

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieure et de la Recherche Scientifique
Université Ahmed Draia - Adrar
Faculté des Sciences et de la Technologie
Département des Mathématiques et Informatique



Mémoire de fin d'étude, en vue de l'obtention du diplôme de Master
en informatique

Option : Systèmes Intelligents

Thème

La classification des images satellitaires par l'apprentissage
profonde (deep learning)

Préparé par

Abdelaziz HABBA et Omar ISHAK

Encadré par

Mr. OUAHAB Abdelwhab

Membres du jury :

Président : Mr. CHOUGUEUR Djilali

Examineur : Mr. BENATIALLAH Djeloul

Examineur : Mr. MAMOUNI EL Mamoune

Année Universitaire 2018/2019

Résumé

Au fil du temps, l'homme a développé les moyens qu'il a utilisés pour identifier la terre et révéler sa richesse, Alors Passer de l'enquête à l'aide d'outils simples reposant sur le contact direct à la photographie aérienne utilisant un avion Jusqu'aux atteindre les satellites , Ce qui lui a permis de localiser, cartographier, découvrir des richesses et identifier des problèmes environnementaux , En plus de l'espionnage et des fonctions scientifiques sans contact direct avec eux.

Il a également mis au point des programmes et du matériel permettant de classer les images prises par satellite, ce qui nous permet de déterminer le terrain et de déterminer les lieux et les changements qui se produisent au niveau du sol.

Le Deep Learning ou « apprentissage profond » est une famille d'algorithmes d'apprentissage (machine learning) pour entraîner des réseaux de neurones composés de plusieurs couches internes et potentiellement un grand nombre. Ces techniques ont permis des progrès importants et rapides dans les domaines de l'analyse du signal sonore ou visuel et notamment de la reconnaissance faciale, de la reconnaissance vocale, de la vision par ordinateur, du traitement automatisé du langage, traitement et classification des images.

Dans ce travail, nous avons pu obtenir de bons résultats et nous avons apporté plusieurs améliorations en apportant plusieurs modifications pour obtenir les meilleurs résultats.

ملخص

طور الانسان عبر الزمن الوسائل التي استخدمها في التعرف على الارض والكشف عن ثرواتها , فانقل من المسح باستخدام الادوات البسيطة التي تعتمد على التماس المباشر الى التصوير الجوي باستخدام الطائرات حتي وصل الى الاقمار الصناعية , التي مكنته من تحديد المواقع ورسم الخرائط والكشف عن الثروات وتحديد المشكلات البيئية , اضافة الى وظائف تجسسية وعلمية دون تماس مباشر معها. كما طور برامج ومعدات من اجل تصنيف الصور التي تلتقطها الاقمار الصناعية التي تسمح لنا بتحديد التضاريس وتحديد الاماكن و التغييرات التي تحدث على مستوى الارض

التعلم العميق عبارة عن مجموعة من خوارزميات التعلم الآلي لقيادة الشبكات العصبية التي تتألف من عدة طبقات داخلية وربما عدد كبير. هذه التقنيات حققت تقدما كبيرا وسريع في مجالات تحليل الإشارات الصوتية أو البصرية وخاصة التعرف على الوجه ، التعرف على الكلام ، رؤية الكمبيوتر ، معالجة اللغة الآلية ، معالجة الصور وتصنيفها .

في هذا العمل، تمكنا من الحصول على نتائج جيدة، كما قدمنا عدة تحسينات من خلال تقديم العديد من التغييرات للحصول على افضل النتائج .

Dédicaces

Je dédie ce noble travail ...

*À mon père et à ma mère qui m'ont soutenu et
encouragé dans les moments les plus difficiles...*

*À mes frères et à mes sœurs pour leur patience et
leur affection...*

À toute ma famille et à tous mes amis...

HABBA

ABD ELAZIZ

Dédicace

Je dédie ce modeste travail

A Mon Cher père « Mohamed »

Qui est le meilleur père dans ce monde, grâce à son encouragement, sa confiance et son soutien moral et matériel et pour son amour infini en exprimant mes gratitude, mon profond amour et ma passion

A ma chère mère « Alia »

En témoignage de ma profonde gratitude et de mon incontestable reconnaissance, pour tous les sacrifices qu'elle me contente, toute la confiance qu'elles m'accordent et tout l'amour dont elle m'entoure.

A mes chers frères « sidi Ali » « Abderrahmane » « Abdelkrim » « moussa » « Ahmed »

A mes chers sœurs « Fatma » « Taousse » « Aicha » « Oiar da » « Halima » « Akida » « Zainab » « Khadidja » « Samia » « Khaoilla »

A mon fiancé « Badjaloul Sara » qui m'a aidé dans les moments difficiles

A toutes les familles « Ishak, Abdaïem, Taleb Alli, Badjaloul, Ben Oumar, Lekssassi, Dartagnon »

A tout mes amies surtout « Ayoub Ismaïl, Idriss, Aboelmajde, Haitham, Adam, Ahmed, zoubir, ossama »

A mon Binôme qui m'a supporté durant toute la durée de la réalisation de ce projet et a su excusé mes retards, « HABBÂ Abdelaziz »

Ishak Omar

Remerciements

Merci beaucoup

Merci ...

Dieu m'éclaire et me guide sur le chemin de la bonté ...

Mes plus sincères remerciements vont à mes supérieurs
hiérarchiques: Mr. OUAHAB Abdelwhab

Qui m'a guidé et m'a soutenu

Sans hésitation ni détente du choix et atteindre ce thème.

Je suis reconnaissant

Enfin, je remercie tout le cadre éducatif et administratif sans
lequel j'ai la tâche Ce ne sera pas fait.

Je remercie tous mes professeurs chacun par son nom pour
leur patience et leur abnégation tout au long de mes études.

Sommaire

Résumé	II
Dédicaces	III
Remerciement	V
Sommaire.....	VI
Liste des figures	IX
Liste des abbreviation	X
Introduction générale	01

Chapitre 1 : Télédétection

I. Introduction	04
II. La télédétection.....	04
II.1. Définition de la télédétection	04
II.2. Les différentes étapes de la télédétection.....	04
II.2.1. Source d'énergie ou d'illumination.....	04
II.2.2. Rayonnement et atmosphère	04
II.2.3. Interaction avec la cible.....	04
II.2.4. Enregistrement de l'énergie par le capteur	05
II.2.5. Transmission, réception et traitement.....	05
II.2.6. Interprétation et analyse.....	05
II.2.7. Application	05
III. Rayonnement électromagnétique.....	05
III.1. Définition.....	05
III.2. Le spectre électromagnétique.....	06
III.3. Propriétés des ondes électromagnétiques.....	07
IV. Système d'acquisition.....	09
IV.1. Les capteurs.....	09
IV.2. Les plates formes.....	09
V. Imagerie satellitaire.....	11
V.1. Introduction.....	11
V.2. Définitions.....	11
V.2.1. Résolution (spatiale).....	12
V.2.2. Valeurs des pixels.....	12
V.3. Les paramètres influant sur le contenu d'une image satellitaire.....	13
V.3.1. La définition de l'image.....	13
V.3.2. Les canaux disponibles.....	13
V.3.3. La date de prise de vue.....	13
V. 4. Les prétraitements.....	13
V.4.1. Modification d'histogramme.....	13
V.4.2. Réduction de bruit.....	14
V.4.3. Rehaussement de contraste.....	14
V.5. Les traitements.....	14
V.5.1. Les traitements statistiques.....	14
V.6. Buts de traitements.....	15

Sommaire

VI. Classification.....	15
VI.1. Définitions.....	15
VI.2. Méthodes de classification.....	16
VI.2.1. Images texturales.....	17
VI.2.2 : Image des régions.....	17
VI.2.3 : Image des contours.....	17
VI.2.4 : Image « masque ».....	17
VI.2.5: Image du MNT (Modèle Numérique de Terrain).....	17
VI.2.6 : SIG (Système d'information géographique).....	17
VI.2.7 : Image du réseau routier et/ou hydrographique.....	17
VI.3. Techniques de classification.....	17
VI.3.1. Procédure de classification.....	17
VI.3.2. La classification non dirigée.....	17
VI.3.3. La classification dirigée.....	18
VI.4 : Marche à suivre pour une classification supervisée.....	18
VI.4.1.Marche à suivre pour une classification non supervisée.....	19
VII. Conclusion.....	19

Chapitre 2 : L'apprentissage profond (deep learning)

I. Introduction.....	22
II. Définition de l'apprentissage profond (deep learning).....	22
III. Pourquoi le deep learning ?.....	23
III.1 Domaines d'application de l'apprentissage profond.....	24
III.2 Architectures de réseaux de neurones profonds.....	24
III.2.1. Les réseaux de neurones convolutifs.....	24
III.2.2. L'entraînement d'un nouveau CNN est difficile.....	25
III.2.2.1. La couche de convolution (CONV).....	25
III.2.2.2.Couche de pooling (POOL).....	26
III.2.2.3.Couche de correction (ReLU).....	27
III.2.2.4.Couche entièrement connectée(FC).....	28
III.2.2.5.Couche de perte (LOSS).....	28
III.2.2.6.Perceptron.....	28
III.2.2.7.Choix des hyperparamètres.....	29
III.2.2.7.1- Nombre de filtres.....	29
III.2.2.7.2- Forme de filtres.....	29
III.2.2.7.3- Forme de Max pooling.....	29
III.2.3. Réseau de neurones récurrents.....	30
III.2.3.1. C'est quoi un RNN?.....	30
III.3.Avantages de CNNs.....	31
IV.Conclusion.....	32

Sommaire

Chapitre 3: Implémentation et résultat

I. Introduction.....	34
II. Ressources matérielles et logicielles	34
II.1. Ressources matérielles.....	34
II.2. Ressources logicielles	34
II.2.1 Python	34
II.2.2. PyCharm	35
II.2.3. Anaconda	35
II.2.4. TensorFlow	36
II.2.4.1. Pourquoi TensorFlow ?	37
II.2.5. PyQt	37
III. Description de CNN:	38
IV. Base d'image:	40
V. L'architecture de l'application.....	41
V.1. Base d'apprentissage	42
V.2. L'apprentissage CNN	42
V.3. Sélectionner une région	43
V.4. Evaluation de résultat :.....	44
VI. Conclusion	45
Conclusion générale	46
Références bibliographiques.....	47

Liste des figures

Chapitre 1:

Figure 1.1: Les étapes de télédétection.....	05
Figure 1.2 : Rayonnement électromagnétique.....	06
Figure 1.3: Domaines spectraux.....	07
Figure 1.4: Propriété des ondes électromagnétiques.....	08
Figure 1.5: Capteur passif et actif.....	09
Figure 1.6: Plates formes (avion, satellite).....	10
Figure 1.7: Image satellitaire « Oran » (LANDSAT 5 -TM).....	11
Figure 1.8: Résolution spatiale.....	12
Figure 1.9: Principe de classification.....	16
Figure 1.10: Représentation des méthodes de classification.....	16

Chapitre 2:

Figure 2.1: La relation entre l'intelligence artificielle, le ML et le deep learning.....	22
Figure 2.2 : Le procédé du ML classique comparé à celui du Deep Learning	24
Figure 2.3: Les réseaux de neurones convolutifs.....	25
Figure 2.4: Exemple d'une convolution 2D	26
Figure 2.5: Average pooling et Max pooling	27
Figure 2.6: Un réseau de neurones convolutif	28
Figure 2.7: Exemple de max pooling (2×2).....	30
Figure 2.8: Un RNN	31

Chapitre 3:

Figure 3.1: Logo Python	34
Figure 3.2 : Interface JetBrains PyCharm	35
Figure 3.3: Interface Anaconda	36
Figure 3.4: logo TensorFlow	37
Figure 3.5: interface Qt Designer	38
Figure 3.6: La structure CNN	39
Figure 3.7: Image UC Merced land-use Dataset (RGB)	41
Figure 3.8: Architecteur de l'application	41
Figure 3.9: Apprentissage Image	42
Figure 3.10: Insérer l'image	43
Figure 3.11: Prendre une région de photo	43
Figure 3.12: Classification des images	44
Figure 3.13: la résultat de classification	44

Liste des Abbreviation

PIR : le proche infrarouge

MIR : le moyen infrarouge

MNT : Modèle Numérique de Terrain

SIG: Système d'information géographique

AI : Artificielle Intelligence

ML : Machine learning

NTIC : nouvelles technologies de l'information et de la communication

CNN : Convolutional Neural Network

ReLU : abréviation de Unités Rectifié linéaires

MLP: Multilayer perceptron

RNN: Recurrent Neural Network

TDNN : réseau neuronal à retard temporel

GPU: Graphics processing unit

API: Application programming interface

NFC :Near Field Communication

Introduction générale

Avant l'ère spatiale, l'homme n'avait jamais pu embrasser l'ensemble d'un hémisphère d'un seul coup d'œil. Jamais, en fait, il n'avait pas pu avoir une vision globale du monde qu'il habite.

Mais de l'espace, on peut observer la terre et utiliser ces observations. Donc il a fallu la mise en orbite des véhicules spatiaux comme les satellites qui portent des instruments (Capteurs) en orbite.

autour de la terre. Ces instruments nous permettent d'observer tout le monde entier. Ils balaient la surface de la terre et enregistrent des mesures d'énergie qui sont ensuite utilisées par des programmes informatiques pour créer des images.

Les images satellitaires, ou images de télédétection sont des images prises à partir d'un satellite qui orbite bien au-dessus de la terre. On appelle également les images satellitaires des images Remote est le terme anglais qui signifie à distance ou à grande distance et Sensing signifie découvrir ou balayer.

La télédétection est l'ensemble des connaissances et techniques utilisées pour déterminer des caractéristiques physiques et biologiques des objets par des mesures effectuées à distance, sans contact matériel avec ceux-ci. Les caractéristiques d'un objet peuvent être retrouvées à sa réflectivité spectrale unique (est souvent appelée signature spectrale) basée sur ses propriétés telles que sa composition chimique et physique, sa température, sa texture, etc.

L'objectif de la télédétection est de tirer partie des informations contenues dans la signature spectrale pour identifier les objets observés. De ce point de vue, la télédétection a montré son potentiel pour l'acquisition des données d'un premier lieu et l'extraction d'informations nécessaires pour la modélisation d'un environnement en second lieu.

Ce que l'on appelle image hyperspectrale est en réalité toute une série d'images de la même scène, mais prises de plusieurs dizaines de longueurs d'onde. Une telle mesure fournit de nombreuses informations sur les propriétés physiques des objets observés. Bien que pour l'analyse d'images hyperspectrale, la connaissance du système et des mécanismes de l'imagerie soit essentielle. Cette connaissance nécessite parfois une étude détaillée et précise qui prenne en compte différents éléments physiques du capteur ainsi que les conditions atmosphériques. C'est pourquoi nous préférons bâtir un modèle général comportant tous les

Introduction générale

paramètres de l'imagerie. À partir de ce modèle, nous sommes capables de détecter, classifier et d'identifier les différentes classes existant dans la scène.

Pour des fins de classifier une image hyperspectrale, une stratégie supervisée suppose que nous possédons des informations à priori sur les zones d'intérêt. Dans ce mémoire, nous allons appliquer une approche neuronale en tant que la méthode supervisée, en vue de définir les différentes classes qui appartiennent à cette image.

L'apprentissage profonde s'appuie sur un réseau de neurones artificiels inspirés du cerveau humain. Ce réseau est constitué de dizaines et de centaines de "couches" de neurones, qui reçoivent et traduisent chacune les informations de la couche précédente. Par exemple, le système apprendra à reconnaître les caractères avant de traiter des mots dans du texte ou une image de la surface de la terre, ou à déterminer s'il existe un visage sur une image avant que quiconque ne soit découvert.

L'objectif principale de notre travail est de classifier des images satellitaires en ensemble de classe s en utilisant les techniques de l'apprentissage profond.

CHAPITRE : 1

Téledétection

I. Introduction:

La télédétection est une technique qui permet, à l'aide d'un capteur, "d'observer" et d'enregistrer le rayonnement électromagnétique, émis ou réfléchi, par une cible quelconque sans contact direct avec celle-ci. Le traitement et l'analyse des informations véhiculées par le rayonnement enregistré permettent d'accéder à certaines propriétés de cette cible : géométriques (position, forme et dimensions), optiques (réflexion, transmission, absorption, etc.) et physicochimiques (température, teneur en eau, chlorophylle foliaire, phyto-masse, matière organique du sol,...), etc.

II. La télédétection :

II.1. Définition de la télédétection :

La télédétection est l'ensemble des techniques qui permettent, par l'acquisition d'images, D'obtenir de l'information sur la surface de la Terre (y compris l'atmosphère et les océans), Sans contact direct avec celle-ci. La télédétection 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.

II.2. Les différentes étapes de la télédétection :

Le processus de la télédétection se résume dans les étapes suivantes :

II.2.1. Source d'énergie ou d'illumination (A):

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

II.2.2. 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 fait dans le trajet entre la cible et le capteur.

II.2.3. 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.

II.2.4. 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 pour être enfin enregistrée.

II.2.5. Transmission, réception et traitement (E):

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

II.2.6. Interprétation et analyse (F):

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

II.2.7. Application(G) :

Lorsque l'on exploite l'imagerie satellitaire pour observer certaines caractéristiques ou certains événements particuliers, on peut parler d'applications du satellite. [Sma, 2004]

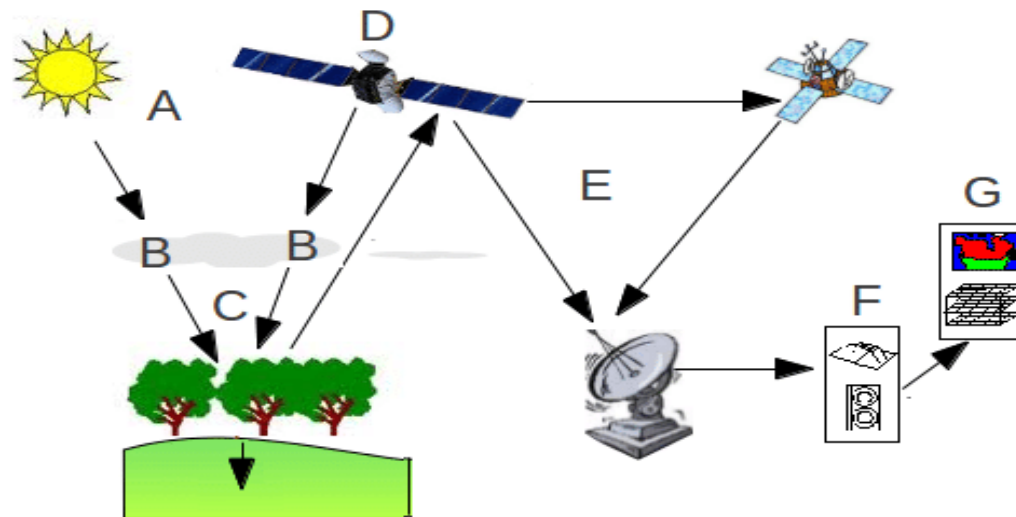


Figure 1.1: Les étapes de télédétection

III. Rayonnement électromagnétique :

III.1. Définition :

Un rayonnement, c'est de l'énergie qui se propage dans l'espace ou dans la matière sous forme de champs électriques et magnétiques perpendiculaires entre eux. Le rayonnement électromagnétique

est un élément important dans la télédétection car il est le seul moyen porteur de l'information sur les objets détectés. Selon la théorie des ondes, tout rayonnement électromagnétique possède des propriétés fondamentales. Le rayonnement électromagnétique est composé d'un champ électrique (E) et d'un champ magnétique (M). Le champ électrique varie en grandeur et est orienté de façon perpendiculaire à la direction de propagation du rayonnement. Le champ magnétique est orienté de façon perpendiculaire au champ électrique. Les deux champs se déplacent à la vitesse de la lumière (c).

III.1.2. Le spectre électromagnétique :

Le spectre électromagnétique est le résultat de la décomposition du rayonnement électromagnétique en ses fréquences constituantes.

En Télédétection, les capteurs peuvent acquérir des informations que l'œil humain ne peut pas voir normalement (en utilisant des rayonnements appartenant à des parties du spectre électromagnétique autres que le visible).

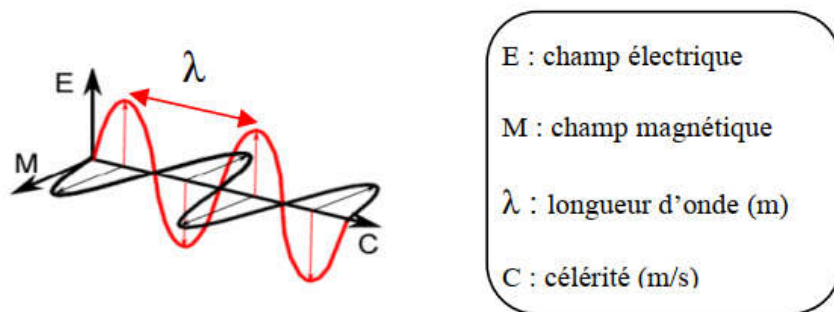


Figure 1.2 : Rayonnement électromagnétique

Les objets reflètent normalement une partie de la lumière qui les atteint.

C'est généralement cette portion de lumière qui donne leurs couleurs aux objets.

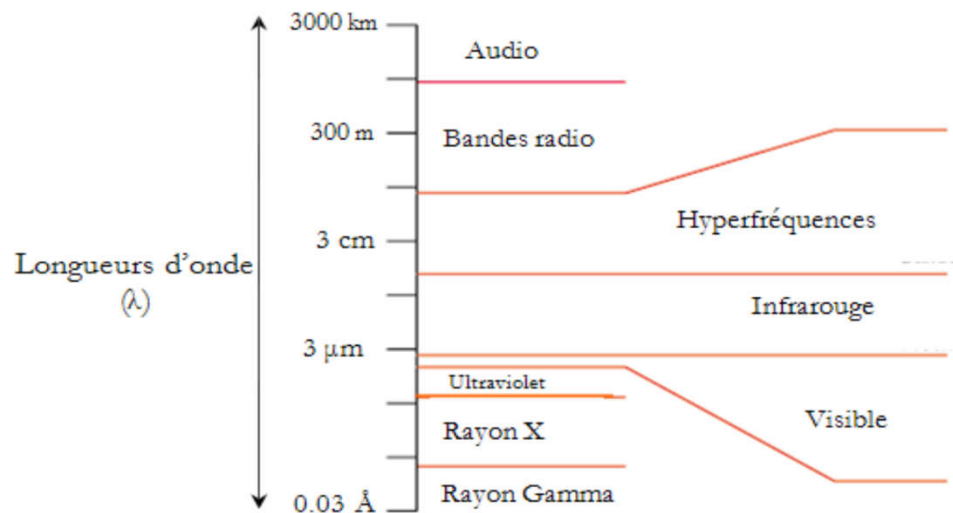


Figure 1.3: Domaines spectraux

Certains objets ne se contentent pas de refléter la lumière qui les atteint, ils émettent en outre des “rayonnements”, surtout à cause de leur température. La lumière est émise et réfléchi par les objets sous forme de rayonnement.

Les ondes électromagnétiques sont caractérisées par différentes longueurs d'ondes.

Pour mesurer le rayonnement émis ou réfléchi par les objets, on mesure leur énergie à différentes longueurs d'ondes.

En télédétection on utilise plusieurs régions du spectre électromagnétique, parmi celles-ci on peut citer : l'ultraviolet, le visible, l'infrarouge, les hyperfréquences . [Sma, 2004]

III.1.3. Propriétés des ondes électromagnétiques :

Avant que le rayonnement utilisé pour la télédétection n'atteigne la surface de la terre, celui-ci doit traverser une certaine épaisseur d'atmosphère. 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 [Ouf, 1998][W, 1]

A- 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 (ou extinction) dans la direction de propagation de l'onde.

B- Emission : C'est la production REM par un corps qui a une température supérieure à zéro degrés Kelvin, c'est à dire 273 °C. 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). [Rab, 2014]

$$\text{L'absorption}(\alpha) = \frac{\text{énergie absorbée}}{\text{énergie reçue}} \quad \alpha \in]0; 1[$$

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 .

$$\text{réflectance}(\rho) = \frac{\text{énergie réfléchi}}{\text{énergie reçue}} \quad \rho \in]0; 1[$$

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. [Rab, 2012]

$$\text{Transmistance}(\tau) = \frac{\text{énergie transmise}}{\text{énergie reçue}} \quad \tau \in]0; 1[$$

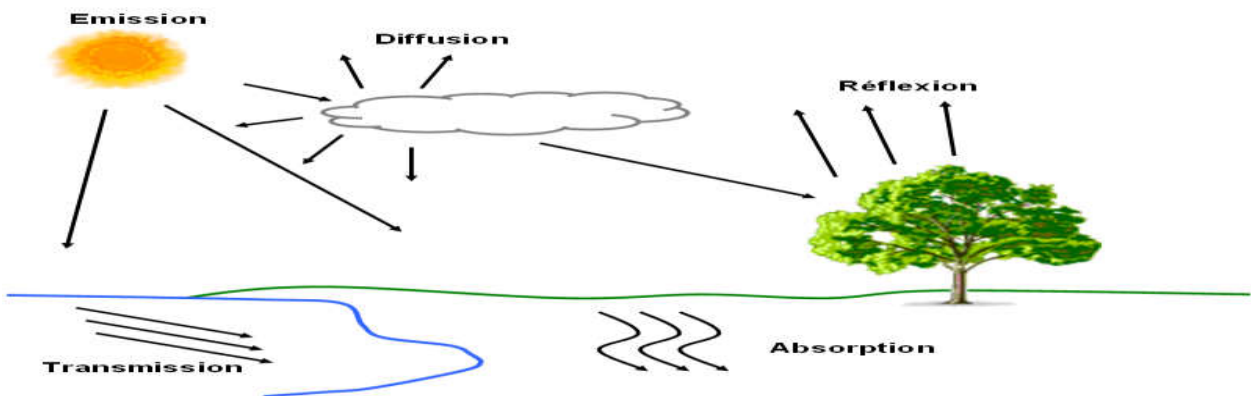


Figure 1.4: Propriété des ondes électromagnétiques

IV. Système d'acquisition :

Le système d'acquisition comporte deux éléments principaux, le premier est le capteur, et le deuxième c'est la plate forme. Cette dernière permet de véhiculer le capteur. [Han, 2010]

IV.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 :

- **Capteurs passifs** : Ils sont en général les capteurs qui utilisent une source d'énergie externe pour observer les objets.
- **Capteurs actifs** : On les appelle capteurs actifs car ils émettent eux-mêmes les rayonnements pour « illuminer » les objets de manière à ce que l'énergie réfléchie puisse être mesurée.

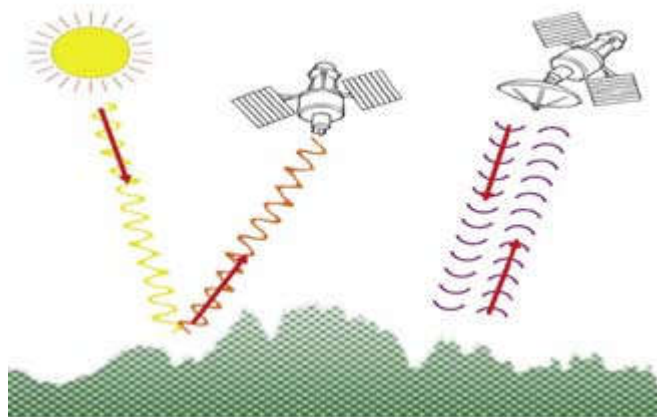


Figure 1.5: Capteur passif et actif

IV.2. Les plates formes :

En l'absence de moyen de nous éloigner de la surface de la terre la télédétection est une technologie relativement nouvelle, car quel que soit le degré de sophistication des instruments utilisés, elle ne permet pas d'acquérir une vision d'ensemble acceptable de notre planète. Parmi les plates formes les plus utilisées, on a les avions et les satellites.

- **Les avions :** Les avions volent à une altitude relativement faible (quelques kilomètres seulement) ils ne peuvent par conséquent prendre en photo (du point de vue de la télédétection) des portions de territoire limitées avec leurs détails.
- **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. [Han, 2010]

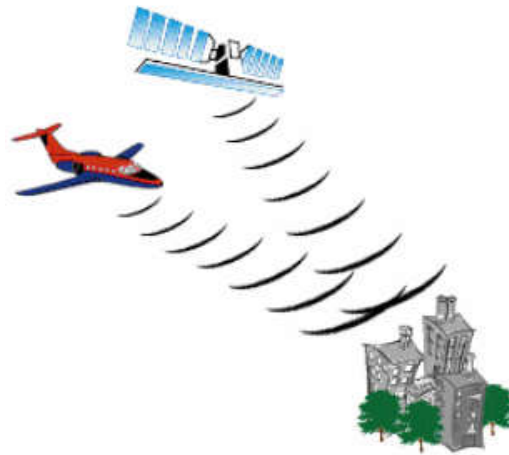


Figure 1.6: Plates formes (avion, satellite)

Remarque:

Deux grands avantages des satellites par rapport aux avions : c'est qu'ils peuvent prendre des images d'étendues extrêmement vastes et que la même surface peut être observée systématiquement à chaque passage du satellite. [Han, 2010]

Suivant les différentes missions des satellites nous pouvons citer :

- Les satellites d'astronomie.
- Les satellites de navigation.
- Les satellites météorologiques.
- Les satellites militaires.
- Les satellites de communication. [Rab, 2014]

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

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, en décrivant une orbite circulaire dans le plan équatorial vers 35790 Km d'altitude. [Sma, 2004]

V. Imagerie satellitaire:

V.1. Introduction:

L'image a deux formats, analogique ou numérique. Dans le premier cas les données sont stockées d'une manière continue, et dans le deuxième, chaque bloque d'information est enregistré d'une façon discrète.

Le format numérique est basé sur une procédure mathématique appelée système binaire (1 et 0). Où 1 représente l'existence d'une impulsion électrique, et 0 en l'absence cette impulsion. Une image satellitaire est une image numérique. La particularité de cette image est quelle est prise par un satellite placé en orbite autour de la planète. Elle est aussi synthétique, flexible, précise et actuelle (renouvellement de l'information automatique ou à la demande). [Han, 2010]

V.2. Définitions:

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. [LANDSAT 5 –TM]

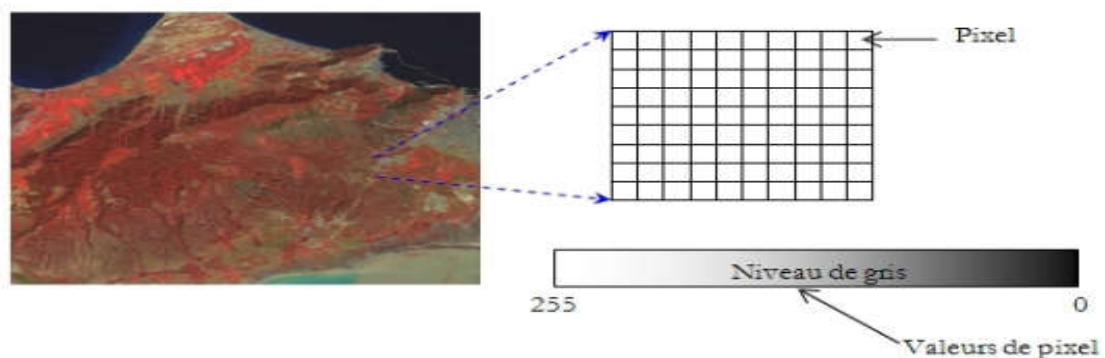


Figure 1.7: Image satellitaire « Oran » (LANDSAT 5 -TM)

Afin d'obtenir des informations différentes, les capteurs observent la terre dans plusieurs bandes de longueurs d'onde, encore appelées "canaux". On dit qu'ils sont multi spectraux. A chaque pixel sont associées plusieurs valeurs radiométriques comprises entre 0 et 255. Mathématiquement c'est une fonction $F(x, y)$ à deux variables entières x et y qui sont respectivement la ligne et la colonne, fournissant une ou plusieurs valeurs entières représentant le niveau de gris du pixel dans les différents canaux. [Han, 2010]

V.2.1. Résolution (spatiale) :

Le premier fait important à connaître concernant une image satellite est sa résolution spatiale. La résolution spatiale d'une image est la plus petite distance entre deux objets adjacents que le capteur puisse identifier.

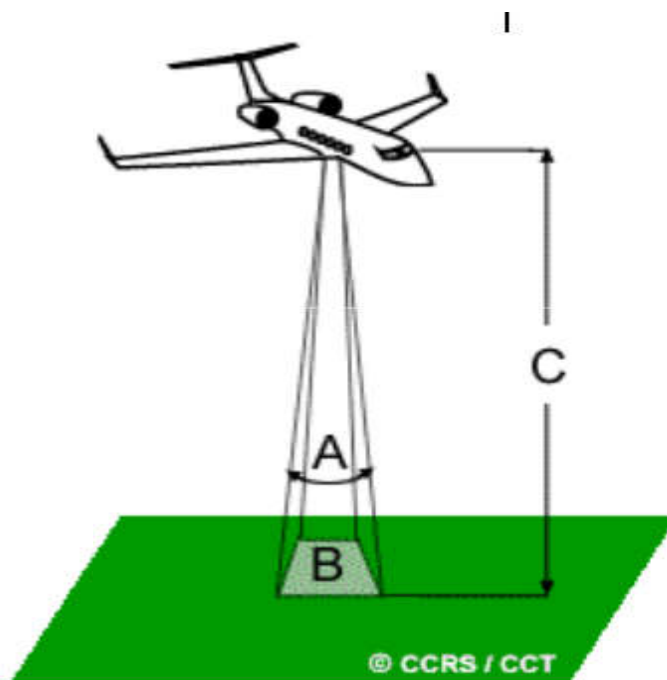


Figure 1.8: Résolution spatiale

V.2.2. 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.

V.3. Les paramètres influant sur le contenu d'une image satellitaire :

Le contenu des images de télédétection est fortement lié à certains paramètres qui dépendent des caractéristiques du satellite. Il s'agit notamment de :

V.3.1. La définition de l'image :

La taille des phénomènes étudiés doit être nettement plus grande que celle des pixels, d'où l'importance de la résolution spatiale du satellite. [Han, 2010]

V.3.2. Les canaux disponibles:

Les valeurs radiométriques d'un pixel dépendent de ses propriétés optiques. Ainsi, dans le cas d'un pixel de végétation, la réponse dans les canaux du visible (vert et rouge pour SPOT) est fonction de la teneur en pigments foliaires des végétaux alors qu'elle dépend surtout de la structure des feuilles dans le proche infrarouge (PIR) et de la teneur en eau dans le moyen infrarouge (MIR).

V.3.3. La date de prise de vue:

Elle doit être en relation avec les thèmes à extraire et de ce fait nécessite une bonne connaissance du terrain. Ce paramètre est d'autant plus important que le thème étudié évolue rapidement. C'est par exemple le cas en agriculture.

V.4. Les prétraitements:

L'étape de prétraitement a pour but de faciliter la segmentation en renforçant la ressemblance entre pixels appartenant à une même région, ou en accentuant la dissemblance entre pixels appartenant à des régions différentes.

V.4.1. Modification d'histogramme:

Le but de cette étape est d'améliorer l'image, pour cela on applique une transformation ponctuelle d'intensité. A tout pixel d'intensité as on associe une intensité $a=T(as)$. Du fait de leur caractère ponctuel, les méthodes de transformation d'histogramme n'affectent pas la forme des régions. Elles en modifient uniquement l'apparence visuelle.

V.4.2. Réduction de bruit:

Le bruit constitue des perturbations qui affectent la qualité de la segmentation et sont généralement désignées sous le terme de bruit d'image.

Ce traitement a pour but de réduire l'amplitude des variations d'intensité dans chaque région, tout en conservant les transitions entre régions adjacentes.

V.4.3. Rehaussement de contraste:

Il arrive toutefois que la transition entre régions (constituant l'image) soit initialement floue, c'est-à-dire, qu'elle s'étale sur plusieurs pixels. Rehausser le contraste de l'image consiste à diminuer l'étendue de la zone de transition sans affecter l'intensité moyenne des régions situées de part et d'autre. limite le risque de fusion intempestive de régions distinctes lors de la phase de segmentation. [Coc, 1995]

V.5. Les traitements :

L'extraction d'informations peut se faire selon plusieurs méthodes, dont nous présenterons ici les deux plus répandues :

- L'approche statistique: d'après les seules valeurs radiométriques du pixel, qui débouche sur de nombreux procédés de seuillage et de classification. [Han, 2010]
- Les autres approches: qui prennent en compte non seulement la radiométrie individuelle d'un pixel mais également celle des pixels de son voisinage immédiat ou ailleurs dans l'image.

V.5.1. Les traitements statistiques :

Parmi les traitements statistiques on distingue:

- **Seuillage d'un histogramme monodimensionnel :**

A partir de l'histogramme d'un seul canal, cette technique consiste à ne retenir que les pixels dont la radiométrie est comprise entre deux valeurs. Elle permet cependant dans certains cas de séparer certains thèmes bien caractéristiques, ces thèmes se traduisant souvent sur l'histogramme par la présence de plusieurs « bosses ».

➤ **La classification :**

Les classifications consistent à découper le nuage de points en cubes, en sphères ou en ellipsoïdes, chacun regroupant les éléments de l'image appartenant à la même classe. Le but de l'opération consiste à réaliser le découpage optimal. Les méthodes de classification se divisent en 2 grands groupes : les méthodes supervisées et les méthodes non supervisées.

V.6. Buts de traitements :

Les intérêts sont nombreux en télédétection : on peut par exemple segmenter automatiquement l'image satellitaire si les limites recherchées (zones administratives, agroclimatiques, etc.) sont déjà stockées. On peut également introduire des critères de décision supplémentaires dans les processus de classification de l'image pour améliorer la thématique finale. Les avantages des traitements numériques se situent aux niveaux suivants :

- Meilleure visualisation
- Souplesse d'édition
- Facilités de cartographie
- Estimation de surface
- Réduction de dimensions Dans notre travail de recherche nous nous sommes intéressés à la segmentation des images satellitaires. [Han, 2010]

VI. Classification :

VI.1. Définitions:

La classification est un processus de reconnaissance des formes. En télédétection, elle consiste à effectuer la correspondance entre une scène de l'image matérialisée généralement par leurs valeurs radiométriques et des classes connues a priori ou non par l'utilisateur. La correspondance est réalisée par des fonctions discriminantes sous forme de règles de décision ou des distances géométriques. [Han, 2010]

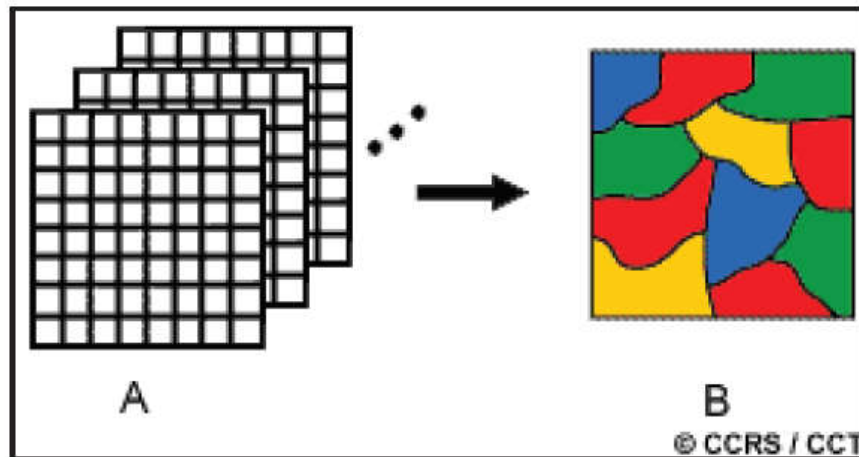


Figure 1.9: Principe de classification

VI.2. Méthodes de classification :

Les méthodes de classification proposées intègrent un grand nombre de données et de Concepts. L'application autour de la classification est une chaîne assez conséquente et complexe, seul les points les plus importants sont décrits. Le fonctionnement général de l'application de classification peut être résumée par la (figure1.10)

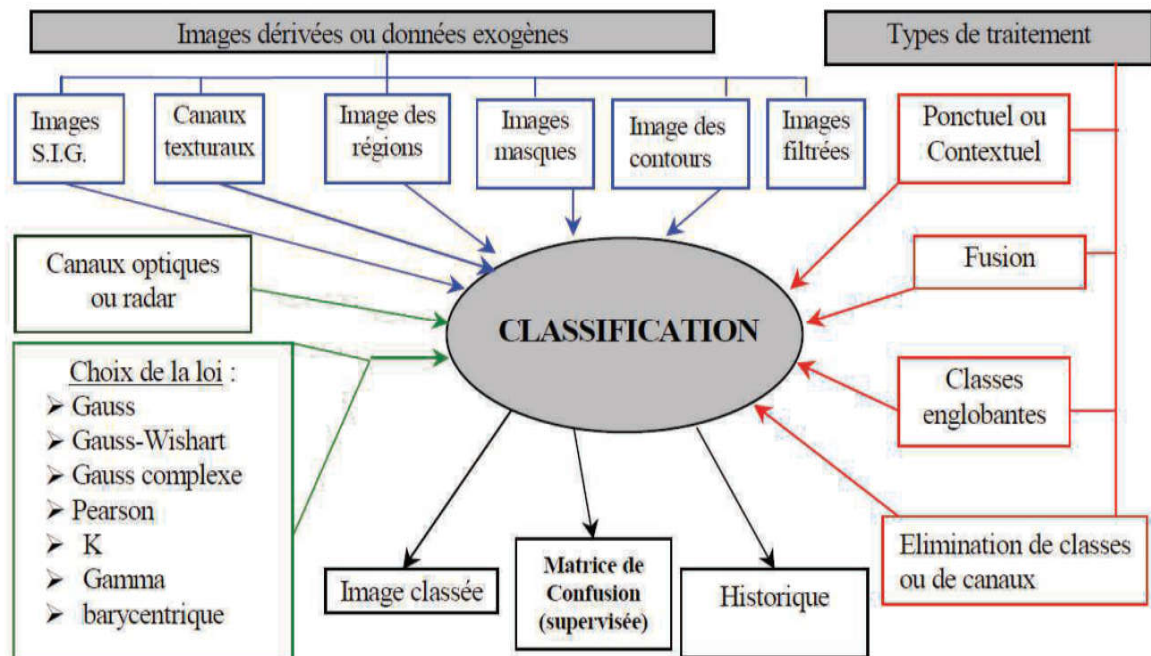


Figure 1.10: Représentation des méthodes de classification

Les images dérivées introduites dans la classification sont les :

VI.2.1. Images texturales : obtenues -1 par le calcul d'écart type ou de coefficient de variation sur les images sur un voisinage centré sur le pixel traité, -2 par les modules d'analyse du paysage.

VI.2.2 : Image des régions: pour introduire la connaissance de l'expert par l'emplacement de certaines classes dans certaine zone, afin de faciliter la détection des classes.

VI.2.3 : Image des contours : d'aiguiller la décision d'affectation du pixel à une classe en analysant son voisinage.

VI.2.4 : Image « masque »: permettent d'établir des zones de l'image à ne pas classer en raison du recalage des images, par exemple, ce qui permet de ne pas tenir compte des zones non commune de plusieurs dates.

VI.2.5: Image du MNT (Modèle Numérique de Terrain).

VI.2.6 : SIG (Système d'information géographique) : permettent de discriminer directement certaines classes à partir d'éléments contenus dans le SIG.

VI.2.7 : Image du réseau routier et/ou hydrographique.

VI.3. Techniques de classification:

Pour classifier des images de télédétection, plusieurs approches sont disponibles

1- Celles, supervisées, ou non-supervisée, basées sur le pixel peuvent être réalisée avec de nombreux logiciels d'analyse d'images comme Erdas IMAGINE et ENVI.

2- Celle basée sur l'objet qui peut être réalisée avec le logiciel Ecognition de la firme Definiens.

VI.3.1. Procédure de classification:

Les procédures de classification se catégorisent selon la situation dans laquelle l'opérateur est placé et selon l'objectif suivi. Deux catégories de démarches sont identifiées : la classification non dirigée et la classification dirigée.

VI.3.2. La classification non dirigée :

L'opérateur ne dispose d'aucune information sur la zone d'intérêt ou ne souhaite pas en exploiter dans la phase initiale d'analyse. Aucune information externe à l'image n'est donc introduite. Le

classement des pixels est effectué uniquement selon le critère de la ressemblance spectrale. L'étiquetage thématique est réalisé, s'il est souhaité, après la classification par interprétation des classes spectrales produites.

VI.3.3. La classification dirigée :

Elle pourrait aussi être dénommée classification par échantillonnage. L'opérateur a défini au préalable les classes thématiques qu'il souhaite extraire de l'image et identifié quelques zones de l'image occupées, sans ambiguïté, par les classes d'intérêt. De ces zones, il extrait les signatures spectrales pour chaque classe. Dans la pratique, lors d'une classification d'une image comportant de nombreuses catégories d'occupation du sol, l'analyste exploite souvent une combinaison des deux procédures. La démarche est itérative jusqu'à ce que les résultats soient conformes aux objectifs fixés par l'analyste.

VI.4. Marche à suivre pour une classification supervisée:

La marche à suivre pour procéder à une classification dirigée comporte 10 étapes formulées ici selon une adaptation de celle proposée par Bonn et Rochon [Bon, 1992]:

1. Formuler avec précision les classes thématiques à extraire.
2. Analyser et évaluer l'existence de leur correspondance spectrale.
3. Sélectionner les zones d'échantillonnage (zones considérées représentatives des classes à déterminer).
4. Diviser en deux groupes les zones d'échantillonnage : un groupe pour l'établissement des signatures spectrales (zones d'entraînement), et un autre pour vérifier si la procédure produit bien les classes attendues (zones tests).
5. Établir les signatures spectrales.
6. Analyser les signatures spectrales. Sont-elles disjointes dans l'espace spectral ? Évaluer leur pouvoir discriminant.
7. Choisir l'algorithme de classification.

8. Effectuer une classification des zones d'entraînement. Établir et analyser la matrice de confusion. Cette première évaluation ne fait qu'orienter l'analyste sur les capacités de la procédure et les difficultés de discrimination entre classes.

9. Effectuer une classification sur les zones tests. Établir et analyser la matrice de confusion. Si les résultats ne sont pas satisfaisants, modifier les zones d'entraînement, pour améliorer leur représentativité ou les classes thématiques, ou encore changer la procédure de classification. Cette étape est cruciale, car elle est à la base de l'évaluation de la qualité de l'image classée.

10. Classer l'ensemble de l'image.

VI.4.1 : Marche à suivre pour une classification non supervisée :

1- Les données sont classées en fonction de leurs caractéristiques spectrales, sans aucune information a priori sur la nature des objets à classer. L'interprète intervient pour fixer un nombre de groupes, des seuils, ... (en anglais non-supervised classifications).

2- Le principe de la classification non-supervisée est d'attribuer automatiquement chaque pixel à une classe suivant sa signature spectrale. L'algorithme le plus utilisé est le cluster ISODATA. Cet algorithme utilise de multiples itérations.

3- Tout d'abord, des clusters sont arbitrairement construits. Ensuite, les pixels sont attribués au cluster le plus proche. Puis, la moyenne de chaque cluster est calculée sur la base des nouveaux pixels assignés au cluster. Les clusters sont regroupés si le centre de 2 clusters est plus proche qu'un certain seuil, ou si le nombre de pixels dans une classe est trop petit. Les clusters sont séparés dans le cas où l'écart type dépasse une certaine valeur, ou si le nombre de pixels est 2 fois supérieur au seuil du nombre minimal de pixels. La seconde et troisième étape sont répétées jusqu'à ce que les changements entre les itérations soient moindres. [Han, 2010]

VII. Conclusion :

La télédétection utilise une technologie de pointe. Grâce au satellite ultra sophistiqué et au matériel rigoureusement utilisé la télédétection a tendance à se développer d'avantage. Elle a investi des domaines scientifiques divers. Dans les domaines des sciences spatiales et surtout en ce qui concerne le traitement de l'imagerie satellitaire la télédétection est un moyen scientifique d'avant-garde qui permet de donner des informations de plus en plus précises que précieuses sur les objets, les surfaces et les profondeurs les plus inexplorés de notre univers. Les chapitres qui suivent

s'appuient sur l'aspect théorique des techniques qui seront utilisées pour la classification de l'image de télédétection.

CHAPITRE : 2

Deep learning
(L'apprentissage profond)

I. Introduction :

L'intelligence artificielle est une discipline scientifique recherchant des méthodes de solution de problèmes à forte complexité logique ou algorithmique. L'apprentissage automatique champ d'étude de l'intelligence artificielle. Par conséquent, L'apprentissage profond (en anglais deep learning, deep structured learning, hierarchical learning) est un ensemble de méthodes d'apprentissage automatique tentant de modéliser avec un haut niveau d'abstraction des données grâce à des architectures articulées de différentes transformations non linéaires. Dans ce chapitre nous allons présenter tout d'abord les notions en relation avec l'apprentissage profond.

Le Deep Learning est un nouveau domaine de recherche du ML, qui a été introduit dans le but de rapprocher le ML de son objectif principal : l'intelligence artificielle. Il concerne les algorithmes inspirés par la structure et le fonctionnement du cerveau. Ils peuvent apprendre plusieurs niveaux de représentation dans le but de modéliser des relations complexes entre les données.

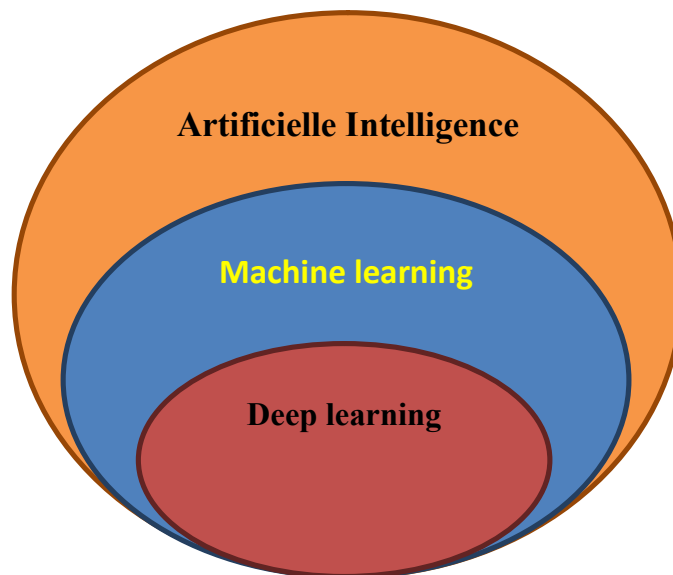


Figure 2.1: La relation entre l'intelligence artificielle, le ML et le deep learning

II. Définition de l'apprentissage profond (deep learning):

L'apprentissage profond (« *deep learning* ») est un ensemble de techniques d'apprentissage automatique qui a permis des avancées importantes en intelligence artificielle dans les dernières années.

Dans l'apprentissage automatique, un programme analyse un ensemble de données afin de tirer des règles qui permettront de tirer des conclusions sur de nouvelles données. L'apprentissage

profond est basé sur ce qui a été appelé, par analogie, des « réseaux de neurones artificiels », composés de milliers d'unités (les « neurones ») qui effectuent chacune de petites opérations simples. Les résultats d'une première couche de « neurones » servent d'entrée aux calculs d'une deuxième couche et ainsi de suite.

Par exemple, pour la reconnaissance visuelle, des premières couches d'unités identifient des lignes, des courbes, des angles... des couches supérieures identifient des formes, des combinaisons de formes, des objets, des contextes...

Les progrès de l'apprentissage profond ont été possibles notamment grâce à l'augmentation de la puissance des ordinateurs et au développement de grandes bases de données (« *big data* »).[W2]

III. Pour quoi le deep learning ?

Les algorithmes de ML décrits dans la première partie fonctionnent bien pour une grande variété de problèmes. Cependant ils ont échoués à résoudre quelques problèmes majeurs de l'IA telle que la reconnaissance vocale et la reconnaissance d'objets.

Le développement du deep learning fut motivé en partie par l'échec des algorithmes traditionnels dans de telle tâche de l'IA.

Mais ce n'est qu'après que de plus grandes quantités de données ne soit disponibles grâce notamment au Big Data et aux objets connectés et que les machines de calcul soient devenues plus puissantes qu'on a pu comprendre le potentiel réel du Deep Learning.

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.

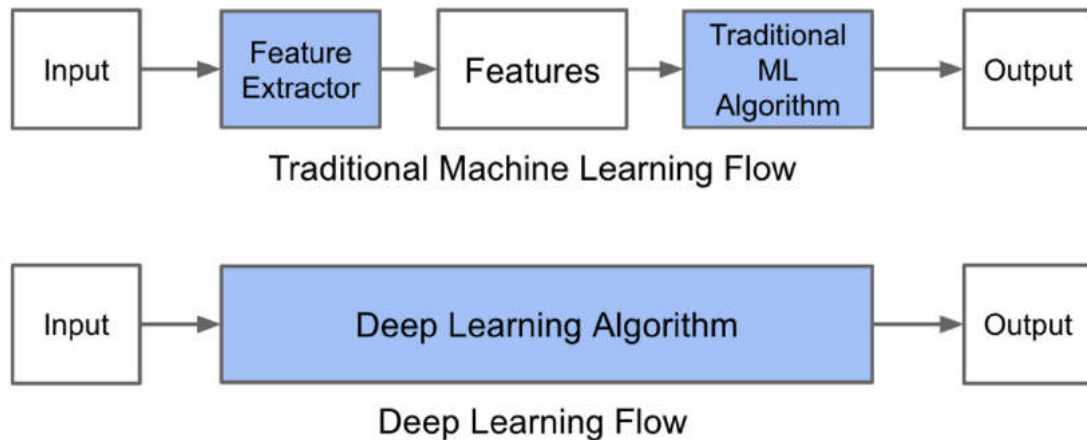


Figure 2.2 : Le procédé du ML classique comparé à celui du Deep Learning

III.1- Domaines d'application de l'apprentissage profonde :

Ces techniques se développent dans le domaine de l'informatique appliquée aux NTIC (reconnaissance visuelle — par exemple d'un panneau de signalisation par un robot ou une voiture autonome — et vocale notamment) à la robotique, à la bio-informatique, la reconnaissance ou comparaison de formes, la sécurité, la santé, etc..., la pédagogie assistée par l'informatique, et plus généralement à l'intelligence artificielle. L'apprentissage profond peut par exemple permettre à un ordinateur de mieux reconnaître des objets hautement déformables et/ou analyser par exemple les émotions révélées par un visage photographié ou filmé, ou analyser les mouvements et position des doigts d'une main, ce qui peut être utile pour traduire le langage des signes, améliorer le positionnement automatique d'une caméra, etc... Elles sont utilisées pour certaines formes d'aide au diagnostic médical (ex. : reconnaissance automatique d'un cancer en imagerie médicale), ou de prospective ou de prédiction (ex. : prédiction des propriétés d'un sol filmé par un robot)

III.2- Architectures de réseaux de neurones profonds :

Il existe un grand nombre de variantes d'architectures profondes. La plupart d'entre elles sont dérivées de certaines architectures parentales originales. Il n'est pas toujours possible de comparer les performances de toutes les architectures, car elles ne sont pas toutes évaluées sur les mêmes ensembles de données. Le Deep Learning est un domaine à croissance rapide, et de nouvelles architectures, variantes ou algorithmes apparaissent toutes les semaines. [MMZ, 2017]

III.2.1. Les réseaux de neurones convolutifs :

Les réseaux de neurones convolutifs sont à ce jour les modèles les plus performants pour classer des images. Désignés par l'acronyme CNN, de l'anglais Convolutional Neural Network, ils

comportent deux parties bien distinctes. En entrée, une image est fournie sous la forme d'une matrice de pixels. Elle a deux dimensions pour une image aux niveaux de gris.

La couleur est représentée par une troisième dimension, de profondeur 3 pour représenter les couleurs fondamentales [Rouge, Vert, Bleu].

La première partie d'un CNN est la partie convolutive à proprement parler. Elle fonctionne comme un extracteur de caractéristiques des images. Une image est passée à travers d'une succession de filtres, ou noyaux de convolution, créant de nouvelles images appelées cartes de convolutions. Certains filtres intermédiaires réduisent la résolution de l'image par une opération de maximum local. En fin, les cartes de convolutions sont mises à plat et concaténées en un vecteur de caractéristiques, appelé code CNN.

Ce code CNN en sortie de la partie convolutive est ensuite branché en entrée d'une deuxième partie, constituée de couches entièrement connectées. Le rôle de cette partie est de combiner les caractéristiques du code CNN pour classer l'image.

La sortie est une dernière couche comportant un neurone par catégorie. Les valeurs numériques obtenues sont généralement normalisées entre 0 et 1, de somme 1, pour produire une distribution de probabilité sur les catégories. [MMZ, 2017]

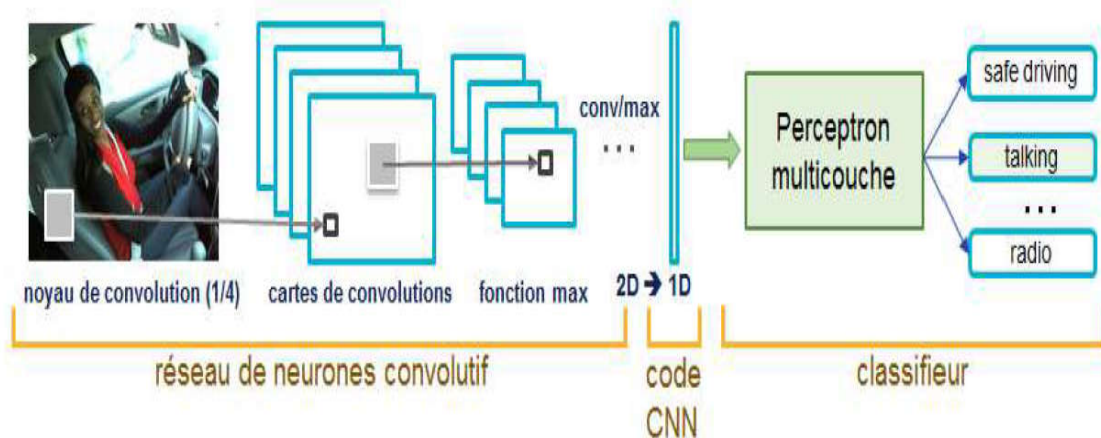


Figure 2.3 : Les réseaux de neurones convolutifs.

III.2.2. L'entraînement d'un nouveau CNN est difficile :

Une architecture CNN est formée par un empilement de couches de traitement indépendants :

III.2.2.1. La couche de convolution (CONV) :

Trois hyper paramètres permettent de dimensionner le volume de la couche de convolution

(aussi appelé volume de sortie) : la 'profondeur', le 'pas' et la 'marge'.

1. 'Profondeur' de la couche : nombre de noyaux de convolution (ou nombre de neurones associés à un même champ récepteur).
2. 'Le pas' contrôle le chevauchement des champs récepteurs. Plus le pas est petit, plus les champs récepteurs se chevauchent et plus le volume de sortie sera grand.
3. 'La marge (à 0)' ou 'zéro padding' : parfois, il est commode de mettre des zéros à la frontière du volume d'entrée. La taille de ce 'zero-padding' est le troisième hyperparamètre. Cette marge permet de contrôler la dimension spatiale du volume de sortie. En particulier, il est parfois souhaitable de conserver la même surface que celle du volume d'entrée (11).

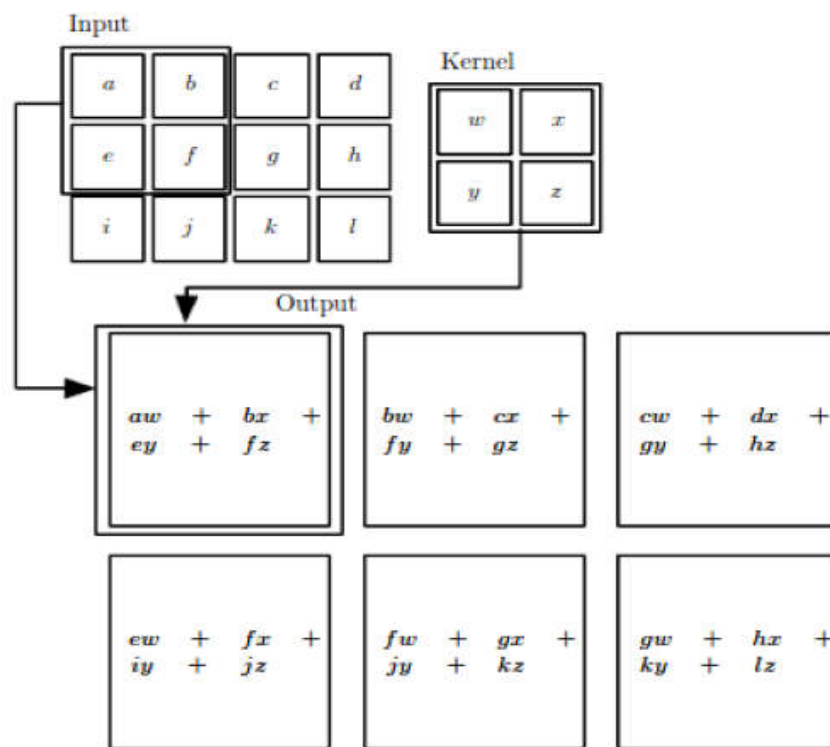


Figure 2.4 : Exemple d'une convolution 2D.

III.2.2.2 Couche de pooling (POOL) :

Une architecture atypique d'un réseau convolutif se compose de trois types de couches différentes. D'abord une couche convolutive pour générer un ensemble d'activations linéaires ensuite, on les fait passer à travers une couche d'activation non linéaire telle que Rectified Linear Unit (ReLU), enfin on utilise la fonction pooling.

- _ Il permet de réduire progressivement la taille des représentations afin de réduire la quantité de paramètres et de calcul dans le réseau et, par conséquent, de contrôler également le surapprentissage. [MMZ, 2017]
- _ Il permet l'invariance aux petites translations
- _ Utile lorsque on préfère savoir si une caractéristique est présente plutôt que la région de sa présence.
- _ Plusieurs type de pooling différent (MAX pooling (très populaire), AVG pooling, ...)

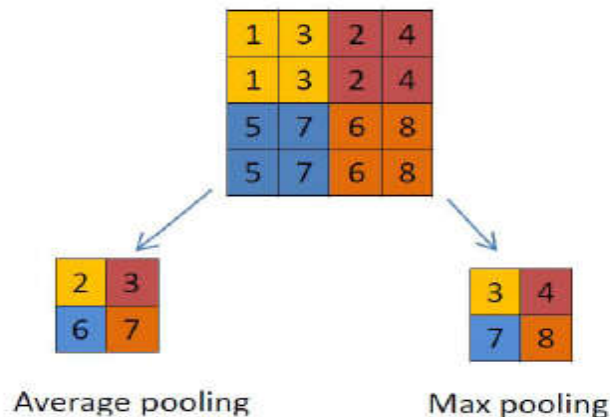


Figure 2.5 : (à gauche) Average pooling même couleur, (à droite) Max pooling : chaque case correspond

III.2.2.3. Couche de correction (ReLU):

Souvent, il est possible d'améliorer l'efficacité du traitement en intercalant entre les couches de traitement une couche qui va opérer une fonction mathématique (fonction d'activation) sur les signaux de sortie. On a notamment :

➤ La correction ReLU (abréviation de Unités Rectifié linéaires) : $f(x) = \max(0, x)$. Cette fonction, appelée aussi « fonction d'activation non saturante », augmente les propriétés non linéaires de la fonction de décision et de l'ensemble du réseau sans affecter les champs récepteurs de la couche de convolution.

- La correction par tangente hyperbolique $f(x) = \tanh(x)$.
- La correction par la tangente hyperbolique saturante : $f(x) = |\tanh(x)|$.
- La correction par la fonction sigmoïde.

Souvent, la correction Relu est préférable, car il en résulte la formation de réseau neuronal plusieurs fois plus rapide, sans faire une différence significative à la généralisation de précision.

III.2.2.4 Couche entièrement connectée(FC):

Après plusieurs couches de convolution et de max-pooling, le raisonnement de haut niveau dans le réseau neuronal se fait via des couches entièrement connectées. Les neurones dans une couche entièrement connectée ont des connexions vers toutes les sorties de la couche précédente (comme on le voit régulièrement dans les réseaux réguliers de neurones). Leurs fonctions d'activations peuvent donc être calculées avec une multiplication matricielle suivie d'un décalage de polarisation.

III.2.2.5 Couche de perte (LOSS) :

La couche de perte spécifie comment l'entraînement du réseau pénalise l'écart entre le signal prévu et réel. Elle est normalement la dernière couche dans le réseau. Diverses fonctions de perte adaptées à différentes tâches peuvent y être utilisées. La perte « Soft max » est utilisée pour prédire une seule classe parmi K classes mutuellement exclusives. La perte par entropie croisée sigmoïde est utilisée pour prédire K valeurs de probabilité indépendante dans [0,1] . La perte euclidienne est utilisée pour régresser vers des valeurs réelles. [MMZ, 2017]

III.2.2.6 Perceptron:

Après avoir extrait les caractéristiques des entrées, on attache à la fin du réseau un perceptron ou bien un MLP. Le perceptron prend comme entrée les caractéristiques extraites et produit un vecteur de N dimensions où N est le nombre de classe ou chaque élément est la probabilité d'appartenance à une classe. chaque probabilité est calculée à l'aide de la fonction softmax dans le cas où les classes sont exclusivement mutuelles.

$$\text{softmax}(z)_i = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

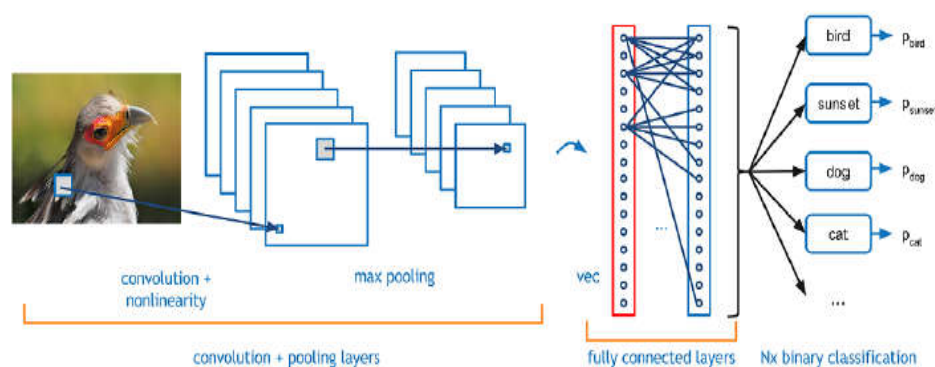


Figure2.6: Un réseau de neurones convolutif qui reçoit une image 2D comme entrée et qui est composé d'une couche convolutive, une fonction d'activation non linéaire, une couche MAX pooling et enfin un perceptron multi Couche

III.2.2.7- Choix des hyperparamètres:

Les CNNs utilisent plus d'hyperparamètres qu'un MLP standard. Même si les règles habituelles pour les taux d'apprentissage et des constantes de régularisation s'appliquent toujours, il faut prendre en considération les notions de nombre de filtres, leur forme et la forme du max pooling.

III.2.2.7.1- Nombre de filtres :

Comme la taille des images intermédiaires diminue avec la profondeur du traitement, les couches proches de l'entrée ont tendance à avoir moins de filtres tandis que les couches plus proches de la sortie peuvent en avoir davantage. Pour égaliser le calcul à chaque couche, le produit du nombre de caractéristiques et le nombre de pixels traités est généralement choisi pour être à peu près constant à travers les couches. Pour préserver l'information en entrée, il faudrait maintenir le nombre de sorties intermédiaires (nombre d'images intermédiaire multiplié par le nombre de positions de pixel) pour être croissante (au sens large) d'une couche à l'autre.

Le nombre d'images intermédiaires contrôle directement la puissance du système, dépend du nombre d'exemples disponibles et la complexité du traitement.

III.2.2.7.2- Forme de filtres:

Les formes de filtre varient grandement dans la littérature. Ils sont généralement choisis en fonction de l'ensemble de données. Les meilleurs résultats sur les images de MNIST (28x28) sont habituellement dans la gamme de 5x5 sur la première couche, tandis que les ensembles de données d'images naturelles (souvent avec des centaines de pixels dans chaque dimension) ont tendance à utiliser de plus grands filtres de première couche de 12x12, voire 15x15.

Le défi est donc de trouver le bon niveau de granularité de manière à créer des abstractions à l'échelle appropriée et adaptée à chaque cas.

III.2.2.7.3- Forme de Max pooling:

Les valeurs typiques sont 2x2. De très grands volumes d'entrée peuvent justifier un pooling 4x4 dans les premières couches. Cependant, le choix de formes plus grandes va considérablement réduire la dimension du signal, et peut entraîner la perte de trop d'information.

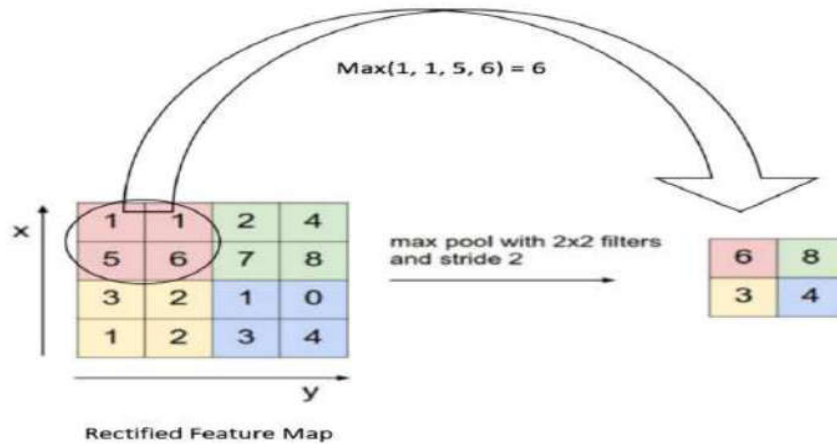


Figure 2.7 : Exemple de max pooling (2×2).

III.2.3. Réseau de neurones récurrents:

Les humains ne commencent pas leur pensées à zéro à chaque seconde. Lorsqu'on lit un livre, on comprend chaque mot en fonction de la compréhension des mots précédents. On n'oublie pas tout et recommençons à réfléchir à nouveau. Nos pensées ont une persistance.

RNN (les réseaux de neurones récurrents) traitent ce problème. Ce sont des réseaux avec des boucles, permettant aux informations de persister.

III.2.3.1. C'est quoi un RNN?:

L'idée derrière les RNN (**Recurrent Neural Network**) est d'utiliser des informations séquentielles. Dans un réseau neuronal traditionnel, nous supposons que toutes les entrées (et les sorties) sont indépendantes les unes des autres. Mais pour de nombreuses tâches, c'est une très mauvaise idée. Si on veut prédire le prochain mot dans une phrase, il faut connaître les mots qui sont venus avant. Les RNN sont appelés récurrents car ils exécutent la même tâche pour chaque élément d'une séquence, la sortie étant dépendante des calculs précédents. Une autre façon de penser les RNN est qu'ils ont une « mémoire » qui capture l'information sur ce qui a été calculé jusqu'ici. En théorie, les RNN peuvent utiliser des informations dans des séquences arbitrairement longues, mais dans la pratique, on les limite à regarder seulement quelques étapes en arrière. Voici à quoi ressemble un RNN typique :

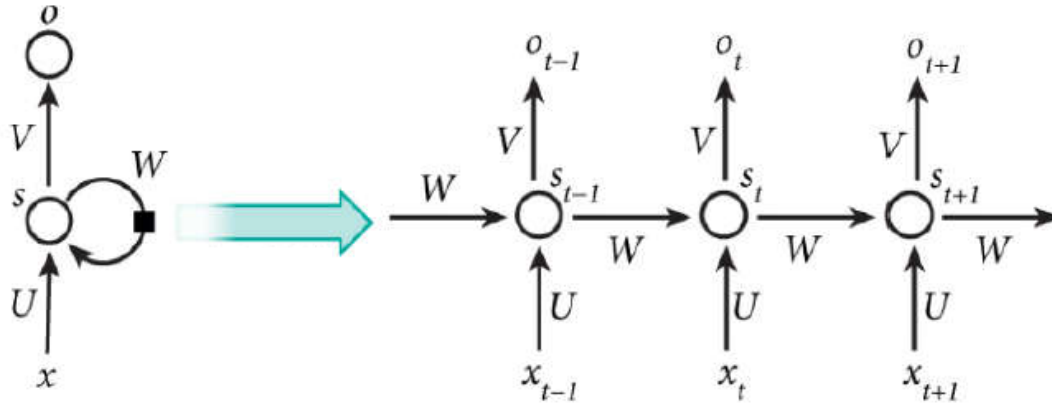


Figure 2.8: (à gauche) Un RNN. (à droite) Sa version déroulée Source : Nature

Le schéma ci-dessus montre un RNN déroulé. En déroulant, nous signifions simplement qu'on montre le réseau pour la séquence complète. Par exemple, si la séquence qui nous intéresse est une phrase de 5 mots, le réseau serait déroulé en un réseau de neurones de 5 couches, une couche pour chaque mot. Les formules qui régissent les calculs dans un RNN sont les suivantes :

_ X_t est l'entrée au moment t .

_ $U; V; W$ sont les paramètres que le réseau va apprendre des données de l'apprentissage.

_ S_t est l'état caché au moment t . C'est la « mémoire » du réseau. s_t est calculé en fonction de l'état

Caché précédent et de l'entrée à l'étape actuelle :

$$S_t = f(Ux_t + Ws_{t-1})$$

Où f est une fonction non linéaire telle que : ReLu ou Hyperbolic tangent (\tanh).

_ O_t est la sortie au moment t . Par exemple, si on veut prédire le prochain mot dans une phrase, ce serait un vecteur de probabilités dans un vocabulaire.

$$O_t = \text{softmax}(Vs_t)$$

III.3. Avantages de CNNs:

Un avantage majeur des réseaux convolutifs est l'utilisation d'un poids unique associé aux signaux entrant dans tous les neurones d'un même noyau de convolution. Cette méthode réduit l'empreinte mémoire, améliore les performances (13) et permet une invariance du traitement par translation. C'est le principal avantage du CNN par rapport au MLP, qui lui considère chaque neurone indépendant et donc affecte un poids différent à chaque signal entrant. Lorsque le volume d'entrée varie dans le temps (vidéo ou son), il devient intéressant de

rajouter un paramètre de temporisation (delay) dans le paramétrage des neurones. On parlera dans ce cas de réseau neuronal à retard temporel (TDNN).

Comparés à d'autres algorithmes de classification de l'image, les réseaux de neurones convolutifs utilisent relativement peu de pré-traitement. Cela signifie que le réseau est responsable de faire évoluer tout seul ses propres filtres (apprentissage sans supervision), ce qui n'est pas le cas d'autres algorithmes plus traditionnels. L'absence de paramétrage initial et d'intervention humaine est un atout majeur des CNN.

IV.Conclusion:

Dans ce chapitre on a présenté les notions importantes qui sont en relation avec l'apprentissage profond (définition, Architectures....etc). Aussi qu'une vision générale sur l'apprentissage profond, tout en donnant en détail la méthode choisie dans notre travail de recherche qui est le CNNs. Le prochain chapitre, traite les détails de la conception, ainsi que la méthode et les outils utilisés pour la réalisation de notre application.

CHAPITRE : 3

Implémentation et résultats

I. Introduction :

Après avoir présenté , les notions générale de télédétection et les techniques de l'apprentissage profonde. Dans ce chapitre , nous présentons les outils utilisées dans notre application, l'architecture de notre approche , l'implémentation et les résultats obtenus.

II. Ressources matérielles et logicielles :

II.1. Ressources matérielles :

Notre système est développé dans un ordinateur dont les caractéristiques techniques, sont les suivantes :

- Processeur : Intel « I5 »
- RAM 6 Go

II.2. Ressources logicielles :

Système d'exploitation : Windows 7 titan 64bit

II.2.1. Langage de programmation : Python est un langage de programmation, un des langages de niveau supérieur, caractérisé par une écriture et une lecture simples, facile à apprendre, utilisant un style de programmation ouvert, et évolutif. Python est un langage polyvalent 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.[W6]



Figure 3.1: Logo Python

II.2.2. PyCharm : est un environnement de développement intégré utilisé pour programmer en Python.

Il permet l'analyse de code et contient un débogueur graphique. Il permet également la gestion des tests unitaires, l'intégration de logiciel de gestion de versions, et supporte le développement web .

Développé par l'entreprise tchèque JetBrains, c'est un logiciel multi-plateforme qui fonctionne sous Windows, Mac OS X et Linux. Il est décliné en édition professionnelle, diffusé sous licence propriétaire, et en édition communautaire diffusé sous licence Apache.[W ,7]

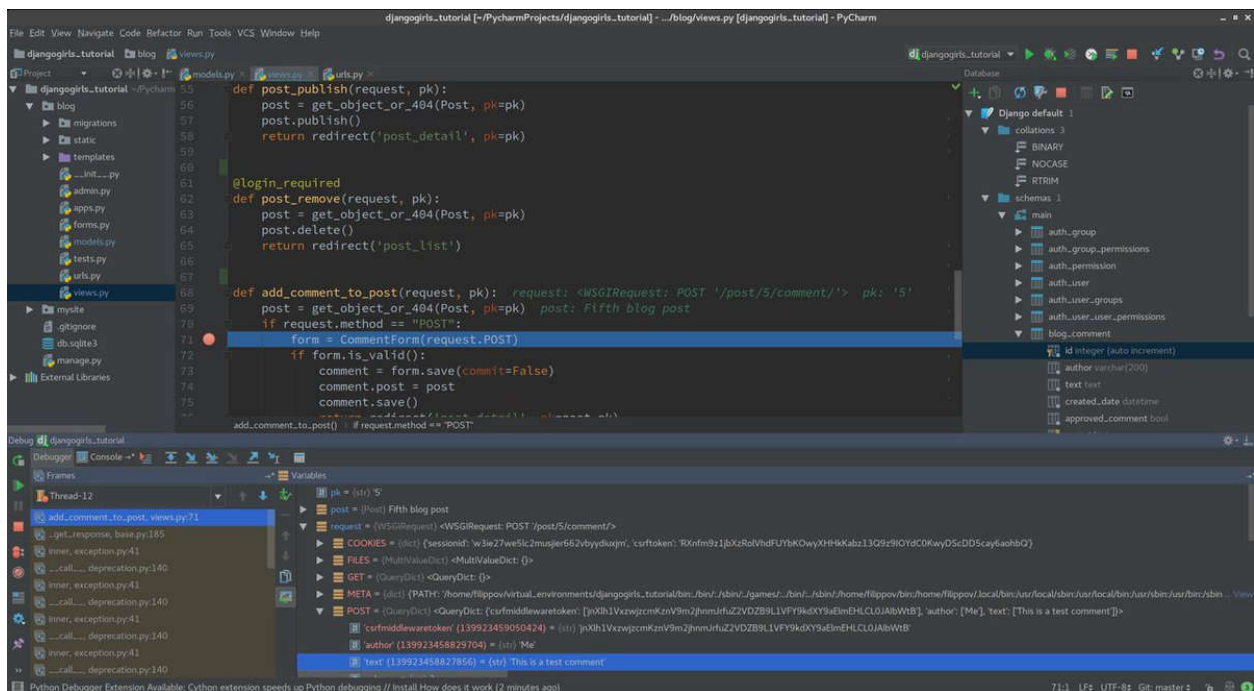


Figure 3.2 : Interface JetBrains PyCharm

II.2.3. Anaconda : est une distribution libre et open source des langages de programmation Python et R 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), qui vise à simplifier la gestion des paquets et de déploiement. Les versions de paquetages sont gérées par le système de gestion de paquets conda. La distribution Anaconda est utilisée par plus de 6 millions d'utilisateurs et comprend plus de 250 paquets populaires en science des données adaptés pour Windows, Linux et MacOS. [W,8]

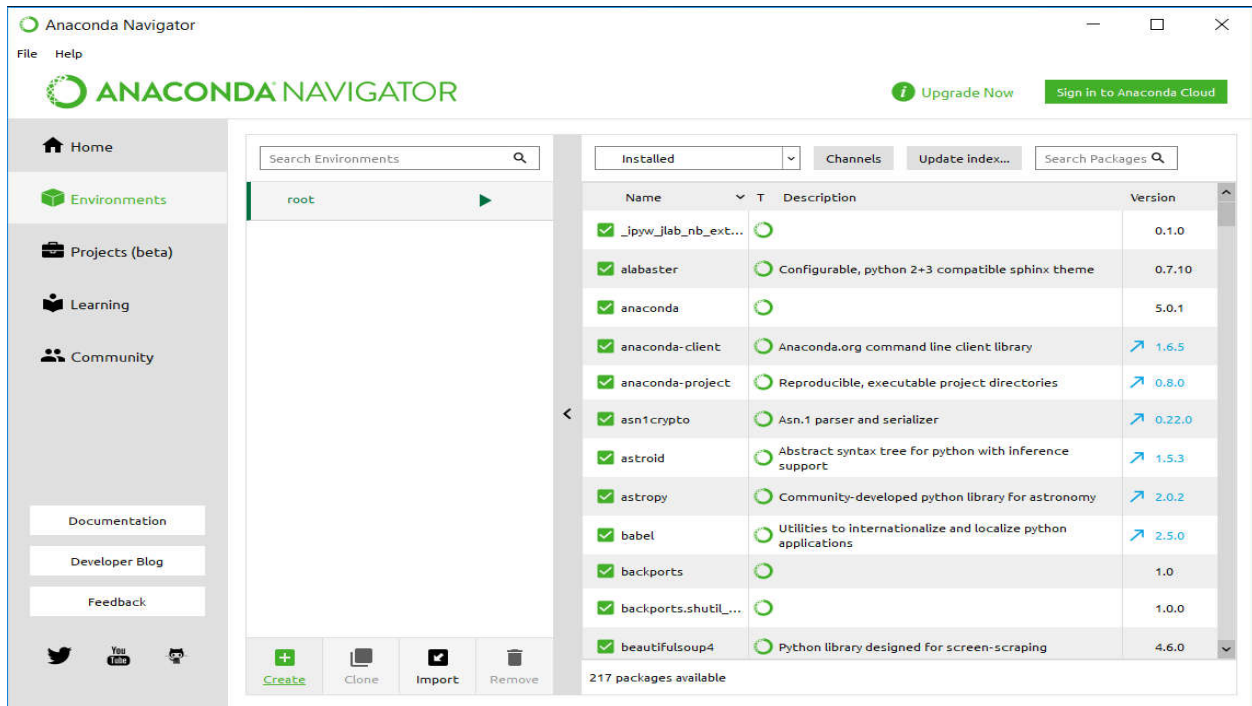


Figure 3.3: Interface Anaconda

II.2.4. TensorFlow : est une bibliothèque de logiciels open source pour le calcul numérique à l'aide de graphiques de flux de données. Les nœuds de graphique représentent des opérations mathématiques, tandis que les arêtes de graphique représentent les tableaux de données multidimensionnels (tenseurs) qui circulent entre eux. Cette architecture flexible vous permet de déployer des calculs sur un ou plusieurs processeurs ou GPU sur un ordinateur de bureau, un serveur ou un périphérique mobile sans réécrire le code. TensorFlow inclut également TensorBoard, une boîte à outils de visualisation de données. [W,9]

TensorFlow a été initialement développé par des chercheurs et des ingénieurs travaillant au sein de l'équipe Google Brain au sein de l'organisation de recherche sur l'intelligence artificielle de Google dans le but de mener des recherches sur le machine learning et les réseaux de neurones profonds. Le système est suffisamment général pour pouvoir s'appliquer à une grande variété d'autres domaines également. [W,9]

TensorFlow fournit des API Python et C stables ainsi que des API rétro compatibles non garanties pour C ++, Go, Java, JavaScript et Swift.



Figure 3.4: logo TensorFlow

II.2.4.1. Pourquoi TensorFlow ?

Que vous soyez expert ou débutant, TensorFlow est une plate-forme de bout en bout qui facilite la création et le déploiement de modèles ML.

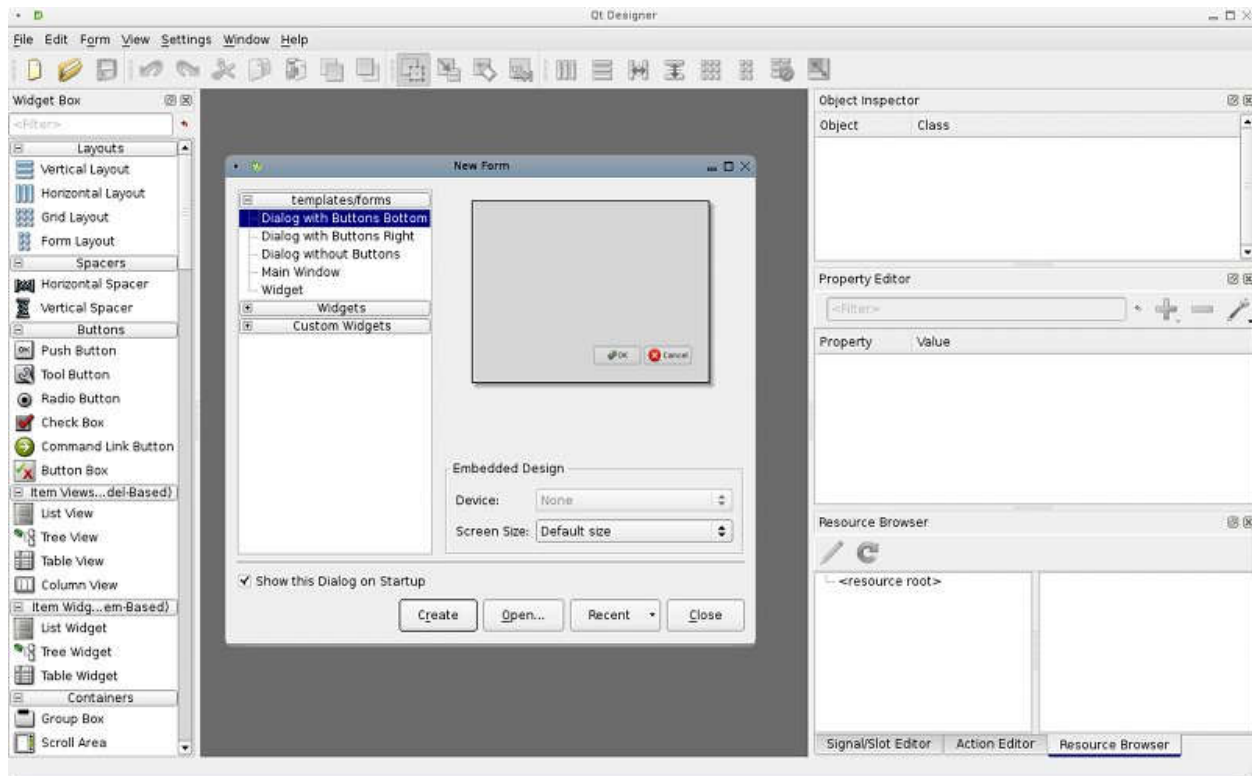
TensorFlow offre plusieurs niveaux d'abstraction afin que vous puissiez choisir celui qui convient le mieux à vos besoins. Créez et entraînez des modèles à l'aide de l'API Keras de haut niveau, ce qui facilite la mise en route de TensorFlow et l'apprentissage automatique.

Si vous avez besoin de plus de flexibilité, une exécution rapide permet une itération immédiate et un débogage intuitif. Pour les tâches de formation de grande taille, utilisez l'API de stratégie de distribution pour une formation distribuée sur différentes configurations matérielles sans modifier la définition du modèle. [W,9]

II.2.5. PyQt : Qt est un ensemble de bibliothèques C ++ multiplates-formes qui implémentent des API de haut niveau pour accéder à de nombreux aspects des systèmes de bureau et mobiles modernes. Ceux-ci incluent les services de localisation et de positionnement, la connectivité multimédia, NFC et Bluetooth, un navigateur Web basé sur Chromium, ainsi que le développement d'interface utilisateur traditionnel. [W,10]

PyQt5 est un ensemble complet de liaisons Python pour Qt v5. Il est implémenté sous la forme de plus de 35 modules d'extension et permet à Python d'être utilisé comme langage de développement d'application alternatif au C ++ sur toutes les plateformes prises en charge, y compris iOS et Android.

PyQt5 peut également être intégré à des applications basées sur C ++ pour permettre aux utilisateurs de ces applications de configurer ou d'améliorer les fonctionnalités de ces applications. [W,10]

**Figure 3.5:** interface Qt Designer

III. Description de CNN:

La convolution est l'acte consistant à prendre les données d'origine et à en créer des cartes. Le polissage est un sous-échantillonnage, le plus souvent sous la forme de "pooling maximal", où nous sélectionnons une région, puis prenons la valeur maximale dans cette région. , et cela devient la nouvelle valeur pour toute la région. Les couches entièrement connectées sont des réseaux de neurones typiques, dans lesquels tous les nœuds sont "entièrement connectés". Les couches convolutives ne sont pas entièrement connectées comme un réseau de neurones traditionnel.

La structure CNN de base est la suivante: Convolution -> Pooling -> Convolution -> Pooling -> Couche entièrement connectée -> Sortie [W,5]

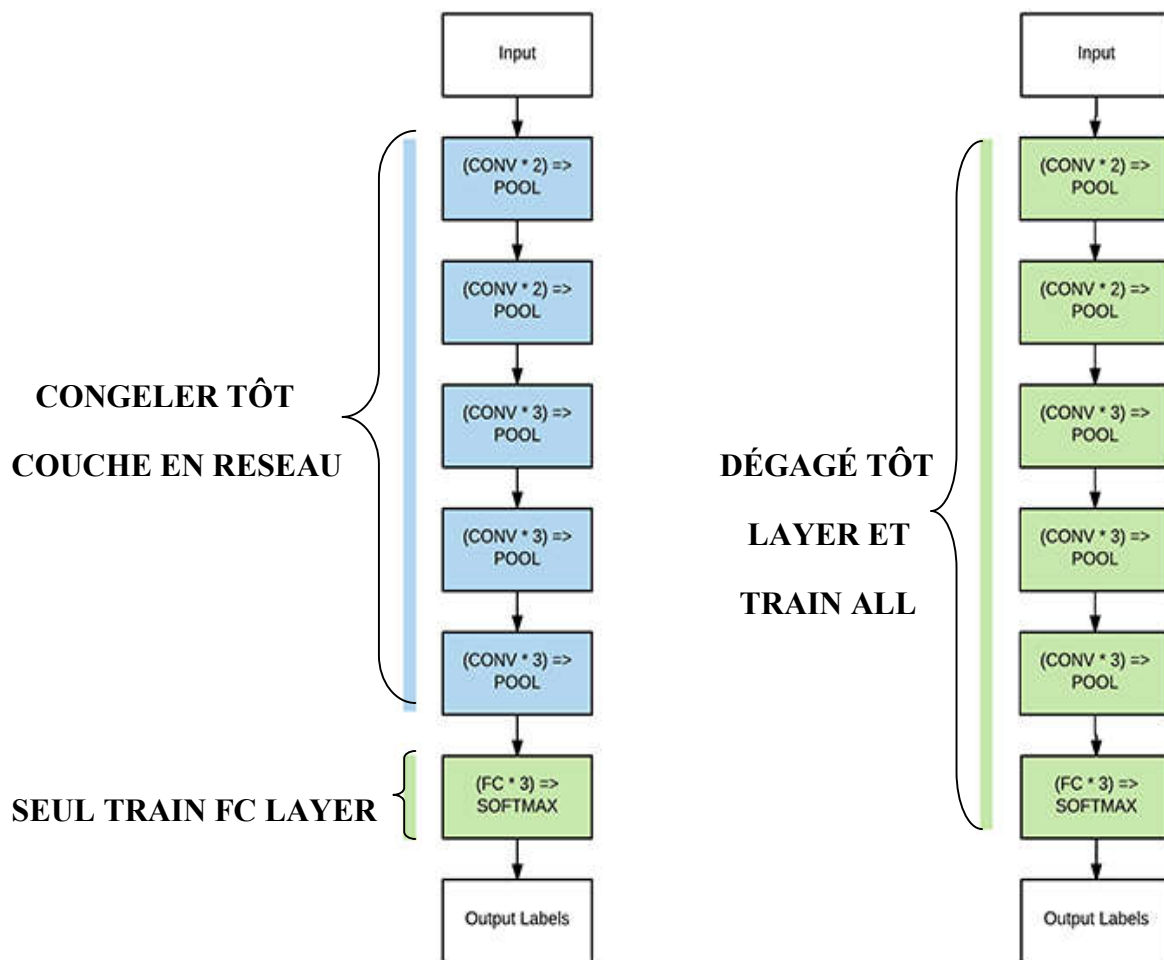


Figure 3.6: La structure CNN[W,5]

Figure 3: Gauche: lorsque nous commençons le processus de réglage précis, nous gèlons toutes les couches CONV du réseau et ne permettons au dégradé que de se propager à travers les couches FC. Cela permet à notre réseau de se «réchauffer». À droite: une fois que les couches FC ont eu la chance de se réchauffer, nous pouvons choisir de débloquer toutes les couches ou une partie de celles-ci plus tôt dans le réseau et permettre à chacune d'entre elles d'être également ajustées.

Les données de formation sont propagées en avant sur le réseau comme nous le ferions habituellement; cependant, la rétropropagation est arrêtée après les couches FC, ce qui permet à ces couches de commencer à apprendre des motifs à partir des couches CONV hautement discriminantes.

Dans certains cas, nous pouvons décider de ne jamais débloquent le corps du réseau car notre nouvelle tête FC peut obtenir une précision suffisante.

Cependant, pour certains jeux de données, il est souvent avantageux de permettre également aux couches CONV d'origine d'être modifiées au cours du processus de réglage précis (Figure 3.6, à droite).

Une fois que la tête FC a commencé à apprendre les modèles de notre jeu de données, nous pouvons suspendre l'entraînement, dégeler le corps et poursuivre l'entraînement, mais avec un très petit taux d'apprentissage - nous ne voulons pas modifier nos filtres CONV.

La formation est ensuite autorisée à se poursuivre jusqu'à obtenir une précision suffisante.

Le réglage précis est une méthode extrêmement puissante pour obtenir des classificateurs d'images sur vos propres ensembles de données personnalisés à partir de CNN pré-formés (et est encore plus puissant que l'apprentissage par transfert via l'extraction d'entités).

IV. Base d'image:

Notre projet s'appuie sur la classification des images satellitaires à l'aide du réseau de neurones à convolution (CNN). La base de donnée utilisée pour l'apprentissage est appelée UC Merced. On peut télécharger gratuitement cette base à partir l'internet[W ,4]. Cette base contient 2100 images des 21 catégories. Chaque catégorie contient 100 image. Les images de cette base sont en forme RGB. La taille de chaque image est 256*256

Dans note projet , nous utilisons 85 images de chaque classe pour l'apprentissage et 15 image pour le test.

La figure ci-dessus illustre des images apprenant dans la base UC Merced.



Figure 3.7 : UC Merced land-use Dataset (RGB) .[W ,4]

V. L'architecture de l'application

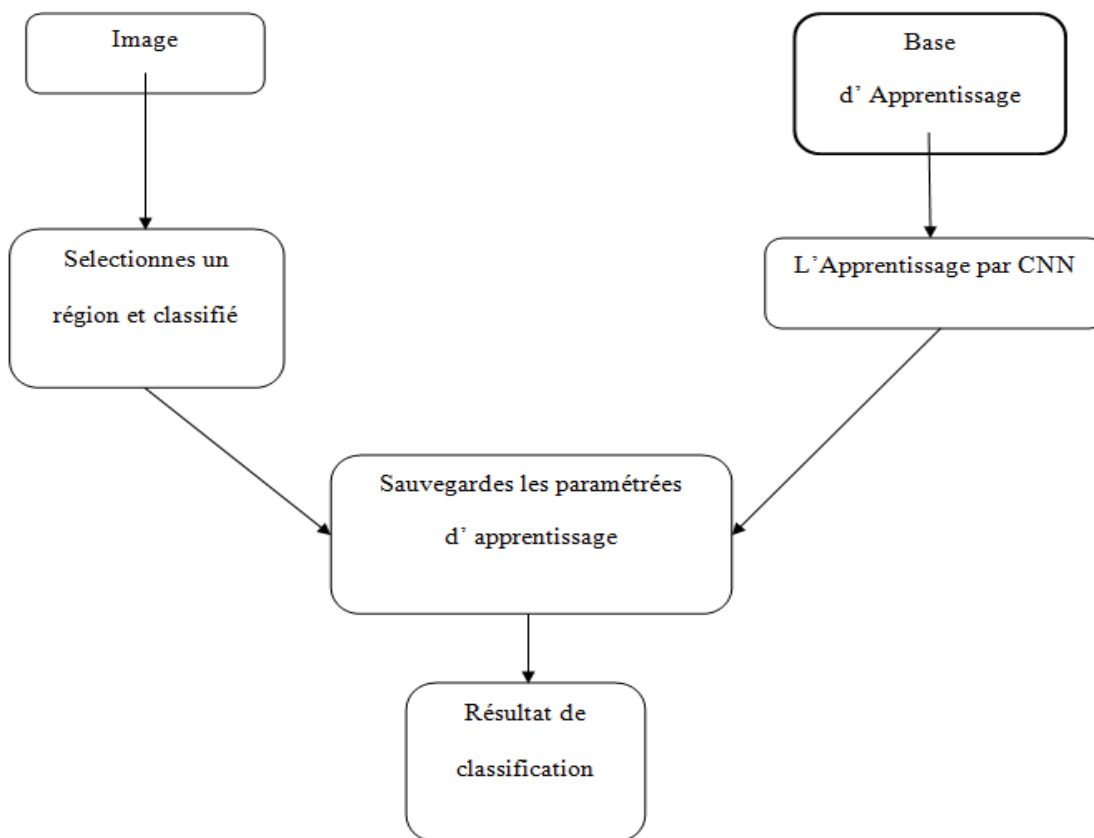


Figure 3.8: Architecteur de l'application

V.1. Base d'apprentissage :

Dans cette étape, on dispose les images pour l'apprentissage . Nous avons utilisé trois classes. Le tableau ci-dessus illustre la structure de notre base d'apprentissage et les classe utilisées.

Classe	Apprentissage	Test
Végétation	85 image	15 image
Route	85 image	15 image
Urbain	85 image	15 image

Tableau : les classe utilisées.

- **Sauvegarder les paramètres de l'apprentissage** : Après l'étape de l'apprentissage les paramètres sont sauvegardés, Ces parmters sont utilisés pour le test et la classification.

V.2. L'apprentissage par CNN :

Dans cette étape les images de l'apprentissage sont l'entrée de CNN

La structure de CNN utilisée est illustrée dans la figure 6. La figure ci-dessus illustre la button "Apprentissage

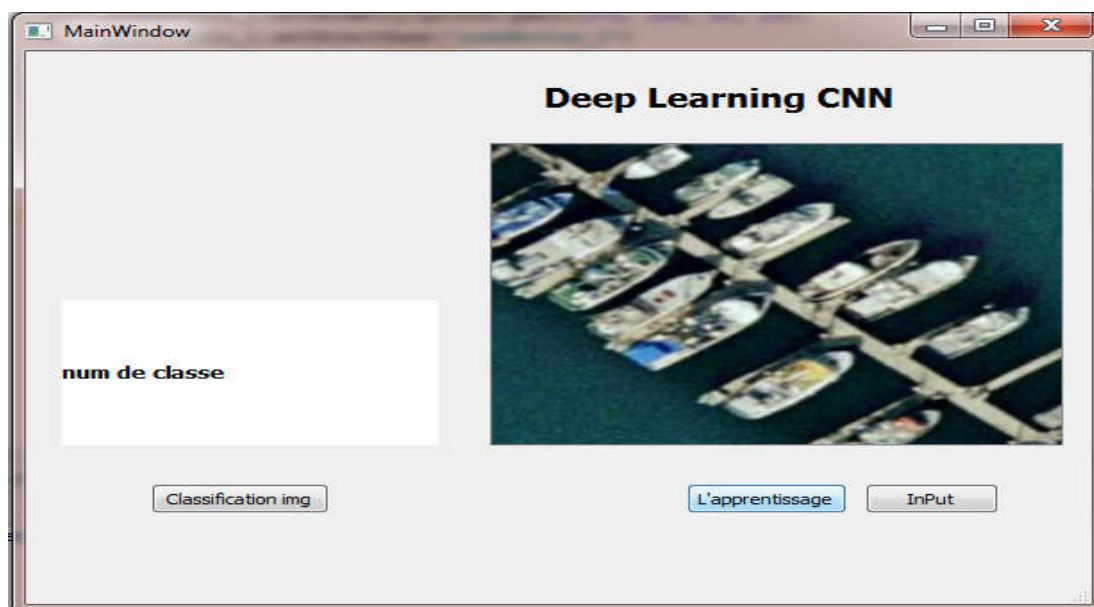


Figure 3.9 : Apprentissage Image

Cette étape consiste à choisir une image.

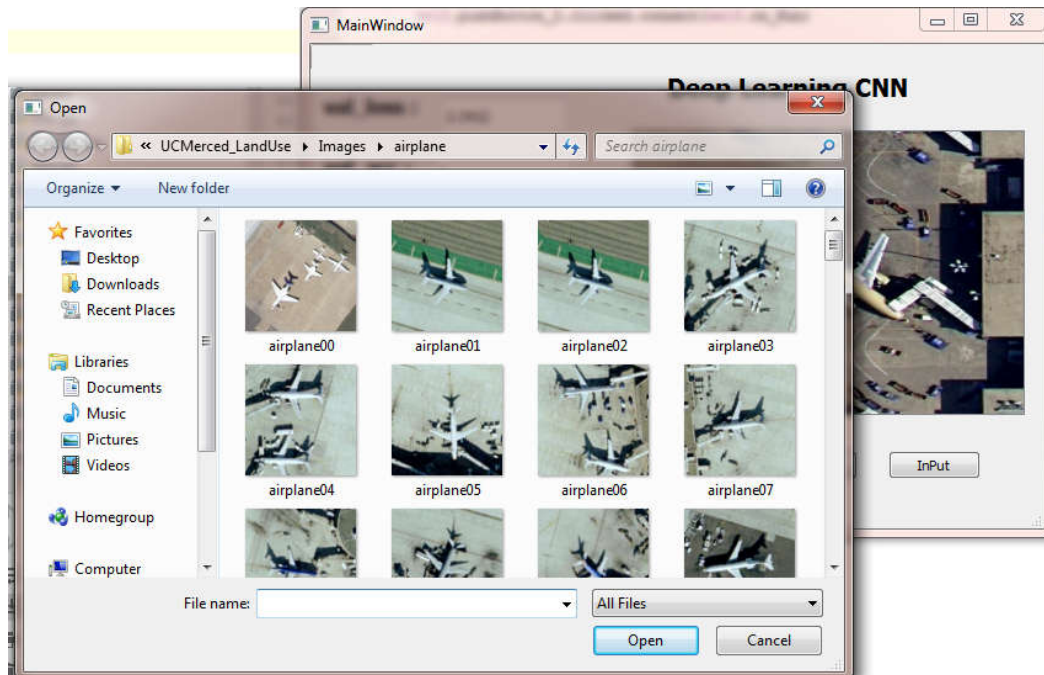


Figure 3.10 : Insérer l'image

V.3. Sélectionner une région :

On sélectionne la position de l'objet de l'image qu'on veut le classifier. La région sélectionnée est convertie en une image. Cette image est considérée comme un exemple à classifié

La figure ci-dessus montre la sélection d'une région sélectionnée.

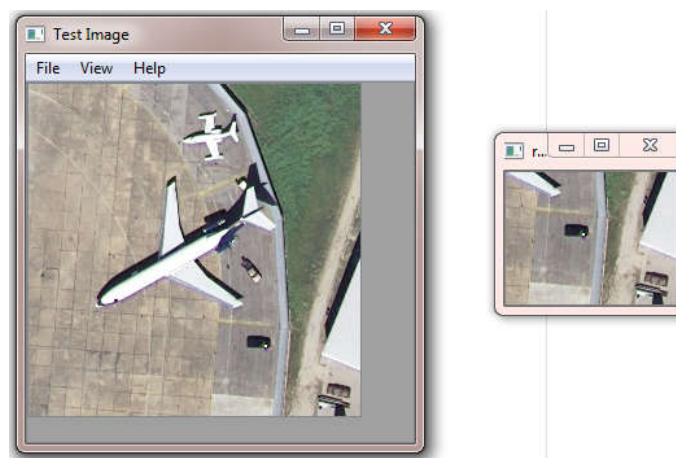


Figure 3.11: Prendre une région de photo

Classification :

Après de sauvegarder les paramètres de l'apprentissage, et de sélection la région à classifié. On peut connaitre la classe de la région sélectionné comme il est montré dans l'interface ci-dessus :

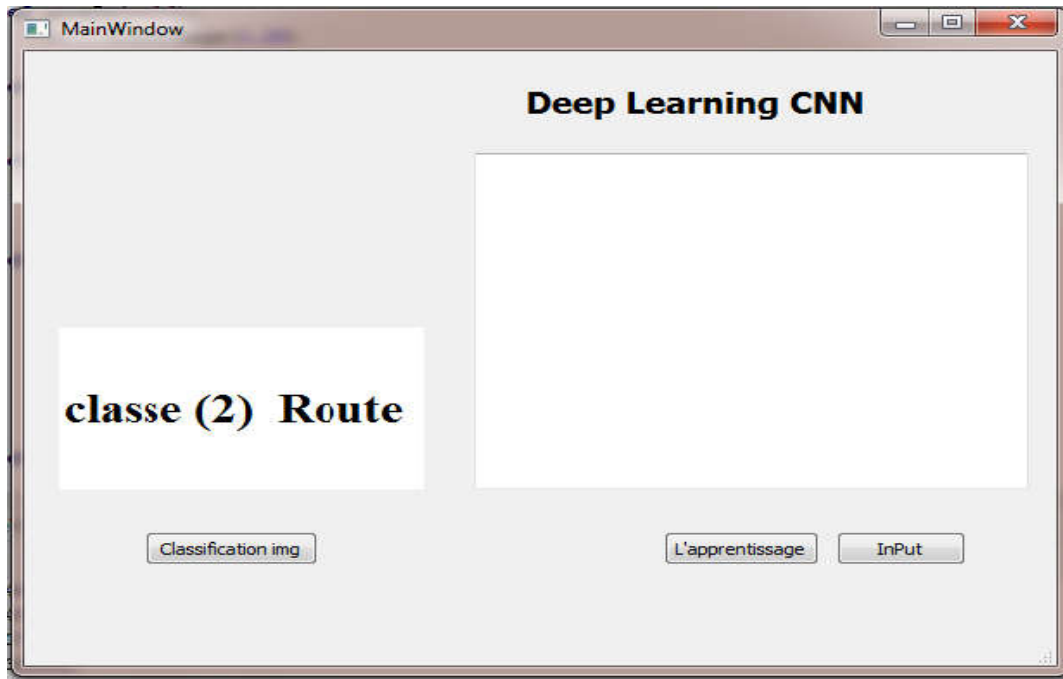


Figure 3.12: Classification des images

V.4. Evaluation de résultat :

L'apprentissage donne un taux de classification = 20%

La figure suivant montre la résultat de classification :



Figure 3.13 : la résultat de classification

La faiblesse de ce taux est à cause que la base de l'apprentissage est insuffisante

VI. Conclusion:

Dans ce, chapitre nous avons expliqué l'approche utilisée et les résultats obtenues. Dans la suite, nous présentons notre conclusions et perspectives

Conclusion générale

La classification d'images est une tâche importante dans le domaine de la vision par ordinateur, la reconnaissance d'objets et l'apprentissage automatique.

Grâce à l'apprentissage en profondeur (Deep Learning) , L'avenir de l'intelligence artificielle dans le développement de grands et très rapides.

Les systèmes d'informations géographiques sont entrain 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 projet, nous avons discuté de la manière dont le processus de télédétection et de la classification des images prises par satellite lors de la discussion des concepts de base des réseaux de neurones dans Réseaux de neurones généraux et thalacophytes en particulier. Nous avons fourni ces Réseaux de neurones en fournissant différents types de couches utilisées.

Pour réaliser notre travail de classification on a utilisé le deep learning, la méthode d'apprentissage qui a montré ses performances ces dernières années et nous avons choisi la méthode CNNs comme méthode de classification, ce choix est justifié par la simplicité et l'efficacité de la méthode.

Le résultat obtenu lors de la phase de test confirme l'efficacité de notre approche.

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

Références bibliographiques

- [Bon, 1992] Bonn, F. et G. Rochon, "*Précis de télédétection*" – Volume 1. Principes et méthodes, Sainte-Foy, Presses de l'Université du Québec/AUPELF, 485 p.
- [Coc, 1995] Coquerez, J., Philipp, S., "*Analyse d'images : filtrage et segmentation*", Masson, Enseignement De La Physique Traitement Du Signal, 1995.
- [Han, 2010] HANNANE 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 projet de fin d'étude. Université des sciences et de la technologie d'Oran Mohamed Boudiaf, Oran, Algérie, 2006.
- [MMZ, 2017] Mr. Mokri Mohammed Zakaria., » Classification des images avec les réseaux de neurones convolutionnels ": mémoire de projet de fin d'étude. Université Abou Bakr Belkaid Tlemcen.
- [Ouf, 1998] K, AOUF & M, RAYAH " Classification d'image par intégration du MNT ". USTO 1998.
- [Rab, 2012] Mr. Rabhi Seddik., " Contribution d'une méta-heuristique bio-inspirée dans le traitement d'image": mémoire de projet de fin d'étude. Université des sciences et de la technologie d'Oran Mohamed Boudiaf, Oran, Algérie, 2006.
- [Sma, 2004] Smaïne El Kahili B., Boudour G., "Conception d'un système hybride, algorithme génétique et réseau RBF pour une classification d'image satellitaire". Université des sciences et de la technologie d'Oran Mohamed Boudiaf, Oran, Algérie 2004.

SET WEB

- [W,1] (04 Août 2015) https://perso-sdt.univ-brest.fr/~delacour/COURS/M210.1/Cours2_Plateforme_Capteurs/
HY Dao " Cours de télédétection n°4645 " Université de Genève
- [W,2] (14 Février 2016) <http://www.psychomedia.qc.ca/lexique/definition/apprentissage-profond>
- [W,3] (15 Mai 2016) <https://discoverymu.info/r%C3%A9seau-neuronal-convolutif-pour-trouver.aspx>
- [W ,4] (28 Oct 2010) <http://weege.vision.ucmerced.edu/datasets/landuse.html>

Références bibliographiques

- [W ,5] (27 Mai 2019) <https://www.pyimagesearch.com/2019/05/27/keras-feature-extraction-on-large-datasets-with-deep-learning/>
- [W ,6] (2019) <https://www.python.org/>
- [W ,7] (2019)<https://plugins.jetbrains.com/pycharm>
- [W ,8] (2019)<https://www.anaconda.com/distribution/>
- [W ,9] (2019)<https://www.tensorflow.org/>
- [W ,10] (2019) <https://riverbankcomputing.com/software/pyqt/intro>