, QuickBird, GF-1 et WorldView-4) et 256 * 256 *8 (WorldView-2 et WorldView-3).

4) Ensemble de données Sans étiquette avec des caractéristiques mixtes : les images de cet ensemble de données se composent de caractéristiques de surface différentes, dont deux ou plus différents types sont généralement les parties principales. Par rapport aux trois ensembles de données thématiques précédents, les images de l'ensemble de données non identifié ont des caractéristiques au sol plus complexes, telles que que les attributs mixtes de vastes zones de végétation et l'eau ainsi que les endroits qui ont des bâtiments, de l'eau et sol. Les images HR Pan ont une dimension de 1024*1024*1, et les images LR MS ont une dimension de 256 *256*4 (Ikonos, QuickBird, GF-1 et WorldView-4) et 256*256*8 (WorldView-2 et WorldView-3).

Les détails sont indiqués dans le tableau suivant :

Satellite		Dimension spatiale	Dimension spectrale	Dimension Taille	Nombre d'images de scènes thématiques				
					Urbain	Vert végétation	Scénario eau	Sans étiquette caractéristiques mixtes	Volume de données
IKONOS	Pan	1 m	1 band	1024×1024	60	8	6	126	200
	MS	4 m	4 bands	256×256×4					
QuickBird	Pan	0.61 m	1 band	1024×1024	150	20	42	288	500
	MS	2.44 m	4 bands	256×256×4					
Gaofen-1	Pan	2 m	1 band	1024×1024	5	90	10	305	410
	MS	8 m	4 bands	256×256×4					
WorldView-4	Pan	0.31 m	1 band	1024×1024	90	85	95	230	500
	MS	1.24 m	4 bands	256×256×4					
WorldView-2	Pan	0.5 m	1 band	1024×1024	150	35	145	170	500
	MS	2 m	8 bands	256×256×8					
WorldView-3	Pan	0.31 m	1 band	1024×1024	55	20	20	65	160
	MS	1.24 m	8 bands	256×256×8					

Tableau 4-1 : Les Détails de base des données.

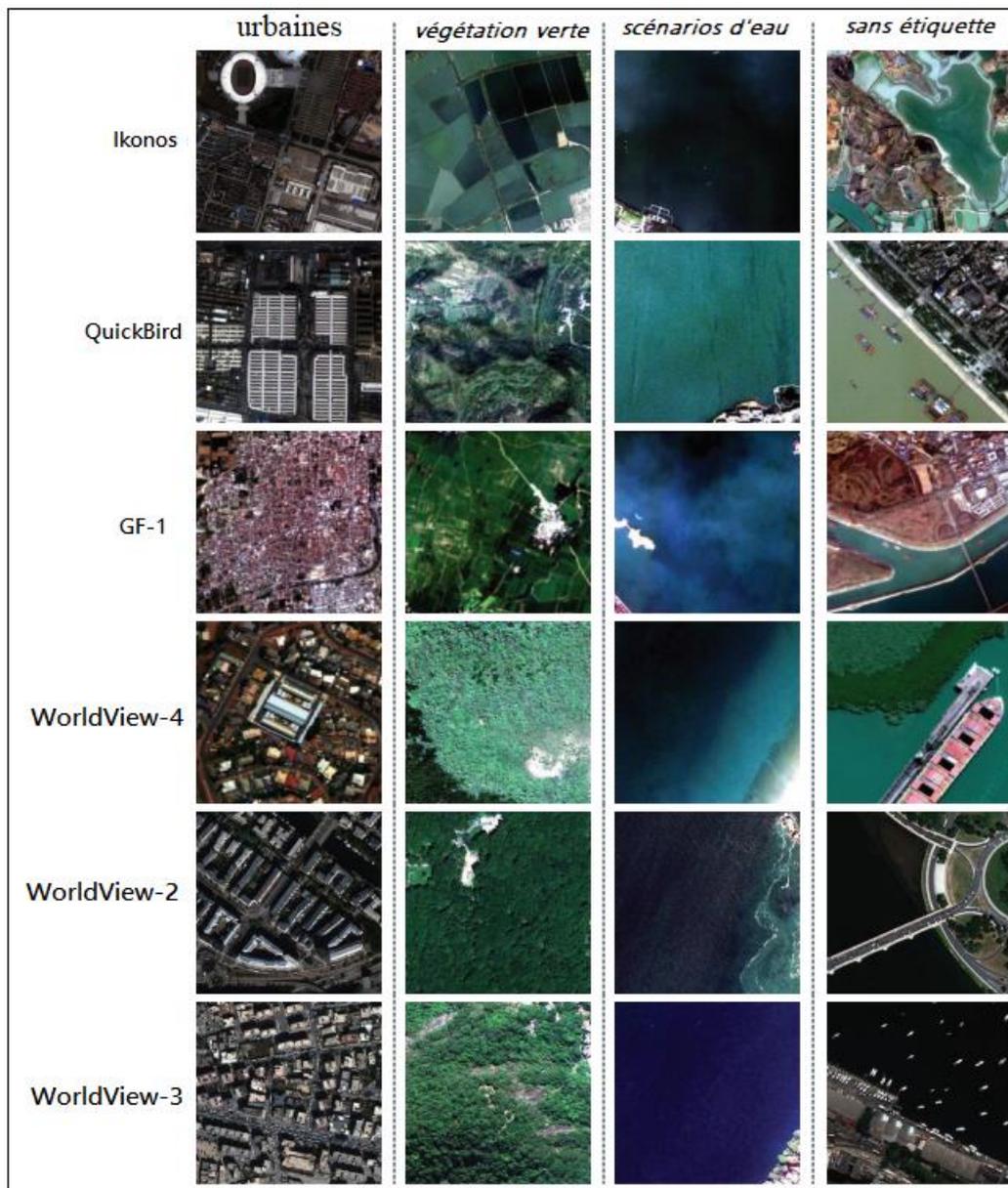


Figure 4.4: Échantillons de l'ensemble de données proposé

4.4.2 Code d'application

ETAPE 1 : télécharger des images PAN et MS...annex1

ETAPE 2 : afficher des images annex1

ETAPE 3 : Prétraitement des données annex1

ETAPE 4 : Préparation du réseau annex1

ETAPE 5 : fusion et afficher résultat annex1

4.4.3 Résultat d'exécution de code :

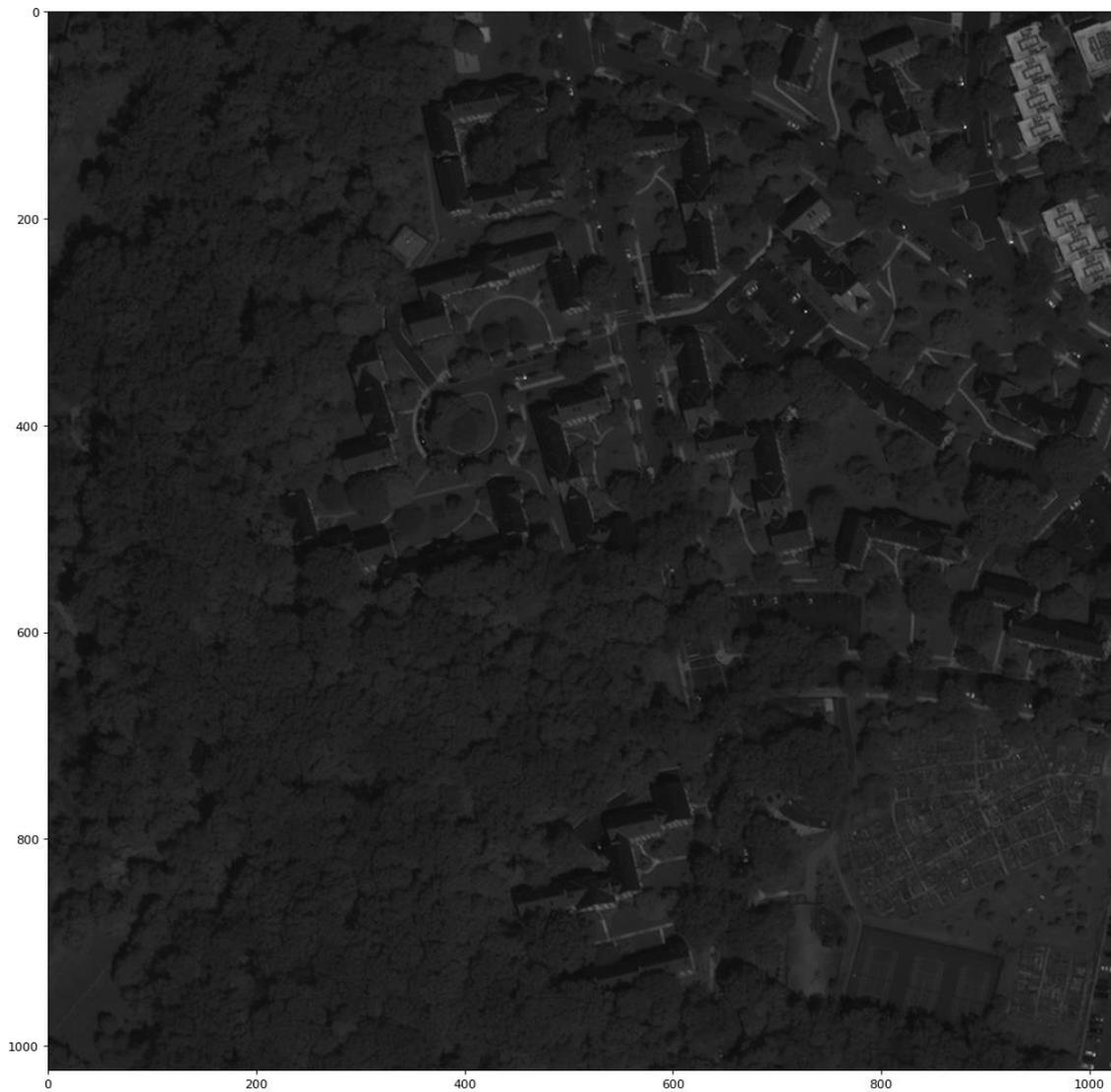


Figure 4.5: image PAN (image Sans étiquette avec des caractéristiques mixtes 1024*1024)



Figure 4.6: image MS (image Sans étiquette avec des caractéristiques mixtes 256*256*4)

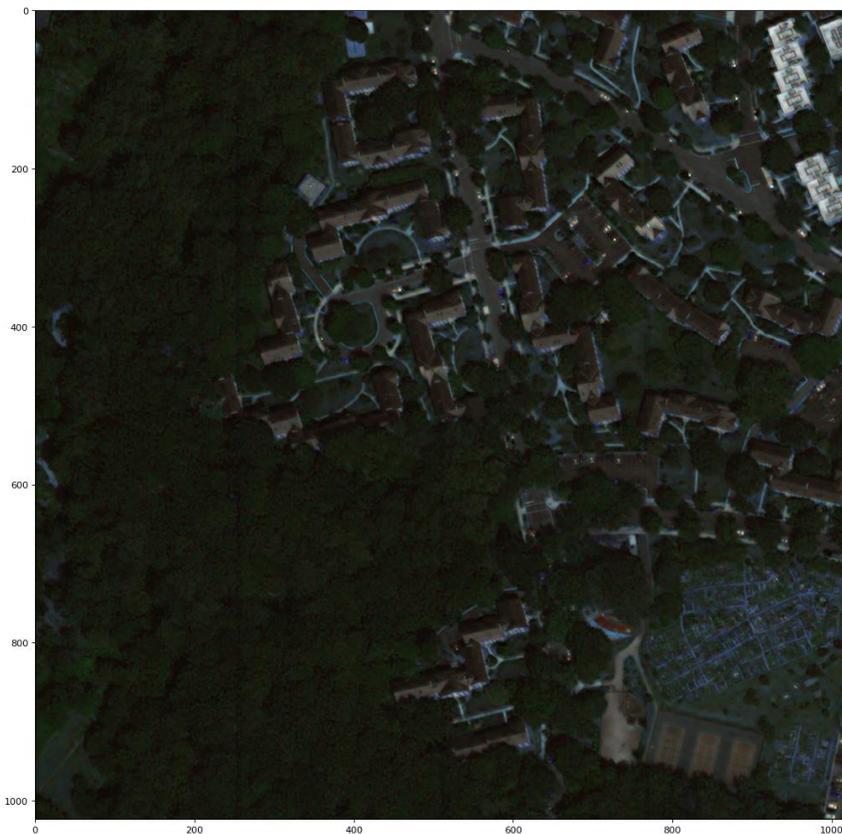


Figure 4.7: Résultat de la fusion PAN COLOR GAN (image MS haute résolution)

4.5 Évaluation de la qualité des méthodes de fusion

Évaluation de la qualité des méthodes et des données de pansharpening produits ont été largement étudiés dans le passé deux décennies. Après une première génération de méthodes empiriques correspondant à une évaluation subjective, une définition objective de qualité [44].

Les méthodes de fusion doivent satisfaire à deux propriétés principales :

La cohérence et la synthèse. La cohérence indique que l'image affinée, dégradée à la résolution MS d'origine, doit être aussi similaire que possible à l'image MS d'origine (spectrale).

La synthèse nécessite que l'image affinée doit être similaire à celle acquis par un capteur MS de plus grande résolution.

Vérification de la propriété de synthèse est possible en effectuant la fusion avec ensembles de données MS et PAN dégradés à des résolutions spatiales inférieurs à ceux d'origine.

Alors que la synthèse est largement adoptée en suivant deux Approche

➤ REDUCED RESOLUTION ASSESSMENT (ÉVALUATION DE LA RÉOLUTION RÉDUITE) (RR) :

L'évaluation RR mesure la similitude de la fusion produit à une référence idéale, c'est-à-dire l'image MS d'origine.

Cela est possible en dégradant les résolutions des deux MS d'origine et le PAN d'origine et en effectuant la fusion à partir de ces données dégradées, le choix du filtre en général défini pour assurer la propriété de cohérence du processus de pansharpening .

Ainsi, il est simple que la réduction de la résolution de l'image MS doit être fait en exploitant des filtres spatiaux.

De plus, le filtre utilisé pour dégrader l'image PAN doit être conçue pour préserver les détails qui aurait été vu si l'image avait été acquise à le RR. Par conséquent.

Plus l'image affinée obtenue est similaire à l'image MS d'origine, plus la qualité mesurée est élevée. Un tel degré de similitude peut être facilement calculé par des index de score qui comparent deux images multicanaux [44].

SPECTRAL ANGLE MAPPER (SAM)

Le mappeur d'angle spectral (SAM) a été introduit à l'origine pour la discrimination des matériaux à partir de leurs spectres de réflectance.

Étant donné deux vecteurs spectraux, \mathbf{v} et $\hat{\mathbf{v}}$, les deux ayant N composants, dans lesquels $\mathbf{v} = [v_1, v_2, \dots, v_N]$ est le vecteur de pixel spectral de référence et $\hat{\mathbf{v}} = [\hat{v}_1, \hat{v}_2, \dots, \hat{v}_N]$ est le vecteur de pixel spectral de test, SAM désigne l'absolu valeur de l'angle spectral entre les deux vecteurs comme

$$SAM(\mathbf{v}, \hat{\mathbf{v}}) = \arccos\left(\frac{\langle \mathbf{v}, \hat{\mathbf{v}} \rangle}{\|\mathbf{v}\|_2 \cdot \|\hat{\mathbf{v}}\|_2}\right)$$

SAM est généralement exprimé en degrés et est égal à zéro si et seulement si le vecteur de test est spectralement identique au vecteur de référence ; c'est-à-dire que les deux vecteurs sont parallèles et peuvent ne diffèrent que par leurs modules. Une dissemblance spectrale globale, ou distorsion, l'indice est obtenu en faisant la moyenne sur la scène entière.

ERREUR RELATIVE GLOBALE ADIMENSIONNELLE DE SYNTHÈSE (ERGAS)

L'indice ERGAS, dont l'acronyme français signifie relative erreur globale sans dimension en synthèse, est un indice qui offre une indication globale de la distorsion vers la référence d'une image multi bande de test :

$$\text{ERGAS} = 100 \frac{d_h}{d_l} \sqrt{\frac{1}{N} \sum_{n=1}^N \left(\frac{\text{RMSE}(n)}{\mu(n)} \right)^2},$$

Où d_h/d_l est le rapport entre les tailles de pixels de PAN et MS, généralement 1/4 pour de nombreux capteurs utilisés pour la fusion panoramique $\mu(n)$ la moyenne (moyenne) de la nième bande de la référence ; et N est le nombre de bandes.

De faibles valeurs d'ERGAS indiquent une similitude élevée entre les données MS fusionnées et de référence.

Q2n :

Q2n est l'extension multi bande de l'image universelle indice de qualité (UIQI) et a été introduit pour l'évaluation de la qualité des images MS pansharpened, d'abord pour quatre bandes et plus tard étendu à 2n bandes. Chaque pixel d'une image avec N bandes spectrales est logé en un nombre hypercomplexe (HC) avec une partie réelle et N-1 parties imaginaires. Soit $z = z(m, n)$ et $z' = z'(m, n)$ la représentation HC des vecteurs spectraux de référence et de test au pixel (m, n). De manière analogue à UIQI, à savoir, $Q2^0 = (Q, Q2^n)$ peut être écrit comme le produit de trois termes :

$$Q2^n = \frac{|\sigma_{z\hat{z}}|}{\sigma_z \sigma_{\hat{z}}} \cdot \frac{2\sigma_z \sigma_{\hat{z}}}{\sigma_z^2 + \sigma_{\hat{z}}^2} \cdot \frac{2|\bar{z}||\bar{\hat{z}}|}{|\bar{z}|^2 + |\bar{\hat{z}}|^2},$$

➤ FULL-RESOLUTION ASSESSMENT (ÉVALUATION À RÉOLUTION COMPLÈTE) (FR)

L'évaluation FR en déduit la qualité de l'image pansharpened à l'échelle de l'image PAN sans recours à une seule image de référence, qui n'est pas disponible, pour résoudre le problème, de nouvelles mesures de distorsion ont été introduites ne mesurerait aucune distorsion si de telles données étaient disponibles [55].

Deux indices de score pour l'évaluation FR sont rapportés dans ce qui suit :

QUALITY WITH NO REFERENCE (QNR)

Le protocole QNR prend des mesures séparées, chacune normalisée en [0, 1] de cohérences spectrale et spatiale. La motivation à ce protocole est que les similitudes

entre les couples de bandes MS et chaque MS bande et PAN devraient être inchangés en moyenne, variant avec la résolution, c'est-à-dire avant et après la fusion.

La distorsion spectrale, D_μ , est calculée entre le Images LRMS et images MS fusionnées, deux ensembles des valeurs UIQI inter bandes sont calculées séparément à bas et hautes résolutions. Les différences des UIQI correspondants les valeurs aux deux échelles donnent la distorsion spectrale introduite par le processus de fusion.

La distorsion spatiale, D_S , combine les valeurs UIQI calculées entre chaque bande MS et l'image PAN dégradée à la résolution de MS et encore entre MS et FR PAN fusionnés. La différence absolue, en moyenne sur toutes les bandes, entre les valeurs UIQI correspondantes donne la distorsion spatiale D_S , l'alignement spatial entre les données MS d'origine et l'image PAN est d'une importance cruciale pour avoir une évaluation précise de cet indice de distorsion. Cette dernière question est considérée comme être résolu par le suréchantillonnage de l'image MS vers le PAN échelle. Ainsi, la mise en œuvre de cet indice peut être effectuée en considérant le problème à l'échelle PAN, c'est-à-dire en exploitant l'image MS suréchantillonnée au lieu de l'image originale. Image LRMS.

Les deux distorsions sont ensuite combinées pour fournir un indice de qualité, appelé QNR $\in [0,1]$ avec 1 étant la meilleure valeur atteignable :

$$QNR = (1 - D_\mu)^\alpha (1 - D_S)^\beta$$

Telle que: $A = \beta = 1$

HYBRID QUALITY WITH NO REFERENCE (HQNR)

Le HQNR emprunte l'indice de distorsion spatiale D_S de QNR et l'indice de distorsion spectrale, $D_\mu^k = 1 - Q2^n(\widehat{MSI} - \overline{MS})$ est l'image \widehat{MSI} pansharpened filtrée par MTF compte tenu d'un rapport de résolution égal à R et \overline{MS} est l'image MS d'origine interpolé à l'échelle PAN, R fois inférieur à la MS échelle. Encore une fois, cette implémentation est réalisée en exploitant l'image MS suréchantillonnée, c'est-à-dire travaillant à l'échelle PAN.

Il est à noter que D_μ^k suit strictement la propriété de cohérence et évite les inconvénients possibles de D_μ , qui calcule les indices de similarité à différentes résolutions. Les deux distorsions sont finalement combinées pour donner l'indice de qualité unique

$$QNR = (1 - D_\mu^k)^\alpha (1 - D_S)^\beta$$

Telle que : $A = \beta$

4.6 Evaluation et Comparaison de résultat

PARAMÈTRES D'EXPÉRIENCE

Pour vérifier efficacement les performances des méthodes de pansharpening, des expériences ont été réalisées en utilisant les données ensemble de différentes images satellites de télédétection,

Des expériences à résolution réduite et à pleine résolution ont été effectuées, l'original image MS a été considérée comme l'image de référence pour l'évaluation

Les indices d'évaluation quantitative populaires de la Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS), angle de mappage de spectre (SAM), indice Q2n Étaient employés.

Dans l'expérience en pleine résolution, en raison de l'absence d'image de référence, l'indice d'évaluation de la qualité sans référence populaire de la qualité sans référence (QNR) a été utilisé, D_m dans le QNR évalue la distorsion spectrale, et D_s dans le QNR évalue la distorsion spatiale.

Dans les expériences, consiste les diverses méthodes de pansharpening, comme indiqué dans le tableau 4-2, ont été évaluées.

METHODES	Expressions
CS	Component substitution
BDSD	Une approche basée sur la CS auto-adaptative, dépendante des données, présentant deux avantages : 1) la solution optimale de l'extraction et de l'injection de détails spatiaux basée sur l'erreur quadratique moyenne minimale 2) l'estimation des paramètres basée sur les régions locales
GS	Une méthode typique basée sur CS basée sur la transformation GS;
GSA	Une méthode GS améliorée en termes de calcul adaptatif du poids de combinaison de bandes spectrales basé sur un algorithme de régression linéaire
Brovvey	Une méthode basée sur le CS sous la compréhension générale.

PRACS	Une méthode basée sur le CS fondée sur le remplacement partiel de la composante d'intensité ; le Pan est simulé sur la base de la moyenne pondérée des bandes Pan et MS au lieu d'utiliser directement le Pan d'origine pour remplacer la composante d'intensité
PCA	Une méthode typique basée sur CS fondée sur la transformation PCA.
MRA	Multi résolutions analysais
ATWT	Une méthode basée sur l'ARM basée sur l'ATWT avec un modèle unitaire
MTF-GLP	Une méthode populaire basée sur l'ARM fondée sur le GLP avec un filtre MTF ; le modèle de détail unitaire est appliqué
ATWT-M2	
MTF-GLP-HPM	Le MTF-GLP avec le schéma spatial-détail du HPM
MTF-GLP-HPM-HPM	Le MTF-GLP-HPM suivi d'une phase PP pour corriger le bruit aux régions de bord
MTF-GLP-CBD	Le MTF-GLP avec le schéma de détails spatiaux du CBD, qui est optimisé localement en l'image dans des zones non superposées
ATWT-M3	L'ATWT avec le modèle, qui est basé sur l'ajustement des moindres carrés pour ajuster les coefficients d'ondelettes de l'image Pan
AWLP	Une généralisation de la méthode de fusion de pansharpening AWL en termes de poids d'injection de détail spatial proportionnel, reposant sur la proportionnalité de la radiance de la bande MS d'origine
VO	methides Variationnelles
FE-HPM	Estimation de filtre basée sur un cadre de déconvolution semi-aveugle et un modèle d'injection HPM
SR-D	Fusion panoramique basé sur une représentation clairsemée des détails injectés

PWMBF	Fusion basée sur un modèle à l'aide de l'ACP et des ondelettes
DL	Deep Learning
PNN	Le modèle PNN est relativement simple : un réseau sériel composé de trois unités convolutives entrelacées par les activations ReLU. Les principaux hyperparamètres architecturaux du PNN sont regroupés dans SUPPORT SPATIAL BANDES D'ENTRÉE BANDES DE SORTIE
PNN-IDX	PNN augmentant l'entrée en incluant plusieurs cartes d'indices radiométriques non linéaires Il s'agit d'indices radiométriques bien établis tels que la différence d'indices végétation/eau, pris en compte pour tous ensembles de données du présent travail, ou la différence normalisée de l'indice de sol et la différence de caractéristiques non homogènes, a performances supérieures du PNN-IDX Sur PNN.
A-PNN-FT	Versions avancées de PNN peut être très différent en termes de complexité intègrent des modules PNN, A-PNN, PNN-IDX
AUTRE	
HPF	Une méthode typique basée sur l'ARM ; dans les expériences, le paramètre par défaut avec un filtre boîte 5×5 a été utilisé
SFIM	Une méthode HPF améliorée basée sur le schéma de détails spatiaux de HPM, qui est calculée en fonction du rapport de l'image MS ré échantillonnée et de l'image Pan de filtrage passe-bas
INDUSION	L'indusion combine l'induction et la fusion ; c'est une méthode de fusion basée sur la technique de mise à l'échelle par induction et bénéficie de plusieurs étapes d'égalisation pour améliorer les performances
IHS	Une généralisation de la méthode traditionnelle de fusion pansharpening IHS pour la fusion de plus de trois bandes MS, on l'appelle aussi la méthode de fusion IHS généralisée rapide

Tableau 4-2 : LES MÉTHODES DE PANARPENING DANS LES EXPÉRIENCES [44]

Résultat de comparaison

Les images utilise (QuickBird satellite) dans la comparaison est comme suivant:



Figure 4.8: image PAN (urbaines 1024*1024)



Figure 4.9: image MS (urbaines 256*256*4)

➤ **REDUCED RESOLUTION ASSESSMENT (ÉVALUATION DE LA RÉOLUTION RÉDUITE) :**



Figure 4.10: Résultats fusionnés (Rr) sur les premières données QuickBird simulées (256 × 256). (a) GT (b) EXP (c) BT-H (d) BDSD (e) C-BDSD (f) BDSD-PC (g) GS (h) GSA (i) PRACS (j) AWLP (k) MTF-GLP (L) MTF-GLP-FS (M) MTF-GLP-HPM (N) MTF-GLP-HPM-H (O) MTF-GLP-HPM-R (P) MTF-GLP-CBD (Q) C-MTF-GLP-CBD (R)MF (S)FE-HPM (T)SR-D (U)PWMBF (V)TV (W)RR (X)PNN (Y)PNN-IDX (Z)A-PNN , PAN-GAN

			ERGAS	Q	Q-avg	SAM	SCC-GT	Time (s)
CS	EXP		3,8347	0,4641	0,4572	4,331	0,6718	
	BT-H		2,3447	0,8059	0,7768	3,9887	0,9224	0,125
	BDSD		2,1781	0,8445	0,8298	3,8711	0,9404	0,1629
	C-BDSD		2,3656	0,8417	0,8288	3,9302	0,9265	0,8224
	BDSD-PC		2,1798	0,8444	0,8296	3,8702	0,9403	0,2836
	GS		2,8493	0,6962	0,6758	4,1863	0,8973	0,0633
	GSA		2,2127	0,8273	0,8099	3,9383	0,9372	0,1063
	C-GSA		2,2402	0,8201	0,8012	3,8881	0,9304	0,3191
	PRACS		2,5646	0,7529	0,7324	3,9821	0,9084	
MRA	AWLP		2,2028	0,8375	0,8218	3,9476	0,9367	0,1527
	MTF-GLP		2,2386	0,8423	0,8273	3,9453	0,9389	0,2068
	MTF-GLP-FS		2,2012	0,8318	0,8146	3,936	0,938	0,1227
	MTF-GLP-HPM		2,2341	0,8428	0,8281	3,9555	0,9396	0,1959
	MTF-GLP-HPM-H		2,3177	0,8135	0,7838	3,9872	0,9238	0,1571
	MTF-GLP-HPM-R		2,1947	0,8335	0,8164	3,9555	0,9382	0,1466
	MTF-GLP-CBD		2,2019	0,8318	0,8144	3,9393	0,9379	0,114
	C-MTF-GLP-CBD		2,2314	0,8241	0,805	3,8886	0,9314	0,4042
	MF		2,407	0,7955	0,7781	4,0275	0,9249	0,1959
VO	FE-HPM		2,3515	0,7996	0,7824	4,0227	0,9396	0,4286
	SR-D		2,3078	0,813	0,203	3,8657	0,929	1,7067
	PWMBF		2,6018	0,7433	0,7238	4,3377	0,914	0,4702
	TV		3,113	0,6673	0,6441	4,2437	0,8611	2,3722
	RR		6,0763	0,5584	0,203	8,9689	0,7883	5,0943
DL	PNN		2,5692	0,8086	0,7924	4,0241	0,9077	4,9837
	PNN-IDX		2,4795	0,8164	0,7908	4,0816	0,9229	0,9934
	A-PNN		2,4903	0,8136	0,791	3,7664	0,9217	0,8626
	PAN-GAN		2,3071	0,8218	0,813	3,4877	0,9333	7,9332
	GT		Le plus petite c'est le meilleur	Le plus grand c'est le meilleur	Le plus grand c'est le meilleur	Le plus petite c'est le meilleur	Le plus grand c'est le meilleur	Le plus petite c'est le meilleur

Tableau 4-3 : les performances ERGAS, Qm,SAM,Scs-GT,Qavg,Time) calculées pour de données à Résolution complet (Rr)

➤ **FULL-RESOLUTION ASSESSMENT (ÉVALUATION À RÉOLUTION COMPLÈTE) :**

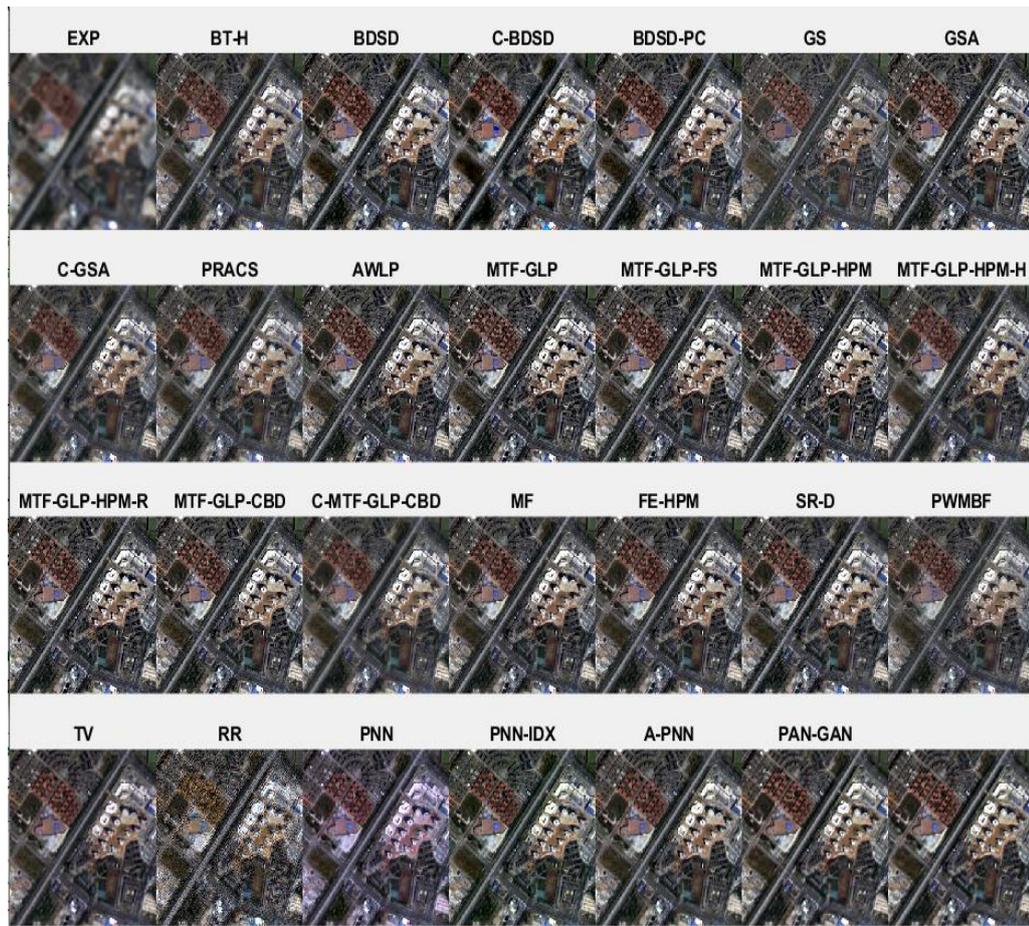


Figure 4.11: Résultats fusionnés (Fr) sur les premières données QuickBird simulées (256×256). (a) EXP (b) BT-H (c) BDS (d) C-BDS (e) BDS-PC (f) GS (g) GSA (h) PRACS (i) AWLP (j) MTF-GLP (k) MTF-GLP-FS (L) MTF-GLP-HPM (M) MTF-GLP-HPM-H (N) MTF-GLP-HPM-R (O) MTF-GLP-CBD (P) C-MTF-GLP-CBD (Q) MF (R) FE-HPM (S) SR-D (T) PWMBF (U) TV (V) RR (W) PNN (X) PNN-IDX (Y) A-PNN, PAN-GAN, (Z)GT

			D lambda	Ds	QNRI
CS	EXP		0,0482	0,2655	0,6991
	BT-H		0,0341	0,1285	0,8418
	BDS		0,0346	0,1645	0,8065
	C-BDS		0,0496	0,125	0,8316
	BDS-PC		0,0343	0,1647	0,8067
	GS		0,1626	0,0828	0,768
	GSA		0,364	0,1878	0,7827
	C-GSA		0,039	0,1841	0,784
	PRACS		0,0358	0,1059	0,9621
MRA	AWLP		0,0265	0,1983	0,7805
	MTF-GLP		0,0419	0,2189	0,7484
	MTF-GLP-FS		0,0349	0,1921	0,7797
	MTF-GLP-HPM		0,0422	0,2185	0,7485
	MTF-GLP-HPM-H		0,0319	0,1282	0,844

	MTF-GLP-HPM-R	0,035	0,1915	0,7802
	MTF-GLP-CBD	0,0349	0,1913	0,7805
	C-MTF-GLP-CBD	0,0336	0,1625	0,8093
	MF	0,0367	0,1616	0,8076
VO				
	FE-HPM	0,0315	0,1684	0,8054
	SR-D	0,0173	0,1049	0,8796
	PWMBF	0,0815	0,1663	0,7658
	TV	0,0797	0,0518	0,8726
	RR	0,0381	0,5964	0,3882
ML				
	PNN	0,0417	0,1107	0,8522
	PNN-IDX	0,0387	0,1404	0,8263
	A-PNN	0,0196	0,1203	0,8625
	PAN-GAN	0,0155	0,0461	0,9391
	GT	Le plus petite c'est le meilleur	Le plus petite c'est le meilleur	Le plus petite c'est le meilleur

Tableau 4-4 : les performances ($D\mu$, Ds , QNR) calculées pour de données à Résolution complet (Fr)

Comme le montrent les tableaux 4-4 et 4-3 et les figures 4-11 et 4-12, les méthodes basées sur DL montrent des bonnes performances dans l'expérience. Parmi eux, la méthode BDSD a d'excellents résultats dans les expériences à résolution réduite et à pleine résolution. Bien que la GSA, Les méthodes MTF-GLP et MTF-GLP-HPM PP sont satisfaisantes performances dans l'expérience à résolution réduite, leur les performances dans l'expérience à pleine résolution sont légèrement plus pauvres. Les performances de les Méthode basée sur DL dans la résolution réduite et pleine les expériences sont robustes.

IHS, ATWT, et relativement dans toutes les méthodes traditionnelles sauf BDSD, montrent des caractéristiques plus floues en ce qui des autres méthodes.

Les méthodes BDSD et Les méthodes basées sur DL montrent une excellente et performances robustes dans les expériences à résolution réduite et à pleine résolution. Les méthodes BDSD, ATWT-modèle 2 (M2), ATWT-modèle 3(M3), MTF GLP HPM, MTF-GLP-CBD montrent performances globalement satisfaisantes. Alors que la meilleure performance est indiquée pour la méthode MTF-GLP-CBD dans l'expérience en pleine résolution, l'approche a eu un des performances plus faibles dans l'expérience à résolution réduite, en particulier dans l'évaluation quantitative ERGAS et SAM indices.

Les méthodes basées sur DL montre la meilleure performance surtout **PAN-GAN** est à des avantages concurrentiels par rapport à la d'autres méthodes de pansharpening dans la résolution réduite expérience. De plus, il affiche des performances robustes en l'expérience en pleine résolution.

En particulier dans les indices d'évaluation quantitative de l'ERGAS, SAM, Ds et QNR. La méthode PAN-GAN montre bonnes performances dans l'expérience en pleine résolution et performances relativement inférieures dans l'expérience à résolution réduite, En termes de caractéristiques spatiales,

4.7 Conclusion :

D'après les résultats obtenus des méthodes d'évaluation de la qualité des images fusionnées, de comparaison reduced résolution tableau 4.3 et les méthodes full-résolution tableau 4.4 on peut conclure que les méthodes basées sur le Deep learning présente de meilleures performances.

La méthode PanColorGAN montre la meilleure performance, car il inclut le caractère aléatoire introduit dans ses rapports de réduction d'échelle qui augmente sa robustesse aux variations de résolution minute entre le panchromatique réduit et images multispectrales.

Les images produites par ces modèles peuvent être utilisé efficacement dans des applications d'ingénierie telles que les tâches de détection d'objet, où une qualité spatiale élevée et une géométrie d'objet précise est requise. Comme les propriétés spectrales sont également préservées à un niveau acceptable

Conclusion générale

5 Conclusion générale :

La fusion d'images est une tâche importante dans le domaine de la télédétection, c'est une méthode potentielle dans les applications de télédétection qui intègrent les informations de combinaisons de panchromatiques, d'images multispectrales, destinés à produire une image composite ayant à la fois des résolutions spatiales et spectrales plus élevées.

L'état de l'art actuel de la fusion d'images en télédétection, présente différentes méthodes de fusion et d'évaluation.

Grâce à l'apprentissage en profondeur (Deep Learning) ; les systèmes d'informations géographiques sont en train de s'imposer comme des outils d'études non incontournables dans le monde, ceci grâce à la puissance de calculs des algorithmes des logiciels.

Dans ce mémoire, nous avons implémenté et tester plusieurs méthodes de fusion et nous avons évalué les performances de la qualité des images fusionnés.

Les images produites par ces modèles présentent de bonne performances et les propriétés spectrales sont également préservées à un niveau acceptable et peuvent être utilisé efficacement dans des applications d'ingénierie pour la télédétection , où une qualité spatiale élevée et une géométrie d'objet précise est requise.

Nous avons aussi fait l'implémentation de la méthode de fusion PanColorGAN qui est un réseau de haute qualité du GAN pour l'amélioration de la netteté des images de télédétection. Le résultat obtenu des images lors de la phase de test confirme l'efficacité de cette méthode.

Notre travail n'est que dans sa version initiale, on peut dire que ce travail reste ouvert pour des travaux de comparaison avec d'autres méthodes de fusion.

Les annexes :

Code de programmer : (1)

```
# Commented out IPython magic to ensure Python compatibility.
```

```
import os
torch
torch.nn as nn
torch.optim as optim
networks import define_G
import
import
import
from
from util
```

```

import patch_sixteen, divide_sixteen, show_image, scale_range
torch.backends.cudnn as cudnn
as np
as plt

import
import numpy
import matplotlib.pyplot

# ["Default Images", "Upload Images"]

dropdown = 'Default Images'
if dropdown == "Upload Images":
    uploaded = files.upload()
    keys=[]
    for fp in uploaded.keys():
        print('User uploaded file "{name}" with length {length} bytes'.format(
            name=fp, length=len(uploaded[fp])))
    keys.append(fp)
    if len(uploaded[keys[0]])>len(uploaded[keys[1]]):
        pan_fn, ms_fn = keys[0], keys[1]
    else:
        pan_fn, ms_fn = keys[1], keys[0]

    ms = tifffile.imread(ms_fn)
    pan = tifffile.imread(pan_fn)
    print(ms.shape,pan.shape)
    ms = ms.astype(np.float32)
    pan = pan.astype(np.float32)

elif dropdown == "Default Images":
    ms_filename = 'dg_msimage200.npy'
    pan_filename = 'dg_panimage200.npy'
    ms = np.load(ms_filename).astype(np.float32)
    pan = np.load(pan_filename).astype(np.float32)
    print(ms.shape,pan.shape)

## Data Display

show_image(ms)
show_image(pan)

## Parameters

reduced_res=False
use_gpu=True

## Data Preprocessing

from skimage.transform import resize
ms_norm = np.array([scale_range(i, -1, 1) for i in ms.transpose((2,0,1))])
pan_norm = scale_range(pan, -1, 1)

```

```

if reduced_res:
    ms_down = [resize(i,(64,64), 3) for i in ms_norm]
    ms_up = [resize(i, (256, 256), 3) for i in ms_down]
    ms_norm = np.clip(ms_up,-1.0,1.0)
    pan_down = resize(pan_norm, (256, 256), 3)
    pan_norm = np.clip(pan_down,-1.0,1.0)
else:
    ms_up = [resize(i, (1024, 1024), 3) for i in ms_norm]
    ms_norm = np.clip(ms_up,-1.0,1.0)

inp = np.concatenate((ms_norm,np.expand_dims(pan_norm,axis=0)),axis=0)
if not reduced_res:
    inp = divide_sixteen(inp)
inp = torch.tensor(inp)

## Network Preparation

model_weights = "PanColorGANRD.pth"
netG = define_G(5, 4, 32, 'batch','leakyrelu', False, 'ResizeConv', 'colorunet', 'SE', 6, [0],
n_downsampling=2)
#netG.load_state_dict(torch.load(model_weights, map_location=lambda storage, loc:
storage))
netG.load_state_dict(torch.load(model_weights))
netG.eval()
if use_gpu:
    netG = netG.cuda()
else:
    netG = netG.cpu()

if use_gpu:
    out = netG(inp.cuda().view(-1,5,256,256)).detach().cpu().numpy()
else:
    out = netG(inp.view(-1,5,256,256)).detach().numpy()
if not reduced_res:
    out = patch_sixteen(out)

#show image MS HR

show_image(out.transpose(1,2,0))

```

Définition de Clustering :(2)

Tout d’abord, le clustering, ou partitionnement, consiste à identifier des groupes dans les données (voir figure 3.4). Cela permet de comprendre leurs caractéristiques générales, et éventuellement d’inférer les propriétés d’une observation en fonction du groupe auquel elle appartient.

On appelle partitionnement ou clustering un problème d’apprentissage non supervisé pouvant être formalisé comme la recherche d’une partition $Ck = e^k$ des n observations $\{x_i\}_{i=1,\dots,n}$. Cette partition doit être pertinente au vu d’un ou plusieurs critères à préciser. [31]

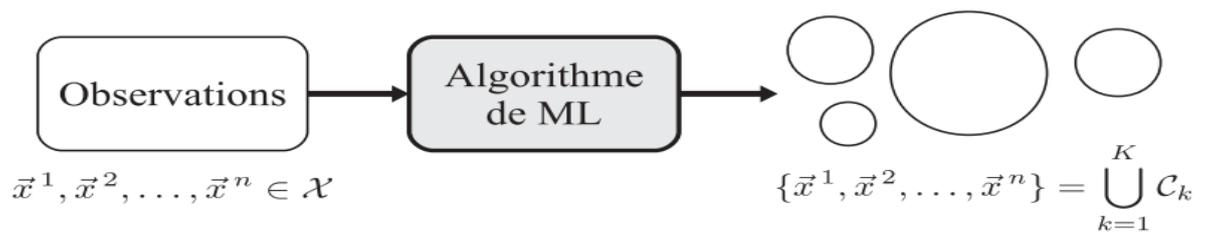


Figure 3.4 : Partitionnement des données, ou clustering.

Définition de réduction de dimension : (3)

On appelle réduction de dimension un problème d'apprentissage non supervisé pouvant être formalisé comme la recherche d'un espace Z de dimension plus faible que l'espace X dans lequel sont représentées n observations $\{x_i\}_{i=1, \dots, n}$. Les projections $\{z_i\}_{i=1, \dots, n}$ des données sur Z doivent vérifier certaines propriétés à préciser. [31]

1 Références bibliographiques

- [1] télédétection, Larousse, URL "<https://www.larousse.fr/dictionnaires/francais/77051>"
- [2] Un cours tutoriel du Centre canadien de télédétection, "*Notions fondamentales de télédétection*", Ressources naturelles Canada, URL "https://www.nrcan.gc.ca/sites/www.nrcan.gc.ca/files/earthsciences/pdf/resource/tutor/fundam/pdf/fundamentals_f.pdf" , vol. 105, no. 5. 2012.]
- [3] Bonn , Rochon, "Les ondes électromagnétiques", URL <https://e-cours.univ-paris1.fr/modules/uved/envcal/html/rayonnement/1-rayonnement-electromagnetique/1-2-les-ondes-electromagnetiques.html>.
- [4]"Le spectre électromagnétique" , <https://ecours.univparis1.fr/modules/uved/envcal/html/rayonnement/1-rayonnement-electromagnetique/1-4-spectre-electromagnetique.html>
- [5] A. HABBA, O. ISHAK. "La classification des images satellitaires par l'apprentissage profonde (deep learning) «. Thèse de doctorat. Université Ahmed Draïa-Adrar, Algeria, 1-58 2019.
- [6] F.E.Gonzalez,J.M.Ruiz,F.M.Acosta, "MANUEL de TELEDETECTION SPATIALE TELECAN" ,Universidad de Las Palmas de Gran Canaria. URL https://www.grss-ieee.org/wp-content/uploads/2014/07/FR_TUTORIAL_COMPLETO.pdf
- [7] H. Amir Mokhtar. "Conception d'un système hybride entre la Méta-heuristique Electromagnétique et l'algorithme de colonie de fourmis pour la classification des images satellitaires". Mémoire de MAGISTER. Université des sciences et de la technologie d'Oran Mohamed Boudiaf. 2010
- [8] Mr R. Seddik. "Contribution d'une méta-heuristique bio-inspirée dans le traitement d'image «. Mémoire de Magistère. Université des sciences et de la technologie d'Oran Mohamed Boudiaf.2014.
- [9]https://www.esa.int/SPECIALS/Eduspace_FR/SEM756S7NWF_0.html
- [10]https://www.esa.int/SPECIALS/Eduspace_FR/SEMZ76S7NWF_0.html
- [11] <https://www.geoimage.com.au/satellite/worldview-2>
- [12] <https://www.satimagingcorp.com/satellite-sensors/worldview-3/>
- [13] <https://earth.esa.int/eogateway/missions/worldview-4>
- [14]<https://directory.eoportal.org/web/eoportal/satellite-missions/g/gaofen-1>

- [15] <https://www.universalis.fr/encyclopedie/teledetection/4-les-applications-de-la-teledetection/>
- [16] Mme G. Djaouher. "Comparatif de méthodes avancées pour la fusion d'images satellites". Mémoire de MAGISTER. Université d'Oran des sciences et de la technologie USTO-MB.2015.
- [17] M. BELKACEMI. "Analyse de texture d'images de télédétection et vectorisation des paramètres de texture". Mémoire d'ingénieur d'état en Electronique. Université Saad Dahleb de Blida .2007.
- [18] F. MULLER, S.DE BÉTHUNE." LA FUSION D'IMAGES".Bulletin de la Société Géographique de Liège, 38, 2000/1, 79-94.
- [19] M. CHIKR EL MEZOUAR. "Fusion d'images en télédétection satellitaire. Thèse de Doctorat. L'INSA de Rennes sous le sceau de l'Université européenne de Bretagne et l'Université Djillali Liabes de Sidi Bel-Abbes .2012
- [20] "Les images de télédétection". URL " <https://eo.belspo.be/fr/les-images-de-teledetection> "
- [21] F.Z. BENHALOUCHE. "Méthodes de démixage et de fusion des images multispectrales et hyperspectrales de télédétection spatiale". Thèse Doctorat. Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier) Cotutelle internationale avec l'Université des Sciences et de la Technologie d'Oran - Mohamed Boudiaf (Algérie).2018
- [22] X. MENG, Y. XIONG, F. SHAO, H. SHEN, W.SUN, G. YANG, Q. YUAN, R. FU, H. ZHANG. "A Large-scale Benchmark Data Set for Evaluating Pansharpening Performance".IEEE GEOSCIENCE AND REMOTE SENSING M 36 AGAZINE.2020.
- [23] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rougé, "A variational model for P+XS image fusion," International Journal of Computer Vision, vol. 69, pp. 43–58, 2006.
- [24] DataScientest . " qu'est-ce que le Deep Learning ?".URL "<https://datascientest.com/deep-learning>"
- [25]. M.BENARD. "Potocole de recherche documentaire, Protocole de recherche terminologique (L'apprentissage automatique appliqué à l'évaluation de la traduction automatique «. Mémoire de Master 2. Université Paris Diderot (Paris 7).2019.
- [26] FUTURA TECH. "Deep Learning : qu'est –ce que c'est ? ".URL "<https://www.futura-sciences.com/tech/definitions/intelligence-artificielle-deep-learning-17262/>"
- [27] Analyse d'images par méthode de Deep Learning appliquée au contexte routier en conditions météorologiques dégradées // Docteur Khouloud DAHMANE// l'Université Clermont Auvergne

- [28] [www.demos.fr/Intelligence artificielle](http://www.demos.fr/Intelligence%20artificielle)
- [29] Eloïse Berthier, vendredi 8 mars 2019
- [30] Livre Introduction Au Machine Learning Chloé-Agathe Azencott
- [31] Livre PRÉSENTATION DU MACHINE LEARNING
- [32] Livre Principaux algorithmes du Machine Learning
- [33] Deep Learning pour la classification des images / Moualek Djaloul Youcef / Université Abou Bakr Belkaid Tlemcen Faculté des Sciences Département d'Informatique[g]
- [34] www.ionos.fr/search-engine-marketing/deep-learning-vs-machine-learning/
- [35] Les Réseaux de Neurones Artificiels /Y. Djeriri /University of Sidi-Bel-Abbes /Department Cai [p]of Electrical Engineering
- [36] www.juripredis.com
- [37] <https://moncoachdata.com/blog/comprendre-les-reseaux-de-neurones/>
- [38] A New Deep Learning Based Multi-Spectral Image Fusion Method / Jingchun Piao, Yunfan Chen and Hyunchul Shin /Department of Electrical Engineering, Hanyang University, Ansan 15588, Korea
- [39] Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion
- [40] Xiangyu Liu (Beijing Advanced Innovation Center for Big Data and Brain Computing), Yunhong Wang (Beijing Key Laboratory of Digital Media, School of Computer Science and Engineering, Beihang University, Beijing 100191, China), Qingjie Liu./ PSGAN: A GENERATIVE ADVERSARIAL NETWORK FOR REMOTE SENSING IMAGE PAN-SHARPENING
- [41] Yu Liua , Xun Chena,b,* , Zengfu Wangc , Z. Jane Wangd , Rabab K. Wardd , Xuesong Wange, \Deep learning for pixel-level image fusion: Recent advances and future prospects
- [42] ARIAN AZARANG1 , (Student Member, IEEE), HAFEZ E. MANOOCHEHRI1 , (Student Member, IEEE), NASSER KEHTARNAVAZI , (Fellow, IEEE)/Convolutional Autoencoder-Based Multispectral Image Fusion
- [43] Zhenfeng Shao, Member, IEEE, and Jiajun Cai /Remote Sensing Image Fusion with Deep Convolutional Neural Network
- [44] [ieee_geoscience_march2021.pdf](#)
- [45] Remote Sensing Image Fusion with Deep Convolutional Neural Network Zhenfeng Shao, Member, IEEE, and Jiajun Cai

- [46] www.lebigdata.fr/reseau-de-neurones-artificiels-definition
- [47] www.juripredis.com/fr/blog/id-19-demystifier-le-machine-learning-partie-2-les-reseaux-de-neurones-artificiels
- [48] www.oriondata-i.com/home/index.php/fusion-pan-sharpening
- [49] opendatascience.com/an-introduction-to-reinforcement-learning-concepts/
- [50] Ozcelik, F., Alganci, U., Sertel, E., & Unal, G. (2020). Repenser l'affinage panchromatique basé sur CNN : colorisation guidée des images panchromatiques via les GAN. arXiv préimpression arXiv:2006.16644.
- [51] <https://www.python.org>
- [52] <https://docs.anaconda.com/anaconda/navigator>
- [53] [cloud.google.com/colaboratory /gpu](http://cloud.google.com/colaboratory/gpu)
- [54] Manuel MATLAB
- [55] Vaibhav R. Pandit Research Scholar S.G.B. Amravati University, Amravati, India R. J. Bhiwani, PhD Professor B.N. College of Engineering, Pusad, India / Image Fusion in Remote Sensing Applications: A Review
- [56] https://bioinfo.irc.ca/fr/introduction_a_la_regression
- [57] scikit-hep.org/root_numpy/auto_examples/tmva/plot_multiclass.html
- [58] <https://www.actuia.com/keras/>
- [59] www.tensorflow.org
- [60] <https://pytorch.org/>
- [61] <https://pypi.org/project/Theano/>
- [62] nlpforhackers.io/classification-performance-metrics

