

La République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Saad Dahleb Blida 1



Institut d'Aéronautique et des Études Spatiales (IAES)
Département Études Spatiales

Mémoire de fin d'études

En vue de l'obtention du diplôme de Master

en : **Aéronautique**

Option : **Télécommunications Spatiales**

THÈME

*Détection des Navires avec les Images Satellitaires
Radar à Synthèse d'Ouverture*

Proposé et dirigé par :

M. Azmedroub Boussad

Réalisé par :

Mlle Aouadi Sabrina

M. Guettari Fares Mohamed Ali

Soutenu le 09 Juillet 2020, devant le jury composé de :

Mme Mouffok Lila MCA IAES Présidente

M. Azmedroub Boussad MCB IAES Encadreur

M. Tahraoui Sofiane MAB IAES Examineur

Promotion : 2019/2020

Remerciement

Tout d'abord, nous remercions Dieu le tout puissant de nous avoir accordé force, santé et la volonté pour réaliser ce travail.

L'aboutissement de ce travail n'aurait pas vu le jour sans la collaboration et l'aide de plusieurs personnes qui d'une façon directe ou indirecte ont contribué à sa concrétisation. Qu'elles trouvent à travers ces quelques lignes l'expression de nos chaleureux remerciements.

Tout d'abord, ce travail ne serait pas aussi riche sans l'aide et l'encadrement de Monsieur Azmedroub Boussad, on le remercie pour la qualité de son encadrement exceptionnel, pour sa patience, sa rigueur, sa disponibilité, ses conseils et ses commentaires précieux qui nous ont permis, de surmonter les difficultés et de progresser dans ce travail.

Tout notre respect et nos remerciements vont vers les membres du jury, Mme Mouffok Lila et M. Tahraoui Sofiane, qui vont pleinement consacrer leur temps et leur attention afin d'évaluer notre travail, qui espérons le sera à la hauteur de leur attente.

Nous tenons aussi à remercier chacun de nos parents Mme Aouadi Chiaz ,Mme Guettari Nadia ,Ms Aouadi Yahia , Ms Guettari Djamel qui nous ont poussés à faire les efforts nécessaires pour le bon déroulement et l'aboutissement de ce travail.

Nous tenons à exprimer également notre grande considération à l'ensemble de nos enseignants qui ont contribué à notre formation.

Dédicace

UN spéciale dédicace à ma tante préféré H.Guettari mes deux grand mère adoré Z.Rebai et T.Guettari, ma grande sœur M.Guettari, ma binôme S.Aouadi et son petit frère Akram Aouadi ainsi que H.Nouçair, F.Nouçair, K.Nouçair.

Sans oublier H.Nougrara, S.Addi , M.Izri , A.Kirimi , E.Bachtoubdji , Y.Kadour , H.hiber M.Ikhlef Et la mēilleure pour la fin S.Gharbi .

Résumé

LE Radar à Synthèse d'Ouverture SAR (Synthetic Aperture Radar) est un système d'imagerie, qui permet une observation de la surface terrestre en jour et nuit quelque soit les conditions météorologique. L'objectif de ce projet de fin d'étude est l'exploitation des images SAR pour la surveillance de la détection des navires. On a utilisé les images SAR de la région de la baie d'Alger, acquises par le satellite Sentinel 1 de l'Agence Spatiale Européenne. Après la préparation des images et le filtrage de speckle, on a commencé l'exploration des méthodes de détection des cibles avec la méthode CA-CFAR. Pour avoir de meilleurs résultats on a utilisé la méthode Mask R-CNN, qui est une méthode des réseaux de neurones avec l'apprentissage profond. Pour le faire on a construit une base de données des images SAR sur la baie d'Alger, pour faire l'apprentissage à notre modèle de réseau de neurones. En conclusion Mask R-CNN a donné de meilleurs résultats spécialement dans les régions hétérogènes.

Abstract

SYNthetic Aperture Radar (SAR) is an imaging system, which allows observation of the earth's surface day and night regardless of weather conditions. The objective of this end-of-study project is the exploitation of SAR images for the monitoring of ship detection. SAR images of the Bay of Algiers area acquired by the Sentinel 1 satellite of the European Space Agency were used. After image preparation and speckle filtering, exploration of target detection methods with the CA-CFAR method began. For better results we used the Mask R-CNN method, which is a neural network method with deep learning. To do this, we built a database of SAR images on the bay of Algiers, to train our neural network model. In conclusion Mask R-CNN gave better results especially in heterogeneous regions.

Table des matières

Remerciement

Résumé i

Table des matières ii

Liste des Abréviations iv

Table des figures v

Liste des tableaux vii

Introduction Générale 1

1 Images SAR 3

1.1 Imagerie Radar 3

1.2 Radar à Synthèse d'Ouverture 3

1.3 Principe de Formation des Images SAR 4

1.4 Résolution des Images SAR 5

1.4.1 Résolution Radiale 5

1.4.2 Résolution Azimutale 6

1.5 Bruit de l'image SAR 7

2 Méthode CFAR 8

2.1 Introduction 8

2.2 Les systèmes CFAR 9

2.2.1 CA-CFAR 9

2.2.2 L'architecture du CA-CFAR 9

2.3 Conclusion 11

3 Réseaux de neurones	12
3.1 Définition	12
3.1.1 Origine biologique	12
3.2 La structure des réseaux de neurones	15
3.3 Architecture des Réseaux de Neurones	15
3.3.1 Réseaux de neurones a une seule couche	15
3.3.2 Réseau de Neurone multi-couche	17
3.4 Apprentissage des Réseaux de Neurones	19
3.4.1 L'apprentissage supervisé	19
3.4.2 L'apprentissage non-supervisé	20
3.5 Les réseaux de neurones convolutifs	21
3.5.1 Définition	21
3.5.2 L'architecture des réseaux neurones convolutifs	21
3.5.3 La couche de pooling (POOL)	23
3.5.4 Réseaux de neurones de convolution pour la détection d'objets	26
3.6 conclusion	30
4 Analyses et Résultats	31
4.1 Présentation des données	31
4.2 Transfert Learning	32
4.2.1 Transfert Learning pour la détection des navires	32
4.3 Détection des navires avec le deep learning	33
4.4 Application sur la baie d'Alger	36
4.4.1 Images baie d'Alger traitées Via CA-CFAR	38
4.4.2 Images de la baie d'Alger traitées Via Mask RCNN	40
4.5 La comparaison entre CA-CFAR et Mask RCNN	42
Conclusion Générale	44
Bibliographie	46

Liste des Abréviations

SAR	Radar à Synthèse d'ouverture (Synthetic Aperture Radar).
PRF	La fréquence de répétition des impulsions (Pulse Repetiton Frequency).
SLAR	Side Looking Airborne Radars.
CFAR	Taux constant de fausses alertes (Constant False Alarm Rate).
CUT	Cell Under Test .
CA-CFAR	Cell Averaging CFAR.
ROI	Region Of Interest.
PDF	Probability Density Function.
CNN	Convolutional Neural Networks .
MLP	Multi Layers Perceptron .
POOL	La couche de pooling.
ReLU	Unité Linéaire Rectifiée.
FC	Fully Connected.
R-CNN	Regions with CNN features.
YOLO	You Only Look Once.
SSD	Single Shot Detector.
HH	Emission Horizontale , réception Horizontale.
VV	Emission Verticale , réception Verticale.
HV	Emission Horizontale , réception Verticale.
mAP	Mean Average Precision

Table des figures

1.1	Première image SAR acquise en 1957	4
1.2	Antenne mon statique	4
1.3	Géométrie de la visée latérale du SAR	5
1.4	Principe de la compression d'impulsion	6
1.5	Schéma d'un système SAR	6
2.1	Détecteur CA-CFAR pour un bruit Gaussien	9
2.2	Affichage de region of interest	10
2.3	Application du principe CA-CFAR	10
2.4	L'architecture de l'algorithme CA-CFAR	11
3.1	Modèle du neurone biologique	13
3.2	La comparaison entre un neurone biologique et un neurone artificiel	13
3.3	Modèle d'un neurone artificiel	14
3.4	Représentation matricielle du modèle d'un neurone artificiel	14
3.5	Schéma d'un réseau de neurone	15
3.6	Réseaux de neurones a une seule couche	16
3.7	Représentation matricielle d'un réseau de neurone à une couche de S neurones	16
3.8	Structure d'un réseaux de neurones statique a trois couches	17
3.9	Matrice d'un réseau de neurones statique a trois couches	17
3.10	Matrice d'un réseaux de neurones dynamique a trois couches	18
3.11	Type d'apprentissage en réseau de neurone.	19
3.12	Architecture d'un CNN (Convolutional Neural Network)	22
3.13	L'architecture d'un réseau neurone convolutif	22
3.14	Schéma du parcours de la fenêtre de filtre sur l'image.	23
3.15	Max pooling avec un filtre 2×2 et un pas de 2.	24
3.16	La fonction ReLu	25

3.17	Architecture R-CNN	27
3.18	Architecture Faster R-CNN [1].	27
3.19	Faster R-CNN : Fenêtre coulissante et résultat [2].	28
3.20	Algorithme YOLO	28
3.21	Single shot detector	29
4.1	Forme des données images et annotations	32
4.2	Detection de navires.	35
4.3	Copernicus Open Access Hub	36
4.4	switch to the area mode	37
4.5	Recherche avancée	37
4.6	Resultats obtenus	37
4.7	Image Sentinel1 de la baie d'Alger dans Google Earth.	38
4.8	Image de la baie d'Alger 600×600	38
4.9	Visualisation des images en variant le facteur de seuil α	39
4.10	Visualisation des images en variant la taille de la fenêtre.	39
4.11	Image VV de la baie d'Alger	40
4.12	Image HV de la baie d'Alger	40
4.13	Detection de la baie d'Alger via Mask RCNN	41
4.14	Detection de la baie d'Alger via Mask RCNN	41
4.15	La comparaison entre CA-CFAR et Mask RCNN.	42

Liste des tableaux

4.1	Les images traitées	34
-----	-------------------------------	----

Introduction Générale

Dans un contexte de dégradation de la sécurité au niveau mondial, les menaces sécuritaires provoquées par divers trafics (stupéfiants, armes), par l'immigration clandestine, par la pêche illégale ou par la pollution par hydrocarbures ont montré l'importance d'une surveillance des zones maritimes. Ainsi, la surveillance maritime constitue aujourd'hui un énorme enjeu dans le domaine de la sécurité et dans celui de l'économie. Cette surveillance peut être réalisée par le biais de radars ou de façon aéroportée. Mais l'étendue du territoire à surveiller rend la télédétection aérospatiale bien adaptée à cette problématique. C'est dans ce cadre que s'inscrit cette thèse. Nous nous intéressons particulièrement à la surveillance et la détection de navires.

Il existe deux manières principales pour la cartographie de la terre : l'observation par des systèmes d'acquisition d'images optiques ou par des systèmes radar. L'utilisation des systèmes radar a plus d'avantages que l'autre puisque la production des images est possible de jour comme de nuit grâce à leur fonctionnement indépendant des conditions d'illumination ainsi que des conditions atmosphériques.

Le Radar à Synthèse d'Ouverture, appelé Synthetic Aperture Radar (SAR) en anglais est un système d'imagerie radar qui permet d'atteindre en pratique des résolutions spatiales élevés. Il est soit aérien ou satellitaire.

Les capteurs embarqués sur des porteurs aériens fournissent de meilleures résolutions tandis que les porteurs spatiaux sont plus stables et ont la possibilité de capter des zones plus étendues.

Pour obtenir une détection parfaite, La reconnaissance de formes est entrée dans une nouvelle ère avec le développement des algorithmes d'apprentissage en profondeur, ou Deep-Learning dans la communauté scientifique comme dans le monde industriel, grâce aux avancées majeures réalisées ces 5 dernières années essentiellement dans le traitement

des images.

Le document est organisé comme suit :

Le premier chapitre comprend la présentation du radar à synthèse d'ouverture, la formation des images SAR et leurs caractéristiques.

Le deuxième chapitre parle de l'une des méthodes la plus utilisée pour la détection des navires CFAR(Constant False Alarm Rate) dont le CA-CFAR (Cell Averaging CFAR).

Le troisième chapitre est dédié aux réseaux de neurones en général et les réseaux de neurones convolutifs spécifiquement.

Afin de compléter l'étude et se replacer dans le contexte industriel et opérationnel de la thèse, il nous a paru important de réaliser des expérimentations de l'ensemble de l'algorithme sur des images complètes pour évaluer les performances et le temps de traitement. Les résultats seront présentés dans le chapitre 4.

Efin nous concluons sur cette étude et proposerons des perspectives sur la suite des travaux.

Chapitre 1

Images SAR

1.1 Imagerie Radar

Le radar est un système permettant de déterminer la distance et la vitesse d'un objet en utilisant des ondes radio. Il a aussi d'autres applications dans plusieurs domaines, une de ces applications est l'imagerie radar grâce aux ondes électromagnétiques de fréquences beaucoup plus courtes que les ondes radio, au lieu d'utiliser la lumière visible pour imager une zone (imagerie optique).

L'avantage de ce système est son fonctionnement à n'importe quelle heure (jour et nuit) comme le radar lui-même est la source d'illumination et aussi certaines bandes de fréquence peuvent pénétrer plusieurs couches permettant de faire des acquisitions par temps couvert. [3]

1.2 Radar à Synthèse d'Ouverture

SAR est un système d'imagerie radar d'un porteur aérien ou satellitaire équipé d'une antenne mono-statique et d'un système d'acquisition et de traitement des signaux rétrodiffusés par la surface terrestre. Il permet d'utiliser le déplacement du porteur pour simuler une antenne virtuelle de longueur très importante afin d'affiner la résolution de l'image le long de la trace (en Azimut). Les premiers systèmes SAR ont été développés au début des années 50 et les premières images acquises à la fin de cette décennie.

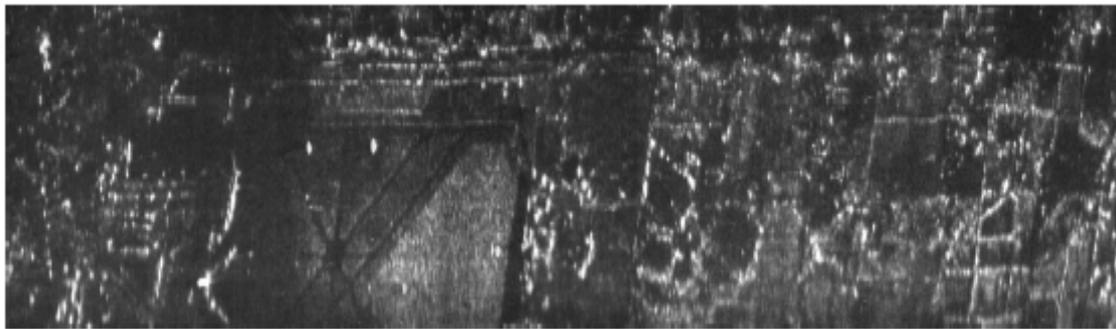


FIGURE 1.1 – Première image SAR acquise en 1957

1.3 Principe de Formation des Images SAR

Un radar imageur émet des ondes électromagnétiques dans une direction donnée. La surface illuminée diffuse l'onde incidente dans toutes les directions, dont la direction de l'émission. La partie qui est ré-émise vers le radar est enregistrée et traitée afin de former l'image radar. L'antenne du radar sert alternativement à l'émission et à la réception des ondes grâce à un dispositif (rotateur ou circulateur) qui permet à l'antenne de commuter d'un mode à l'autre.

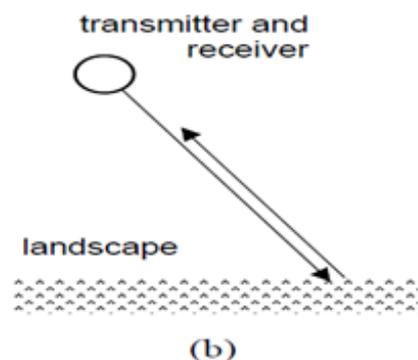


FIGURE 1.2 – Antenne mon statique

L'antenne envoie le signal sous forme de petite impulsion à la fréquence d'opération du radar pour distinguer les longueurs suivant la direction perpendiculaire à la direction de l'antenne. Les impulsions sont répétées suivant une fréquence synchronisée avec la vitesse de déplacement de l'antenne pour que chaque impulsion couvre une partie comme suivant la direction azimutale. On appelle la fréquence de répétition des impulsions par PRF (Pulse Repetition Frequency). Elle envoie des impulsions de durée τ vers le sol, avec une inclinaison déterminée par rapport à la verticale pour éviter que deux points symétriques par rapport à la perpendiculaire ne soient confondus en un seul point dans le système

imageur et à la fréquence de répétition des impulsions.

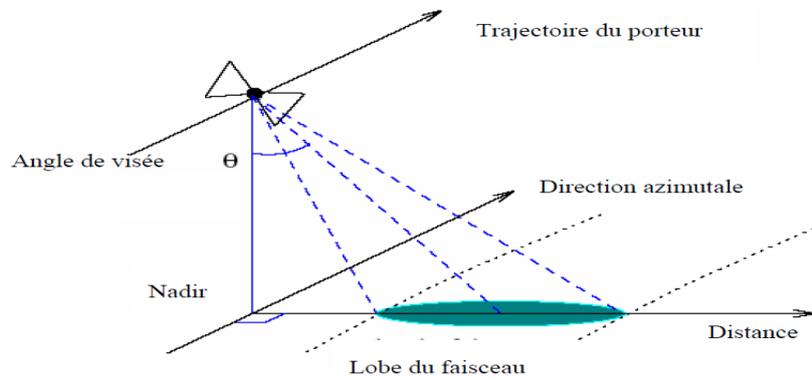


FIGURE 1.3 – Géométrie de la visée latérale du SAR

1.4 Résolution des Images SAR

La résolution spatiale est le paramètre fondamentale d'un système imageur, c'est la capacité de séparer deux objets voisins selon certaine distance. si la résolution est grande, les deux objets voisins seront placés dans deux cases de résolutions différentes. Sinon, ils seront situés dans la même case. On distingue deux types, la résolution radiale et azimutale.

1.4.1 Résolution Radiale

Donnée par $R_d = \frac{c \tau}{2 \sin \theta}$ la distance radiale au sol R_d dépend de l'angle d'incidence θ et de la durée des impulsions émises par le radar τ . Il y aura pas de résolution spatiale si $\theta = 0$ (directement au-dessous de l'antenne) c'est pourquoi on utilise des systèmes qui émettent par le côté appelés SLAR (side looking airborne radars).

Pour augmenter la résolution, on utilise le signal chirp (varier linéairement en fréquence suivant le temps). Sa formule est donnée par :

$$c(t) = p(t) \cos(w_0 + 0.5\alpha t)t = p(t) \cos(w_0 t + 0.5\alpha t^2) = p(t) \cos(2\pi f_0 t + 0.5\alpha t^2) \quad (1.1)$$

Afin de compresser l'impulsion, on calcule la corrélation entre le signal émit et le signal reçu.

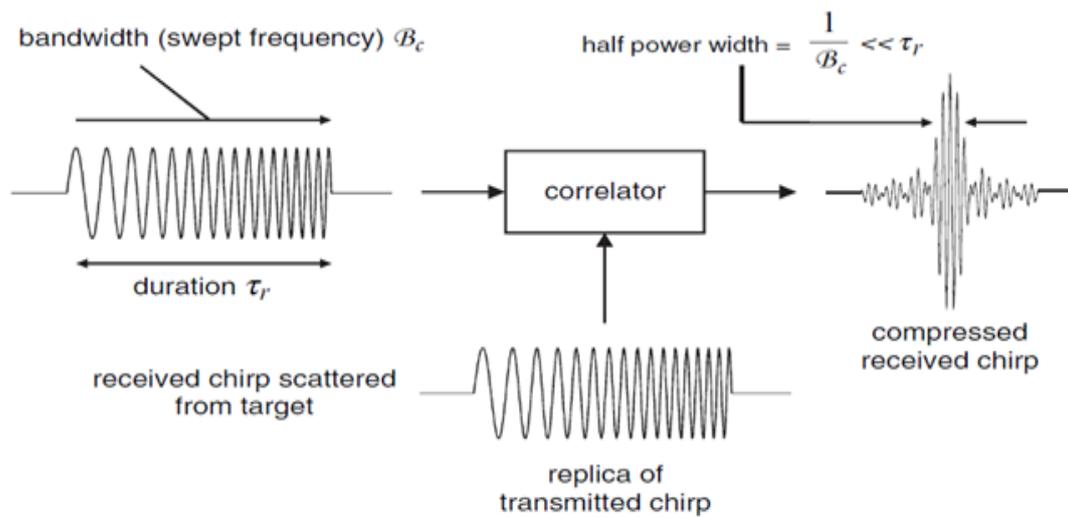


FIGURE 1.4 – Principe de la compression d'impulsion

1.4.2 Résolution Azimutale

Elle est pratiquement égale à la moitié de la longueur réelle en azimut de l'antenne utilisée et donnée par : $R_a = \frac{L}{2}$, tel que L est la longueur de l'antenne.

Pour améliorer la résolution en azimut, la manière la plus efficace serait d'augmenter la taille de l'antenne. La taille de l'antenne nécessaire serait trop élevée pour être réalisable en pratique. C'est ici qu'intervient le principe du système SAR : la synthèse d'ouverture. En intégrant les signaux rétrodiffusés pour chaque point le long de la trajectoire du système radar, on dispose donc d'une antenne synthétique virtuelle dont la taille est beaucoup plus élevée que celle de l'antenne réelle.

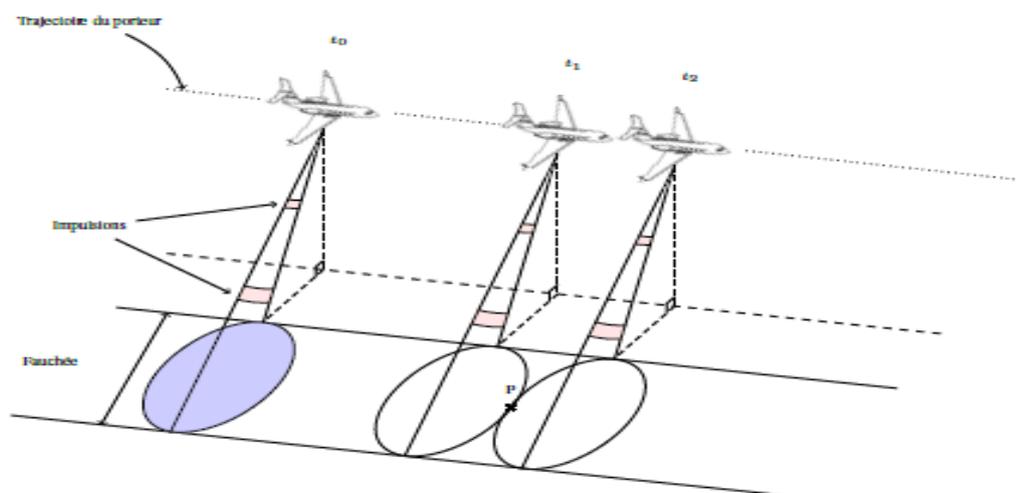


FIGURE 1.5 – Schéma d'un système SAR

Une image SAR comme tout type d'image est un ensemble de points accolés les uns à

côté des autres (pixel). La valeur de ces pixels dépend de l'intensité reçue et l'intensité émise en direction de la surface (rétrodiffusion radar).

1.5 Bruit de l'image SAR

Lorsque l'onde radar est émise en direction de la cible, celle-ci renvoie non pas une seule onde mais plusieurs ondes qui interfèrent au niveau du radar de façon destructive, ces interférences erronent les images produites par le SAR par un bruit appelé speckle. Le speckle a les caractéristiques d'un bruit multiplicatif. Pour bien extraire les informations thématiques et bien exploiter les images il faut réduire ou éliminer ce bruit. D'où le développement de nombreuses méthodes de réduction du speckle, parmi ces méthodes le filtrage des images SAR.

Chapitre 2

Méthode CFAR

2.1 Introduction

Le CFAR (Constant False Alarm Rate ou taux constant de fausses alertes) est un algorithme adaptatif utilisé dans les radars pour distinguer le signal retour d'une cible d'un bruit puissant, de brouillage aussi des interférences.

Dans la plupart des systèmes de détection CFAR à seuil variable, une estimation du niveau ou de la puissance du bruit de fond autour de la CUT (Cell Under Test) est faite pour calculer le seuil.

Pour cela, on prend les cellules voisines de la CUT et on calcule le niveau de puissance moyen. Pour éviter que les résultats soient faussés par la puissance reçue dans la CUT, puis on élimine du calcul les cellules adjacentes.

Pour affirmer la présence d'une cible dans la CUT, il faut que son signal soit supérieur aux cellules adjacentes et au niveau de la puissance moyen calculé, ce système de calcul s'appelle le CA-CFAR (Cell-Averaging CFAR) ou (CFAR par cellule moyennée).

D'autres systèmes sont basés sur le calcul de moyennes respectivement pour les cellules situées à droite et à gauche de la CUT considérée (Greatest Of-CFAR et Smallest Of-CFAR), les valeurs les plus hautes et les plus basses sont intégrées respectivement pour améliorer la détection dans le cas où on se trouve à proximité d'une source bruitée.

Un autre aspect de la détection CFAR est basé sur le classement des cellules de référence pour minimiser l'erreur d'estimation de la puissance du bruit en présence des cibles interférences (détecteurs OS-CFAR, CMLD-CFAR et TM-CFAR).

2.2 Les systèmes CFAR

2.2.1 CA-CFAR

Le paradigme des détecteurs CFAR est celui à niveau moyen, CA-CFAR. Ce dernier a été proposé pour la première fois en 1968 par les chercheurs américains Finn et Johnson.

Cet algorithme estime en temps réel la puissance du bruit qui est égale à la somme ou la moyenne des échantillons dans la fenêtre de référence comme montrer dans la Figure 2.1.

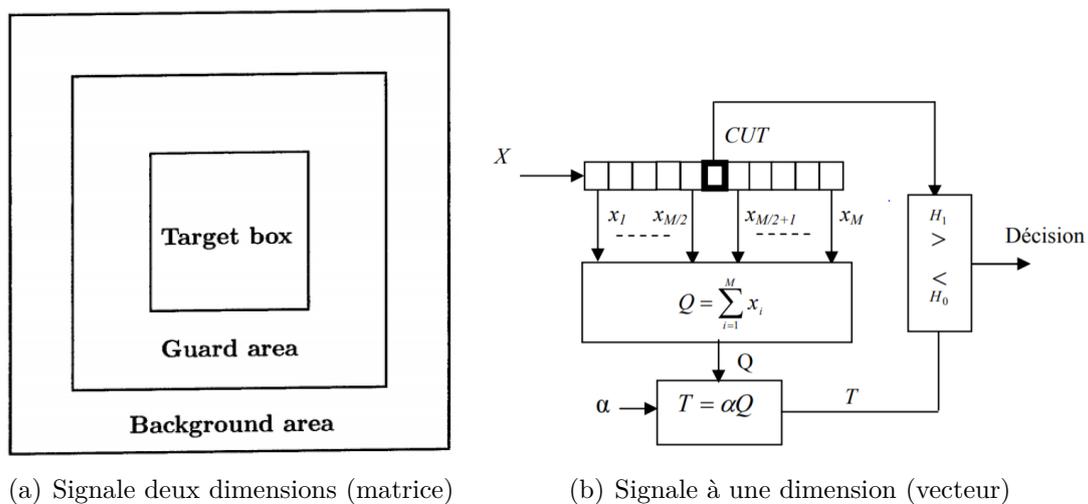


FIGURE 2.1 – Détecteur CA-CFAR pour un bruit Gaussien

2.2.2 L'architecture du CA-CFAR

Implémentation du moyennage des images SAR

Au début on récupère les données de l'image à traiter depuis un fichier et on les saisit sous forme de matrice 3D, puis on balaie l'image avec une fenêtre de taille (3X3)

Après avoir rajouter des zéros tout au tour de l'image pour prendre en considération tout les pixels on calcule la moyenne des parties balayées, ensuite il faut les stocker dans une matrice de taille égale à celle de l'image.

Affichage des ROI

Commençant par l'utilisation du logiciel ENVI pour marquer les ROI (region of interest)

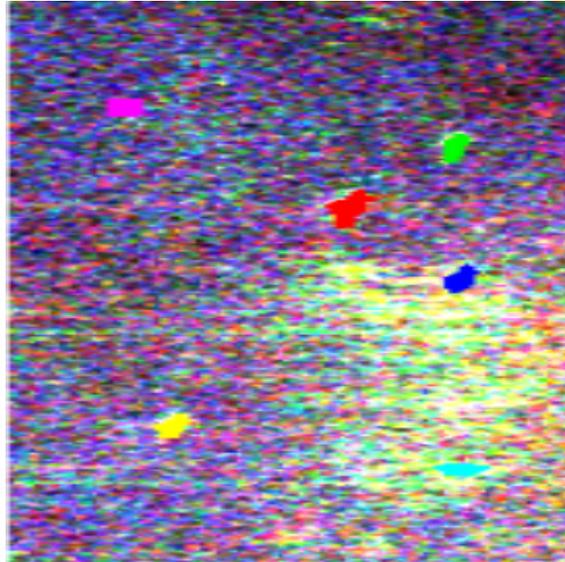


FIGURE 2.2 – Affichage de region of interest

Implémentation du CA-CFAR

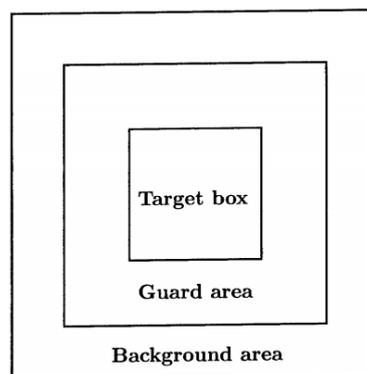


FIGURE 2.3 – Application du principe CA-CFAR

Donc, d'après la Figure 2.3, la région « Background area » est la région utilisée pour le calcul de seuil et prendre une décision sur la région « target box » (la fenêtre de la cible). Comme vous pouvez remarquer, il y a trois fenêtres une dans l'autre. Dans notre cas on va prendre la première fenêtre 15x15, la deuxième fenêtre 35x35 et la troisième fenêtre 55x55.

Pour prendre la décision sur le pixel central on calcule T qui est le seuil comme suit :

$$T = \mu_b \times \alpha \quad (2.1)$$

On va prendre la valeur de $\alpha = 2$, et μ_b est la moyenne des pixels qui sont dans "Background area".

Et tout cela sera récapituler dans la Figure 2.4.

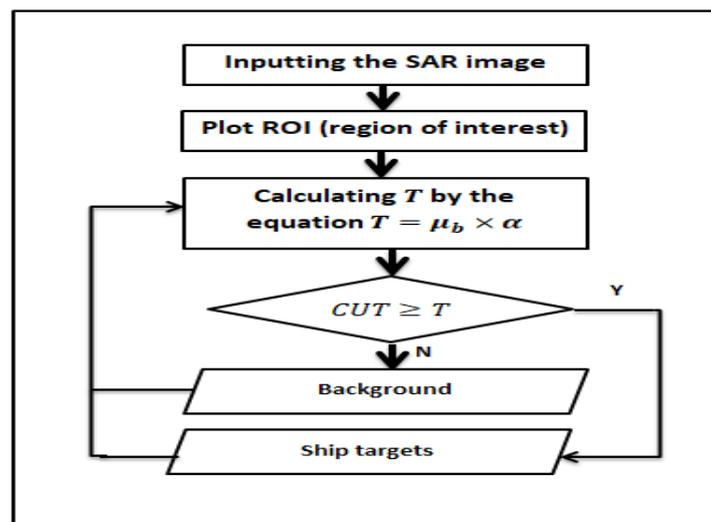


FIGURE 2.4 – L'architecture de l'algorithm CA-CFAR

2.3 Conclusion

La méthode CA-CFAR est parmi les premières méthodes qui ont été proposé pour mettre en valeur l'algorithm CFAR , elle permet à atteindre des résultats très satisfaisantes mais elle n'est pas bien adapté pour les environnement hétérogène.

Chapitre 3

Réseaux de neurones

3.1 Définition

3.1.1 Origine biologique

Deux d'approche ont été adoptées en recherche en Intelligence Artificielle :

- Procéder d'abord à l'analyse logique des tâches relevant de la cognition humaine et tenter de les reconstituer par programme. C'est cette approche qui a été privilégiée par l'Intelligence Artificielle symbolique et la psychologie cognitive classique. Cette démarche est étiquetée sous le nom de cognitivisme.
- Puisque la pensée est produite par le cerveau ou en est une propriété, commencer par étudier comment celui-ci fonctionne. C'est cette approche qui a conduit à l'étude des réseaux de neurones formels. On désigne par connexionnisme la démarche consistant à vouloir rendre compte de la cognition humaine par des réseaux de neurones.

La seconde approche a donc mené à la définition et à l'étude de réseaux de neurones formels qui sont des réseaux complexes d'unités de calcul élémentaires interconnectées.

Le modèle du neurone biologique a été affiné par ces découvertes (McCulloch,1943) :

- Van Leuwenhook (1718) : première description fidèle de ce qu'on appellera plus tard les axones.
- Dutrochet (1824) : observation du corps cellulaire des neurones.
- Valentin : découverte des dendrites.
- Deiters (1865) : image actuelle de la cellule nerveuse.

- Sherrington (1897) : les synapses.
- Les neuro-transmetteurs (première moitié du 20ème siècle).

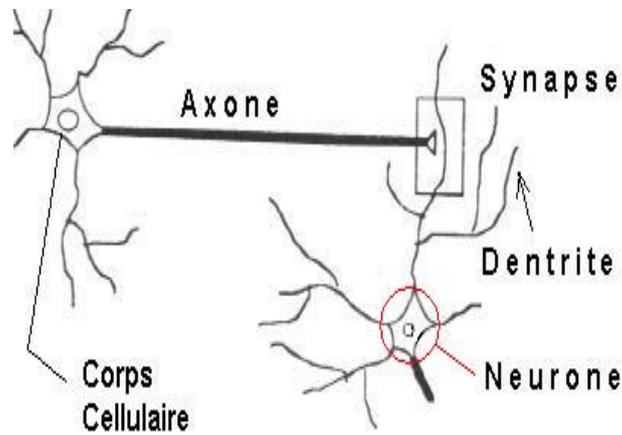


FIGURE 3.1 – *Modèle du neurone biologique*

Des dispositifs nommés neurones formels inspirés du modèle biologique (Figure 3.1) sont utilisés en informatique pour modéliser et résoudre des problèmes complexes. Ils comportent N entrées ayant chacune un poids synaptique, une fonction de sortie, et une sortie qui sert à son tour d'entrée à d'autres neurones similaires. La Figure 3.2 montre une comparaison entre le neurone biologique et le neurone artificiel. Dans la cellule nerveuse humaine, la synapse correspond au poids d'un neurone artificiel, le corps cellulaire à une fonction de transfert et l'axone à un élément de sortie.

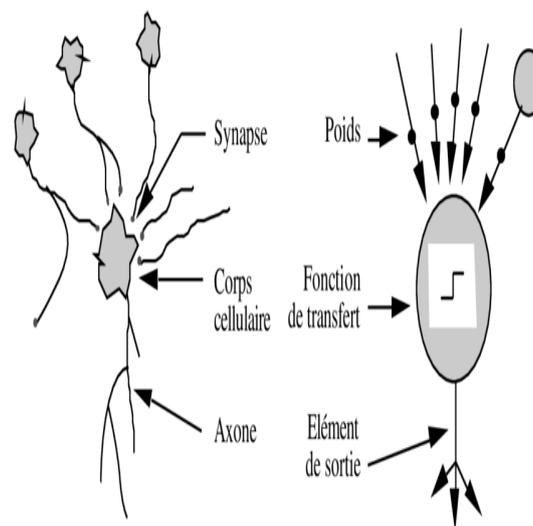


FIGURE 3.2 – *La comparaison entre un neurone biologique et un neurone artificiel*

Tel que la sortie du neurone est donnée par :

$$n = \sum_{i=0}^{n-1} w_{i,j} p_i + b \quad (3.1)$$

$$\alpha = f(n) \quad (3.2)$$

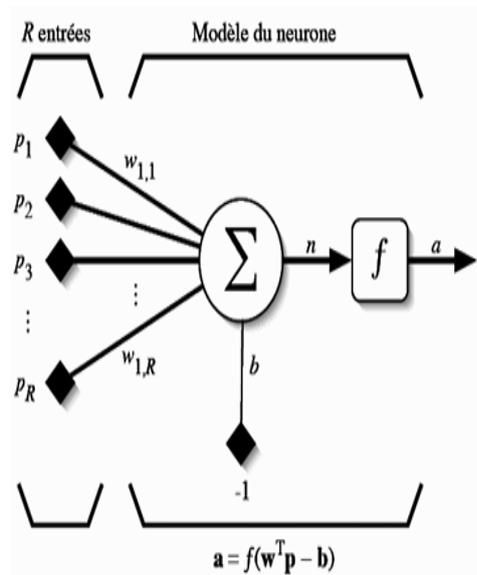


FIGURE 3.3 – Modèle d'un neurone artificiel

Le réseau de neurones sous la forme matricielle sera :

$$\alpha = f(n) = f(\mathbf{w}\mathbf{p} + b) \quad (3.3)$$

\mathbf{W} est la matrice des poids des connexions du neurone de dimension $R * 1$

Où R est le nombre d'entrées du réseau.

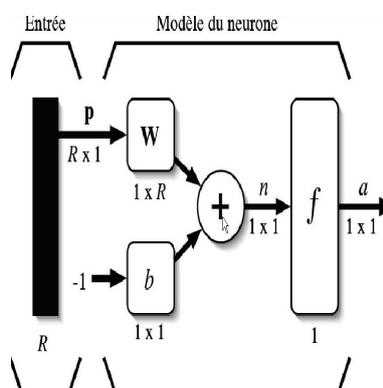


FIGURE 3.4 – Représentation matricielle du modèle d'un neurone artificiel

3.2 La structure des réseaux de neurones

En général on peut dire que les réseaux de neurones peuvent être caractérisés comme des modèles de calcul apprenant, généralisant et organisant des données. Un réseau de neurone artificiel contient un grand nombre d'unités, les neurones, qui communiquent entre eux en s'envoyant des signaux à travers des liens, appelées connexions synaptiques.

Le système de neurone possède trois types de neurones :

- les neurones d'entrée qui reçoivent les données.
- les neurones de sortie qui envoient les données par la sortie du système.
- les neurones cachés, dont les signaux d'entrée et de sortie demeurent dans le système.

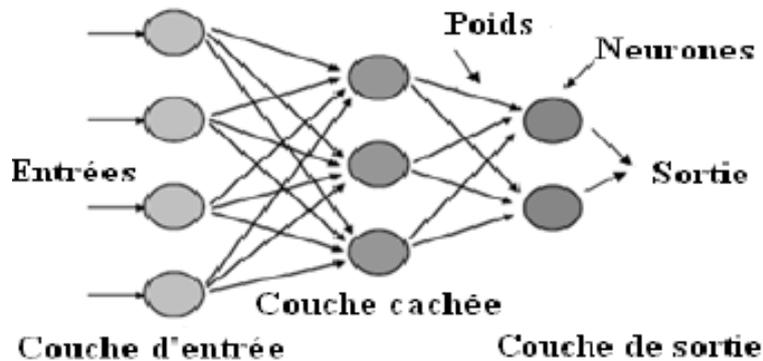


FIGURE 3.5 – Schéma d'un réseau de neurone

3.3 Architecture des Réseaux de Neurones

3.3.1 Réseaux de neurones a une seule couche

Un réseau de neurones est un maillage de plusieurs neurones, généralement organisé en couches. Pour construire une couche de S neurones, il s'agit juste de les assembler.

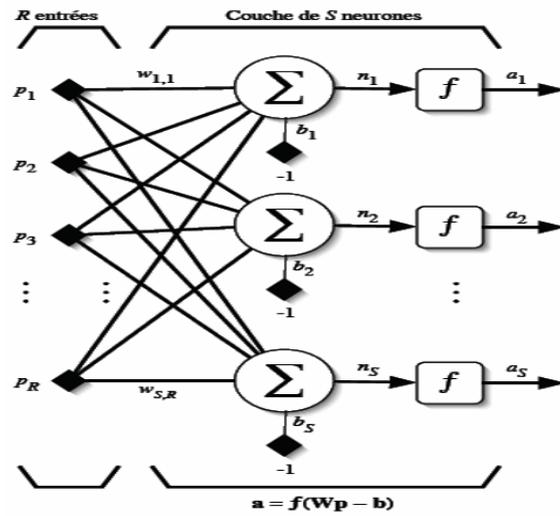


FIGURE 3.6 – Réseaux de neurones à une seule couche

Les S neurones d’une même couche sont tous branchés aux R entrées, on dit alors que la couche est totalement connectée. On obtient la matrice w comme suit :

$$W = \begin{bmatrix} W_{1,1} & \cdots & W_{1,R} \\ \vdots & \ddots & \vdots \\ W_{S,1} & \cdots & W_{S,R} \end{bmatrix} \tag{3.4}$$

Chaque ligne représente le numéro d’une entrée du réseau et chaque colonne représente le numéro d’un neurone de la couche.

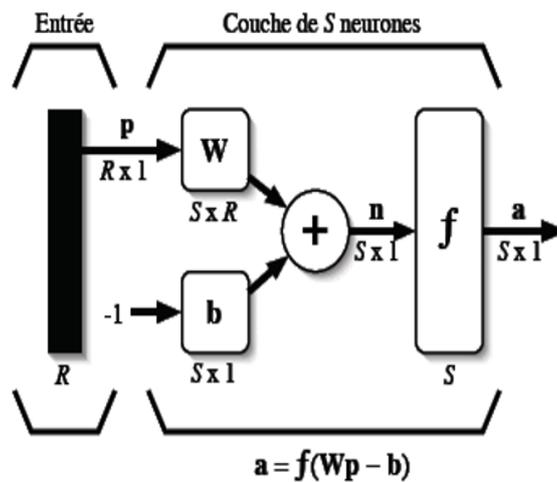


FIGURE 3.7 – Représentation matricielle d’un réseau de neurone à une couche de S neurones

3.3.2 Réseau de Neurone multi-couche

On peut réaliser des réseaux de neurones plus puissants en augmentant le nombre de couche, on enfile autant de couche que l'on veut on fixe aussi un nombre quelconque de neurones sur chaque couche. On distingue deux types de réseau a plusieurs couches : les réseaux statiques et les réseaux dynamiques.

Réseaux statiques

Les neurones de chaque couche sont reliés avec les neurones de la couche suivante, pas avec ceux de la couches précédentes, ils contiennent trois parties, l'entrée, les couches cachées, et la sortie comme illustre la figure suivante :

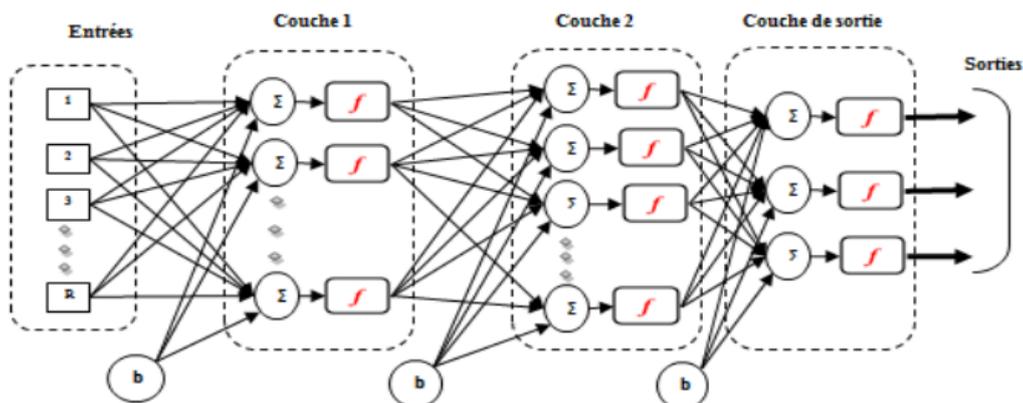


FIGURE 3.8 – Structure d'un réseaux de neurones statique a trois couches

Ce réseau de neurones est caractérisé par le nombre de couches cachées, le nombre de neurones et leurs fonctions d'activation. La forme matricielle de ce réseau :

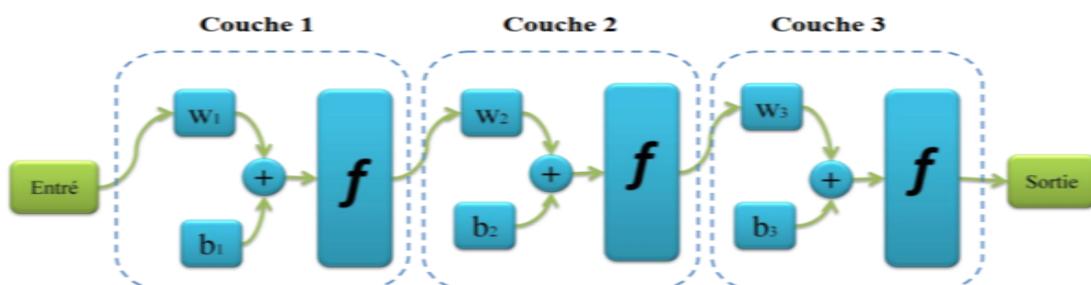


FIGURE 3.9 – Matrice d'un réseau de neurones statique a trois couches

$p(t)$ est la matrice d'entrée du ce réseau a chaque instant t :

$$p(t) = \begin{bmatrix} Y_{1,t} \\ Y_{2,t} \\ \vdots \\ Y_{l,t} \end{bmatrix} \quad (3.5)$$

la sortie des couches 1,2 et 3 est respectivement donnée :

$$a_1(t) = f_1[(W_1 \times p(t)) + b_1] \quad (3.6)$$

$$a_2(t) = f_2[(W_2 \times a_1) + b_2] \quad (3.7)$$

$$a_3(t) = f_3[(W_3 \times a_2) + b_3] \quad (3.8)$$

Où f_1, f_2, f_3 sont les fonctions d'activation des neurones des couches 1, 2 et 3.

les matrices des poids d'un réseau de neurones complètement maillé, allant de la couche i vers la couche j sera de dimension $C_j \times C_i$, C_j est le nombre de neurones de la couche j et C_i est le nombre de neurones de la couche i . $w_{x,y}$ est le poids de connexion entre le neurone x de la couche j et le neurone y de la couche i .

Réseaux Dynamiques

Un réseau de neurones dynamique est un réseau en couches qui possède des retours unitaires dans les couches cachées.

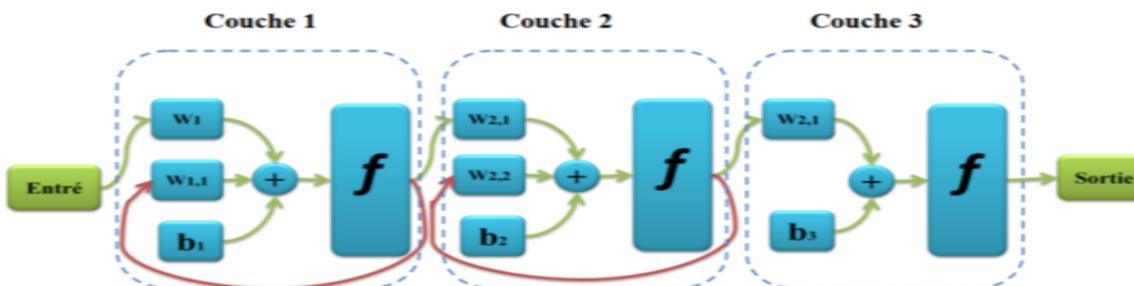


FIGURE 3.10 – Matrice d'un réseaux de neurones dynamique a trois couches

Une connexion arrière est un registre a décalage qui permet d'introduire l'état précédent des sorties des neurones dans le calcul présent.

La sortie des couches 1,2 et 3 est respectivement donnée :

$$f_1 = [(W_1 \times p(t)) + (W_1 \times \alpha_1(t - 1)) + b_1] \quad (3.9)$$

$$f_2 = [(W_2 \times \alpha_1(t)) + (W_2 \times \alpha_2(t - 1)) + b_2] \quad (3.10)$$

$$f_3 = [(W_3 \times \alpha_2(t)) + (W_3 \times \alpha_3(t - 1)) + b_3] \quad (3.11)$$

3.4 Apprentissage des Réseaux de Neurones

L'apprentissage d'un réseau de neurone peut être considéré comme une action de la mise à jour de ses poids des connexions synaptiques, afin de résoudre le problème demandé. L'apprentissage est la caractéristique principale des réseaux de neurones et il peut se faire de différentes manières et selon différentes règles. On peut distinguer deux types d'apprentissage : l'apprentissage supervisé et l'apprentissage non-supervisé

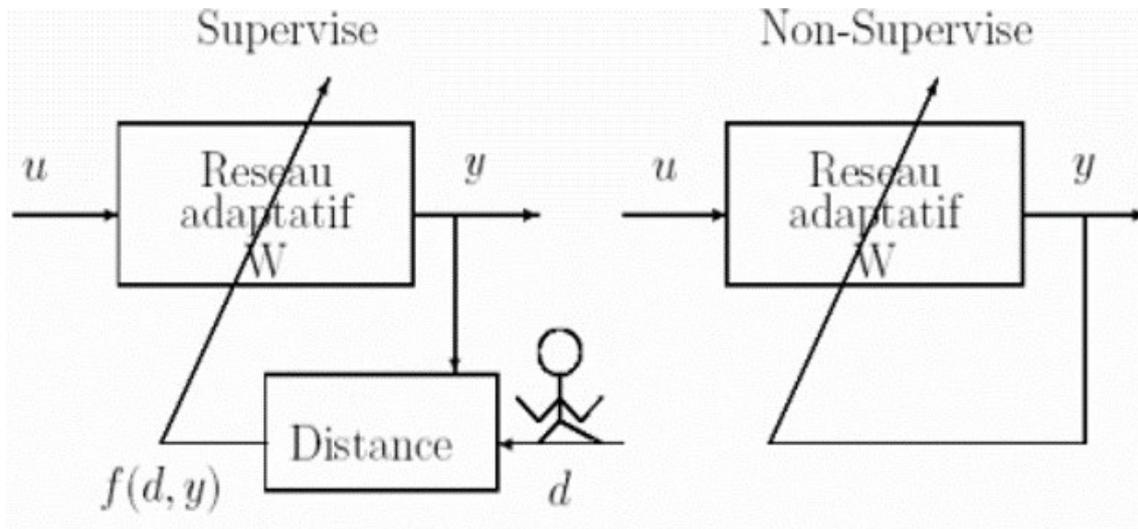


FIGURE 3.11 – Type d'apprentissage en réseau de neurone.

3.4.1 L'apprentissage supervisé

L'apprentissage supervisé ou l'apprentissage associatif : le réseau adaptatif W compare le résultat « y » qu'il a calculé, en fonction $f(d,y)$ des entrées u fournies, et la réponse d attendue en sortie. Ainsi le réseau va se modifier jusqu'à ce qu'il trouve la bonne sortie d , c'est-à-dire celle attendue, correspondante à une entrée u donnée.

3.4.2 L'apprentissage non-supervisé

L'apprentissage non-supervisé ou auto organisation : l'apprentissage est basé sur des probabilités. Le réseau adaptatif W va se modifier en fonction des régularités statistiques de l'entrée u et établir des catégories, en attribuant et en optimisant une valeur de qualité, aux catégories reconnues.

Les deux types d'apprentissage ont pour but d'ajuster les poids de connexions entre les neurones, en s'accordant de certaines règles. Nous citons ci-après les règles les plus répondues :

La règle de Hebb, (Hebb, 1949) :

Cette règle permet de modifier la valeur des poids synaptiques en fonction de l'activité des unités qui les relient. Le but principal est le suivant : si deux unités s'activent en même temps la connexion qui les lie est renforcée (c'est une connexion excitatrice) sinon elle est affaiblie (c'est une connexion inhibitrice)

La règle delta :

Elle calcule la différence entre la valeur de la sortie et la valeur désirée pour ajuster les poids synaptiques. Elle emploie une fonction d'erreur, nommée aussi le moindre carré moyen, basée sur les différences utilisées pour l'ajustement des poids.

La règle d'apprentissage compétitive :

Elle ne concerne qu'un seul neurone. On regroupe les données en catégorie. Les neurones similaires vont donc être rangés dans une même classe en se basant sur des corrélations des données et seront représentés par un seul neurone.

L'architecture d'un tel réseau possède une couche d'entrée et une couche de compétition. Une forme est présentée à l'entrée du réseau et est projetée sur chacun des neurones de la couche compétitive. Le neurone gagnant est celui qui possède un vecteur de poids le plus proche de la forme présentée à l'entrée.

Chaque neurone de sortie est connecté aux neurones de la couche d'entrée et aux autres cellules de sortie (c'est une connexion inhibitrice) ou à elle-même (c'est une connexion excitatrice). La sortie dépend alors de la compétition entre les connexions inhibitrices et excitatrices .

La règle de corrélation en cascade (Fahlman, 1990) :

C'est une technique d'apprentissage qui ajoute progressivement des neurones cachés au réseau jusqu'à ce que l'effet bénéfique de ces nouveaux neurones ne soit plus perceptible.

La règle de rétro-propagation :

Elle s'utilise pour ajuster les poids de la couche d'entrée à la couche cachée. Les poids dans le réseau de neurones sont au préalable initialisés avec des valeurs aléatoires.

On considère ensuite un ensemble de données qui vont servir à l'apprentissage. Chaque échantillon possède ses valeurs cibles qui sont celles que le réseau de neurones doit à terme prédire lorsqu'on lui présente le même échantillon.

3.5 Les réseaux de neurones convolutifs

3.5.1 Définition

En apprentissage automatique, un réseau de neurones convolutifs ou réseau de neurones à convolution (CNN Convolutional Neural Networks) est un type de réseau de neurones artificiels acycliques dans lequel le motif de connexion entre les neurones est inspiré d'un cortex visuel des animaux. Leur fonctionnement est inspiré des processus biologiques, ils consistent en un empilage multicouche de perceptrons dont le but est de pré-traiter de petites quantités d'informations.

Les réseaux neuronaux convolutifs ont de larges applications dans la reconnaissance d'image et vidéo, les systèmes de recommandation et le traitement du langage naturel.

On distingue deux parties, une première partie que l'on appelle la partie convolutive du modèle et la seconde partie, que l'on va appeler la partie classification du modèle qui correspond à un modèle MLP (Multi Layers Perceptron).

C'est un réseau de neurone multicouche et plus précisément c'est un réseau profond composé de multiples couches qui en générale sont organisées en blocs.

3.5.2 L'architecture des réseaux neurones convolutifs

Une architecture de réseau de neurones convolutifs est formée par un empilement de couches de traitement, les parties les plus importantes dans cette architecture sont :

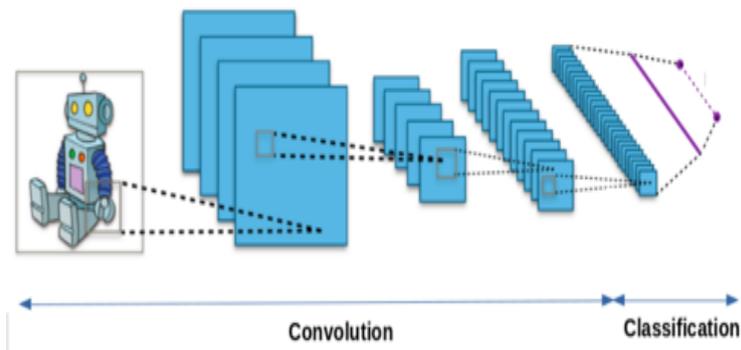


FIGURE 3.12 – Architecture d'un CNN (Convolutional Neural Network)

- 1-la couche de convolution (CONV) qui traite les données d'un champ récepteur.
- 2-la couche de pooling (POOL), qui permet de compresser l'information en réduisant la taille de l'image intermédiaire (souvent par sous-échantillonnage).
- 3-la couche de correction (ReLU), souvent appelée par abus « ReLU » en référence à la fonction d'activation (Unité de rectification linéaire) .
- 4-la couche « entièrement connectée » (FC :fully-connected), qui est une couche de type perceptron.
- 5-la couche de perte (LOSS)

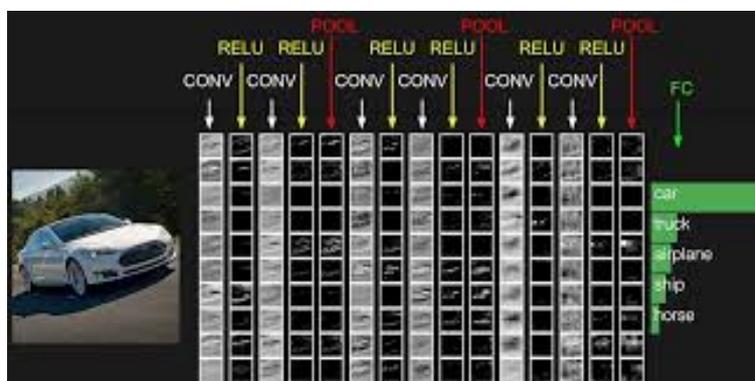


FIGURE 3.13 – L'architecture d'un réseau neurone convolutif

La couche de convolution

La couche de convolution est le bloc de construction de base d'un réseau de neurone convolutif et la composante clé des réseaux de neurones convolutifs.

Son but est de repérer la présence d'un ensemble de features dans les images reçues en entrée. Pour cela, on réalise un filtrage par convolution le principe est de faire "glisser" une fenêtre représentant la feature sur l'image, et de calculer le produit de convolution

entre la feature et chaque portion de l'image balayée. Une feature est alors vue comme un filtre.

Trois hyperparamètres permettent de dimensionner le volume de la couche de convolution :

- Profondeur de la couche : nombre de noyaux de convolution (ou nombre de neurones associés à un même champ récepteur).
- 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.
- La marge (à 0) ou zero 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.

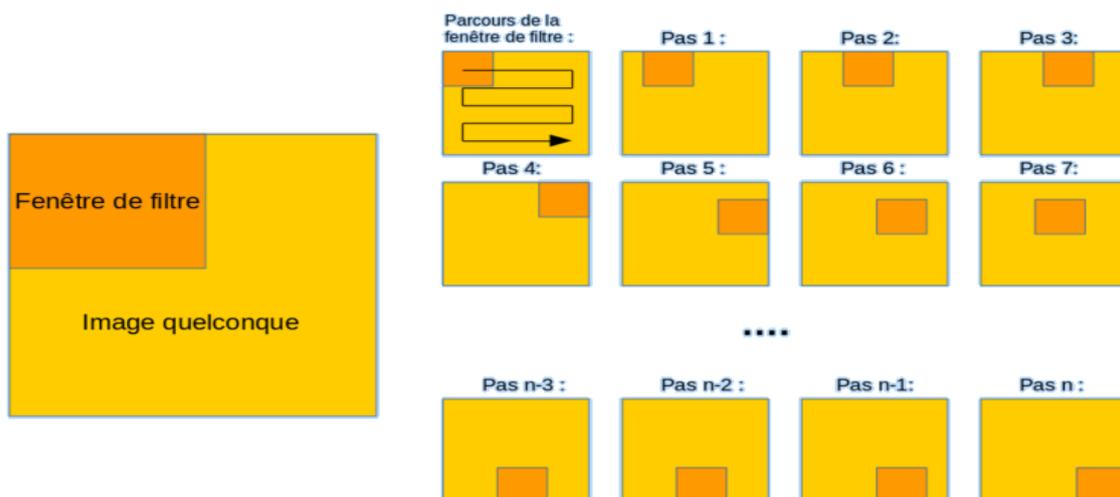


FIGURE 3.14 – Schéma du parcours de la fenêtre de filtre sur l'image.

3.5.3 La couche de pooling (POOL)

Un autre concept important des réseaux de neurones convolutifs est le pooling (« mise en commun »), ce qui est une forme de sous-échantillonnage de l'image souvent placé entre deux couches de convolution. L'image d'entrée est découpée en une série de rectangles de n pixels de côté ne se chevauchant pas (pooling). Chaque rectangle peut être vu comme une tuile. Le signal en sortie de tuile est défini en fonction des valeurs prises par les différents pixels de la tuile.

L'opération de pooling (ou sub-sampling) consiste à réduire la taille des images, tout en préservant leurs caractéristiques importantes. Pour cela, on découpe l'image en cellules régulières, puis on garde au sein de chaque cellule la valeur maximale. En pratique, on utilise souvent des cellules carrées de petite taille pour ne pas perdre trop d'informations. Les choix les plus communs sont des cellules adjacentes de taille 2×2 pixels qui ne se chevauchent pas, ou des cellules de taille 3×3 pixels, distantes les unes des autres d'un pas de 2 pixels (qui se chevauchent donc). On obtient en sortie le même nombre de feature maps qu'en entrée, mais celles-ci sont bien plus petites.

Il est courant d'insérer périodiquement une couche Pooling entre les couches Conv successives dans une architecture ConvNet. Sa fonction est de réduire progressivement la taille spatiale de la représentation afin de réduire la quantité de paramètres et de calculs dans le réseau, et donc de contrôler également le sur-ajustement. La couche de pooling fonctionne indépendamment sur chaque tranche de profondeur de l'entrée et la redimensionne spatialement, en utilisant l'opération MAX, Figure 3.15.

La couche de pooling permet de réduire le nombre de paramètres et de calculs dans le réseau. On améliore ainsi l'efficacité du réseau et on évite le sur-apprentissage. Ainsi, la couche de pooling rend le réseau moins sensible à la position des features , le fait qu'une feature se situe un peu plus en haut ou en bas, ou même qu'elle ait une orientation légèrement différente ne devrait pas provoquer un changement radical dans la classification de l'image.

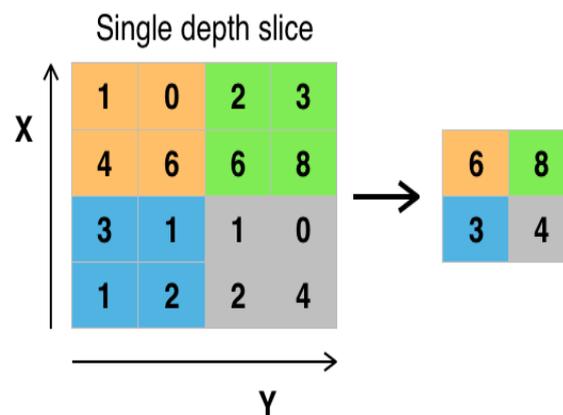


FIGURE 3.15 – Max pooling avec un filtre 2×2 et un pas de 2.

La couche de correction (ReLU)

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 Unité Linéaire Rectifiée) : $f(x) = \max(0, x)$. 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 : $f(x) = |\tanh|$
- La correction par la fonction sigmoïde : $f(x) = (1 + e^{-x})^{-1}$.

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.

La couche de correction ReLU remplace donc toutes les valeurs négatives reçues en entrées par des zéros. Elle joue le rôle de fonction d'activation.

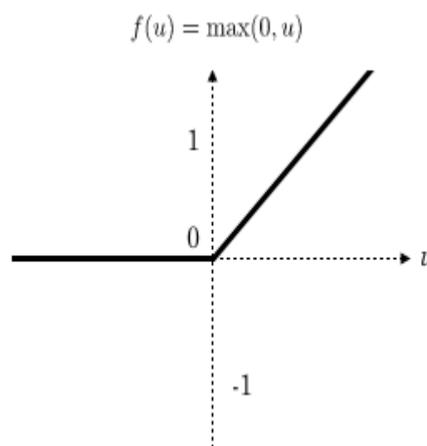


FIGURE 3.16 – La fonction ReLU

Couche entièrement connectée (FC :fully-connected)

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.

Couche de perte (LOSS)

La couche de perte spécifie comment l'entraînement du réseau pénalise l'écart entre le signal prévu et réel. Diverses fonctions de perte adaptées à différentes tâches peuvent y être utilisées telle que la perte « Softmax », la perte « euclidienne » et la perte par « entropie croisée sigmoïde » .

3.5.4 Réseaux de neurones de convolution pour la détection d'objets

De nombreux algorithmes ont déjà vu le jour : YOLO, R-CNN, Fast R-CNN, Faster R-CNN, SSD, RetinaNet ...

R-CNN

Le R-CNN est une architecture de détection d'objet. Le R-CNN commence par extraire des régions intéressantes de l'image, puis il utilise ces régions comme données d'entrée pour un CNN.

Cette séparation en régions permet de détecter plusieurs objets de plusieurs classes différentes dans une même image. Cette solution proposée par Girshick et al.

Dans le R-CNN présentée, les régions sont extraites grâce à une recherche sélective proposée par Uijlings, van de Sande, Gevers et Smeulders, 2013. Celle-ci utilise la structure de l'image et plusieurs techniques de partitionnement pour récupérer toutes les régions intéressantes possibles.

Faster R-CNN

Le Faster R-CNN est une amélioration du R-CNN dans sa précision et sa rapidité à l'entraînement [1].

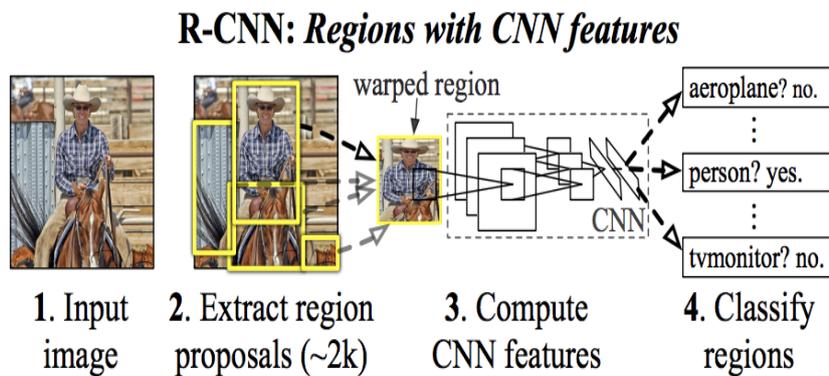


FIGURE 3.17 – Architecture R-CNN

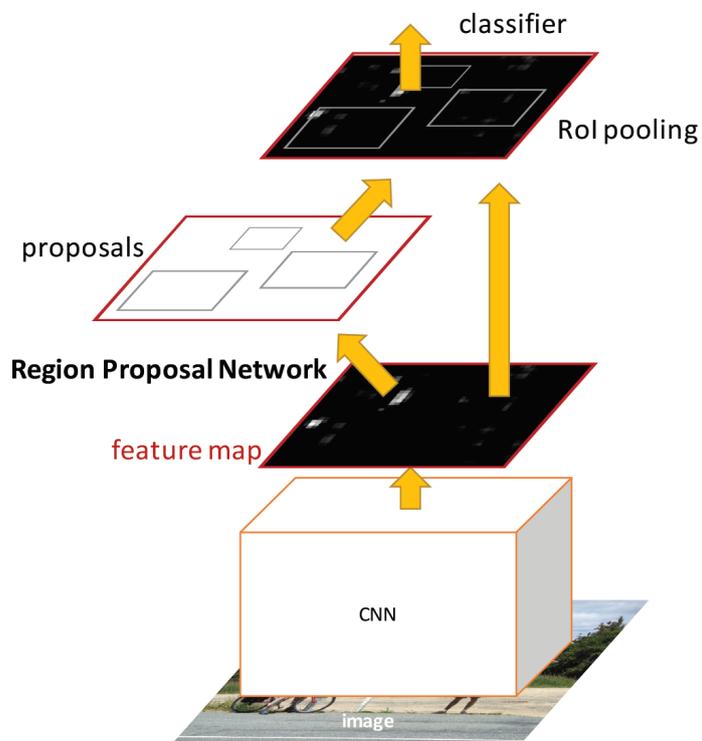


FIGURE 3.18 – Architecture Faster R-CNN [1].

L'architecture Faster R-CNN utilise la même carte de caractéristiques résultant des couches de convolution pour générer les régions intéressantes et pour ensuite les classifier.

Le réseau de propositions de régions utilise des fenêtres coulissantes de tailles et ratio différents pour analyser la carte de caractéristiques en profondeur. Ces changements améliorent significativement la précision et la rapidité de l'architecture comparée au R-CNN [1].

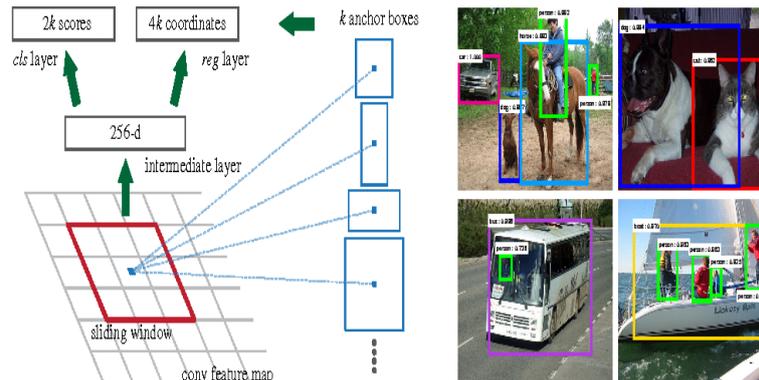


FIGURE 3.19 – Faster R-CNN : Fenêtre coulissante et résultat [2].

YOLO : You Only Look Once

YOLO est un système de détection d'objets en temps réel à la fine pointe de la technologie. Cet algorithme développé par J.Redmon et al [4], repose sur deux étapes qui sont appliquées sur des images de taille prédéfinie lors de l'apprentissage comme :

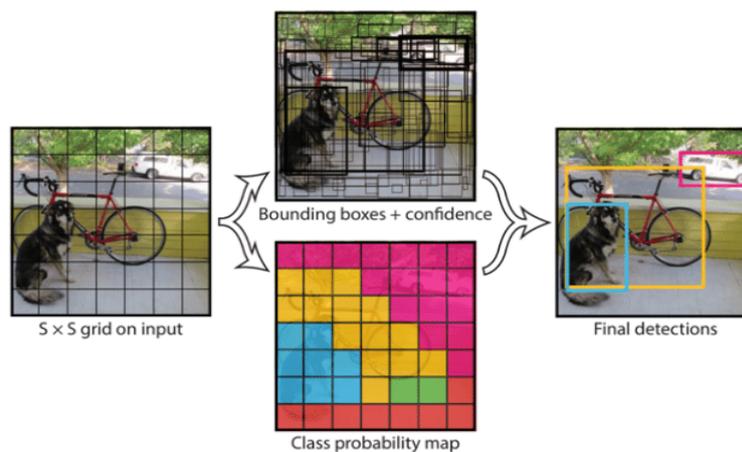


FIGURE 3.20 – Algorithme YOLO

-Une détection d'objets opérée par des réseaux de neurones de convolution.

-Un quadrillage de l'image où l'on prédit la classe de l'objet s'il existe.

L'avantage de cette prédiction est qu'elle peut se faire indépendamment de la première détection mais qu'elle ne cible pas les objets en particulier ; il est donc plus difficile de détecter les plus petits objets ainsi que les objets se chevauchant.

La dernière version YOLOv3 proposée par Redmon et Farhadi en 2018 [5], l'une permettant une classification plus rapide mais avec une perte sur l'erreur de détection. Les contraintes matérielles nous ont imposés pour le moment à n'utiliser que la version la moins performante.

SSD : Single Shot Detector

-On utilise des filtres convolutifs pour estimer directement la probabilité de classification et la boîte de délimitation

-On le fait avec plusieurs résolutions (différentes couches convolutives) pour estimer des objets de tailles différentes

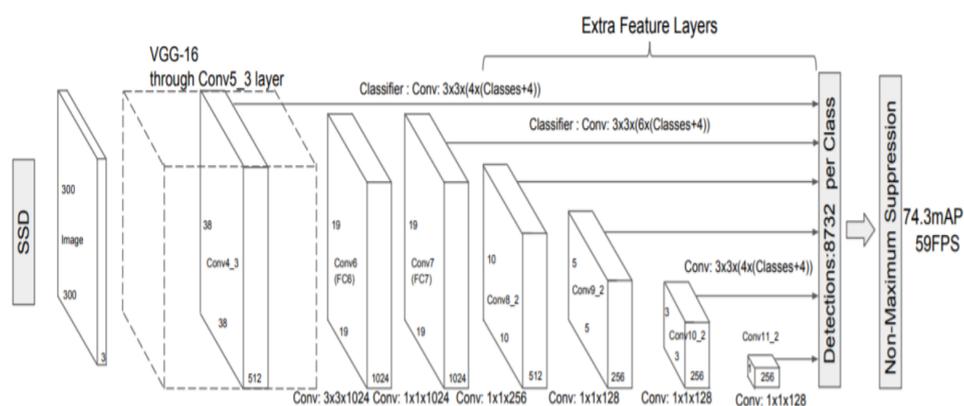


FIGURE 3.21 – Single shot detector

On utilise le même réseau, paramètres et données d'entraînement, SSD obtient une performance inférieure, mais il est beaucoup plus rapide

3.6 conclusion

Dans le domaine de traitement d'images, les réseaux neuronaux permettent de développer rapidement une solution à des problèmes très variés avec une bonne précision, Dont la méthode Mask R-CNN.

Chapitre 4

Analyses et Résultats

4.1 Présentation des données

Notre but est de détecter et surveiller les navires, nous avons utilisé l'intelligence artificielle avec les réseaux de neurones. Pour faire l'apprentissage des réseaux de neurones on a utilisé la base des données des images SAR [6], elle contient 21 237 images captées par le satellite Sentinel1 dont :

- 650 images de polarisation VV (émission verticale, réception verticale)
- 10 293 images de polarisation HH (émission horizontale, réception horizontale)
- 10 294 images de polarisation HV (émission horizontale, réception verticale)

Tel que chaque image est importé avec son annotation, qui indique les positions des navires dans l'images. La Figure 4.1 donne un exemple d'une image qui contient un navire et une capture sur le fichier XML correspondant à cette image avec les coordonnées sur la position de navire.

Pour l'implémentation des méthodes de détection et la manipulation des images on a utilisé le langage de programmation Python qui est très recommandé dans la communauté scientifique et dans le domaine industriel, qui est aussi un langage important pour l'intelligence artificielle. Toutes les opérations de manipulation des images SAR, ont été réaliser avec le langage de programmation Python.



FIGURE 4.1 – Forme des données images et annotations

4.2 Transfert Learning

Les domaines d'application du Transfer Learning sont nombreux. Principalement, les méthodes de transfert de connaissance sont très souvent utilisées pour la reconnaissance d'image ainsi que le traitement automatique du langage. Ces deux domaines d'apprentissage sont très complexes et chronophages.

C'est pour cela que le Transfer Learning apporte un souffle nouveau pour tenter d'optimiser ces traitements en exploitant au maximum des modèles déjà entraînés.

Il est possible maintenant de trouver des réseaux de neurones disponibles sur internet entraînés par d'autres personnes. Ainsi la librairie Caffe permet à des utilisateurs de partager les poids de leurs réseaux de neurones qui permettront à d'autres personnes de gagner du temps en réutilisant ces mêmes poids dans leur modèle.

4.2.1 Transfert Learning pour la détection des navires

Mask R-CNN

Nous présentons un cadre conceptuel simple, flexible et général pour la détection d'objets. Notre approche détecte efficacement les objets d'une image tout en générant simultanément un masque de segmentation de haute qualité pour chaque instance. La méthode, appelée Mask R-CNN[7], étend Faster R-CNN[1] en ajoutant une branche pour prédire un masque d'objet en parallèle avec la branche existante pour la reconnaissance de la boîte de délimitation. Pour la réalisation de notre algorithme de détection nous avons utilisé la librairie Mask RCNN [7].

Nous présentons les meilleurs résultats dans les trois pistes de la suite de défis COCO, y compris la segmentation des instances, la détection des objets de zone de délimitation et la détection de points clés pour la personne [7].

Dans notre cas le modèle dont on a utilisé pour détecter les navires a été conçu sur la base des données Microsoft COCO (Common Objects in Context)[8], et pour cela on a fait appel aux méthodes de transfert learning pour aboutir à notre objectif en utilisant le fichier des poids d'un réseau Mask RCNN avec un apprentissage sur la base des données COCO (mask-rcnn-coco.h5) et par la suite on a adapté ce réseau pour la détection des navires dans les images Sentinel1.

4.3 Détection des navires avec le deep learning

La détection des navires avec un radar à synthèse d'ouverture repose sur la détection directe du navire ou de son sillage.

Durant les premières exécutions, on a souvent rencontré des problèmes concernant les boxes et la détection des navires au bord de la cote donc il fallait augmenter le nombre d'images utilisées dans l'apprentissage ainsi que le nombre d'époch pour avoir des meilleurs résultats. Comme vous pouvez le constater dans la Table 4.1.

D'après les figures 2, 4 et 5 de la Table 4.1, la polarisation qui fournit des résultats assez satisfaisants est la HV, mais il est préférable de combiner entre les différentes polarisations pour que le modèle soit adapter à toutes sortes d'images.

Pour la détection des navires, on privilégie l'imagerie dans la bande à la polarisation HH, puisque le contraste entre le navire et l'eau est habituellement plus élevé dans cette polarisation.

En effet, à cause des ondes capillaires, la rétrodiffusion de la surface de l'océan est plus forte dans le canal VV, ce qui produit un fouillis de fond plus intense que dans la polarisation HH. En contrepartie, on choisit la polarisation VV pour détecter les sillages, puisque le sillage est rarement visible dans les images polarisées [9], comme indiquer sur la Figure 4.2.

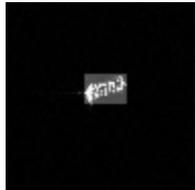
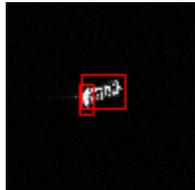
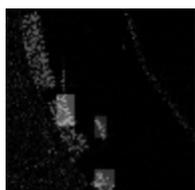
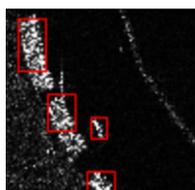
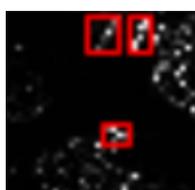
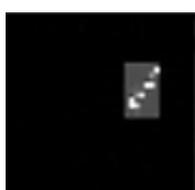
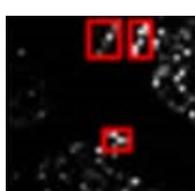
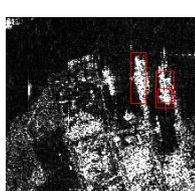
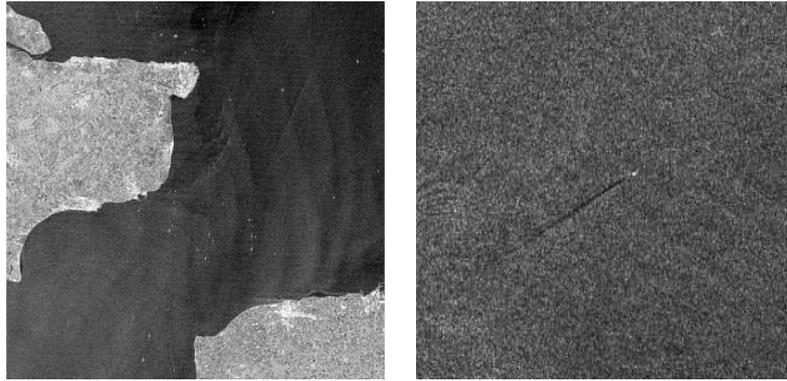
Manipulation	Actuel	Predicted
Sen Ship VV Train=400 Test=500 Epoch=3 Train mAP=0.866 Test mAP=0.881		
Sen Ship VV Train=390 Test=10 Epoch=10 Train mAP=0.801 Test mAP=0.830		
Sen Ship HH Train=500 Test=9793 Epoch=3		
Sen Ship HH Train=500 Test=9793 Epoch=6 Train mAP=0.845 Test mAP=0.753		
Sen Ship HV Train=500 Test=9794 Epoch=10		
Sen Ship 2004 HV 2000 HH 596 VV Train=4500 Test=100 Epoch=6		
Sen Ship 2004 HV 2000 HH 596 VV Train=4500 Test=100 Epoch=10		
Sen Ship 201 HV 600 HH 201 VV Epoch=6 Train=500 Test=502 Train mAP=0.55 Test mAP=0.475		

TABLE 4.1 – Les images traitées



(a) Image HH dans la bande C, collectée par RADARSAT-1 (b) Image VV, collectée par ERS-1

FIGURE 4.2 – *Détection de navires.*

On prévoit que le recours aux données multi-polarimétriques ou complètement polarimétriques permettront d'exploiter différentes combinaisons de polarisation et optimiser la détection des navires. Par exemple, une combinaison de polarisation VV et VH serait optimale pour la surveillance des navires, puisque le canal VH fournit des informations sur les cibles ponctuelles sur un fond de fouillis très sombre. En parallèle, la polarisation VV, fournirait un signal rétro-diffusé adéquat permettant l'analyse des sillages. Puisque le mécanisme de diffusion entre le navire et la mer est une double réflexion.

mAP : Mean Average Precision

AP (précision moyenne) est une métrique populaire pour mesurer la précision de détecteurs d'objets comme Faster R-CNN[1], SSD[10], etc. La précision moyenne calcule la valeur de précision moyenne pour la valeur de rappel sur 0 à 1.

La précision mesure la précision de vos prévisions. c'est-à-dire que le pourcentage de vos prévisions est correct.

Le rappel mesure à quel point vous trouvez tous les points positifs [2] .

$$Precision = \frac{TP}{TP + FP} \quad (4.1)$$

$$Recall = \frac{TP}{TP + FN} \quad (4.2)$$

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4.3)$$

Tel que :

TP = True Positive

TN = True Negative

FP = False Positive

FN = False Negative

Les valeurs de rappel augmentent à mesure que nous descendons le classement des prédictions. Cependant, la précision a un motif en zigzag - elle descend avec les faux positifs et remonte avec les vrais positifs, tel que c'est montré dans la Table 4.1.

4.4 Application sur la baie d'Alger

Pour avoir ces images il faut passer par les étapes suivantes :

- A partir du site de Copernicus Open Access Hub [11], comme indiqué sur la Figure 4.3, on a pu télécharger les images. Les recherches se font suivant le type des images et la région géographique.

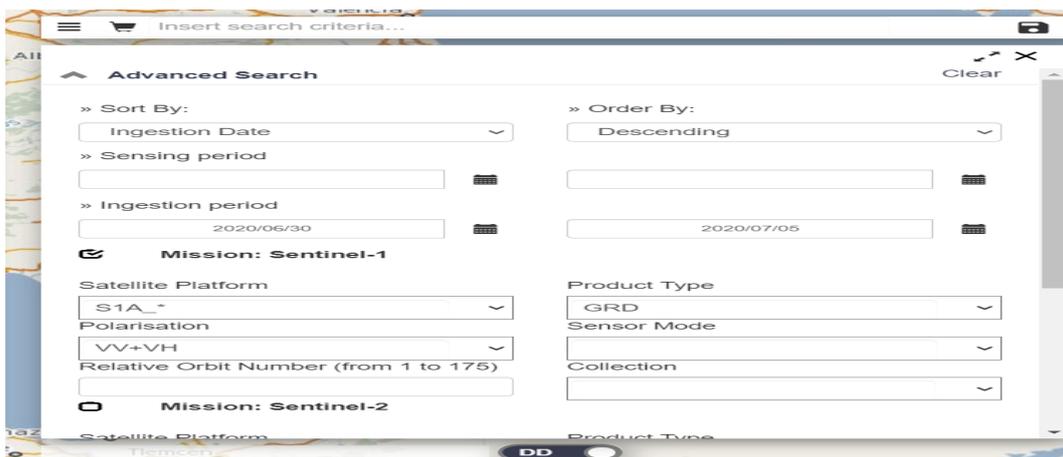


FIGURE 4.3 – Copernicus Open Access Hub

- Pour sélectionner la région à étudier, on clique sur le bouton marqué en rouge (switch to the area), Puis à l'aide du curseur on sélectionne la zone voulue comme indiqué sur la Figure 4.4.

FIGURE 4.4 – *switch to the area mode*

- Afin de trouver les images qui nous intéressent, on doit sélectionner quelques informations tel que la date, la polarisation, le satellite et type de produit, Figure 4.5.

FIGURE 4.5 – *Recherche avancée*

- Choisir l'image à télécharger parmi les résultats. On obtient une grande image qui couvre une grande surface de la côte algérienne (la taille de l'image 1.7GB).

FIGURE 4.6 – *Resultats obtenus*

- Ouvrir l'image avec le logiciel SNAP et zoomer jusqu'à la baie d'Alger, ensuite on découpe cette partie de l'image. Les images sont encore grande avec une dimension de plus de 1000 x 1000 pixels, une capture sur l'étendu de l'image sur Google Earth est présenté dans la Figure 4.7.

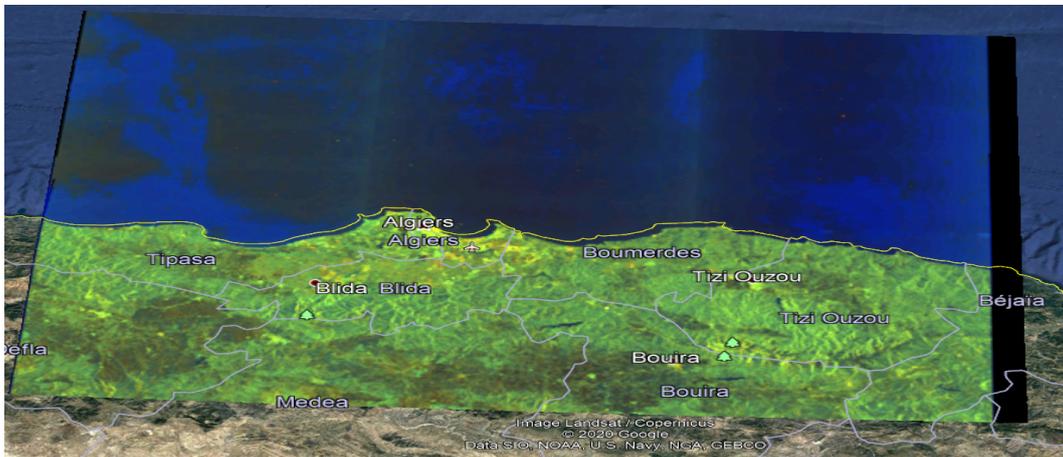


FIGURE 4.7 – Image Sentinel1 de la baie d'Alger dans Google Earth.

- On sélectionne des parties à partir des images de la baie d'Alger avec une dimension de l'ordre de 600 x 600 pixels, pour pouvoir les manipuler et réduire le temps de calcul.



FIGURE 4.8 – Image de la baie d'Alger 600 × 600

4.4.1 Images baie d'Alger traitées Via CA-CFAR

Le facteur de seuil joue un très grand rôle dans le détecteur CA-CFAR, une simple variation peut provoquer des changements radicale au niveau de la détection en s'appuyant sur les figures montrées dans la Figure 4.9, on constate qu'à chaque fois qu'on augmente le α

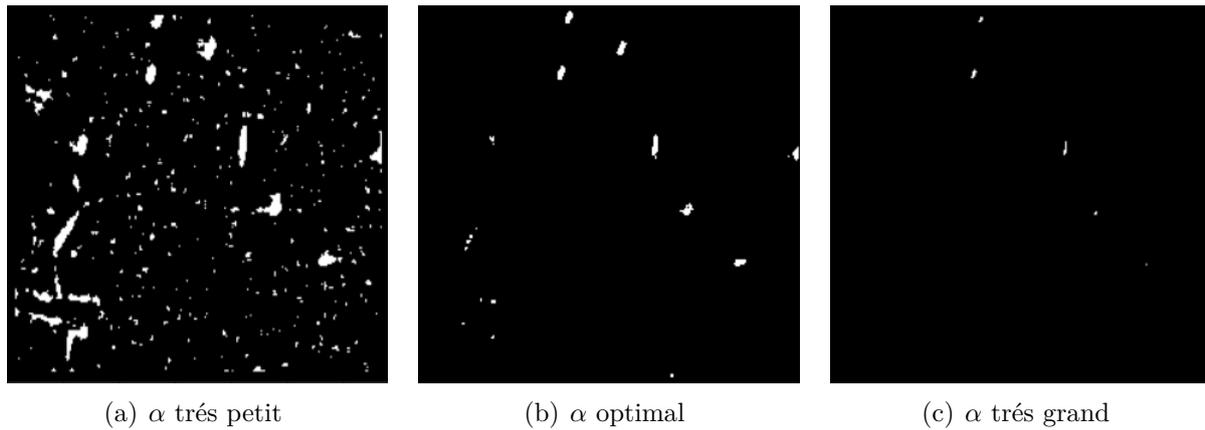


FIGURE 4.9 – Visualisation des images en variant le facteur de seuil α .

le nombre de cellules prises comme cible diminue ce qui implique l'ignorance de quelques cibles confirmées, Figure 4.9 (c) .

Dans le cas contraire le nombre de cellules prises comme cible augmente ce qui nous mène vers le piège de la fausse alarme, Figure 4.9(a) .

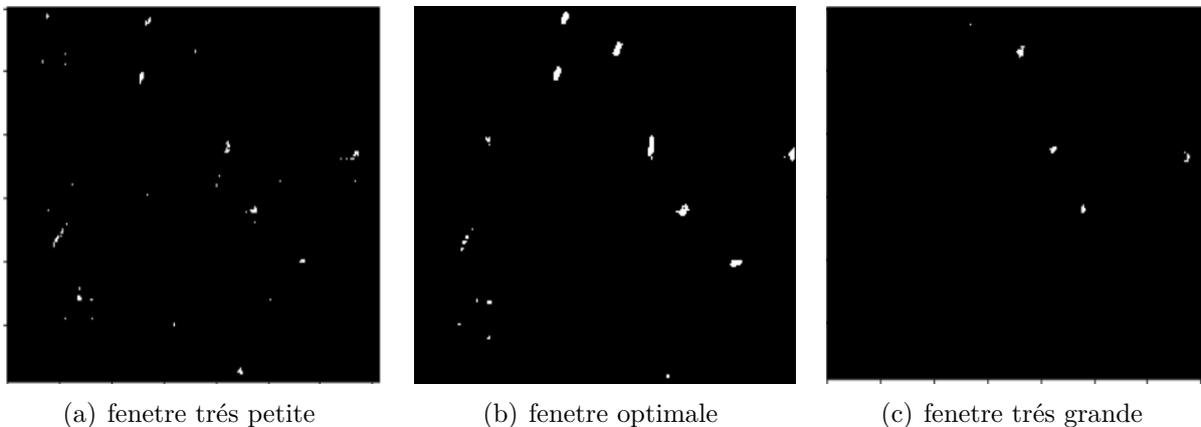


FIGURE 4.10 – Visualisation des images en variant la taille de la fenêtre.

Le choix des tailles des fenêtres est aussi important que la valeur α ce qui implique un choix adapté aux différentes tailles des navires, dans le cas où la taille de la fenêtre est très petite quelque partie de la cible pourrait être considéré comme background comme montre la Figure 4.10 (a) et dans le cas contraire on pourrait tomber dans le piège de la fausse alarme en prenant une partie du background comme cible, Figure 4.10(c).

Pour appliquer ce traitement, il faut passer par les mêmes étapes cités dans le chapitre 2, en tenant en considération les connaissances théorique et pratique on doit choisir un

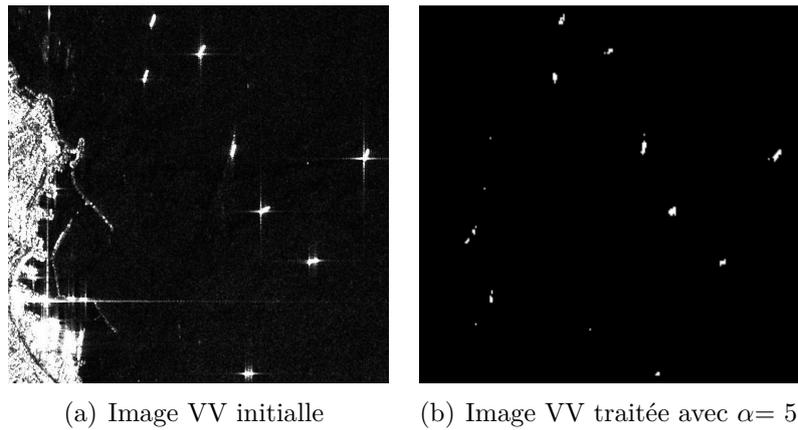


FIGURE 4.11 – Image VV de la baie d'Alger

facteur de seuil α et des tailles de fenêtre optimales afin d'obtenir des résultats adéquats comme montrent les Figures 4.11 et 4.12.

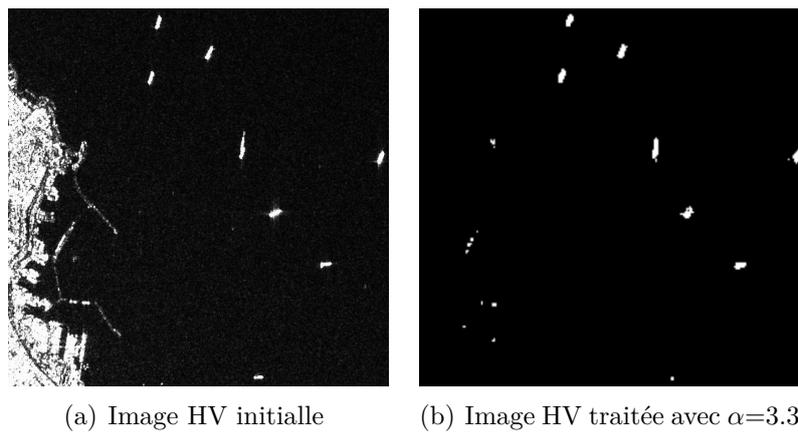
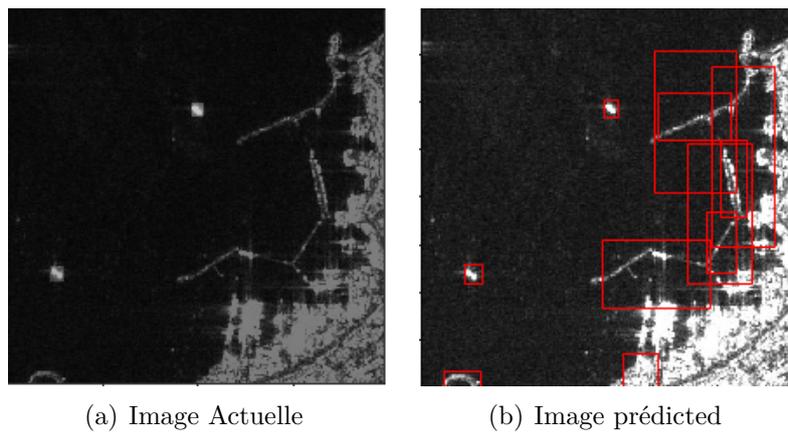


FIGURE 4.12 – Image HV de la baie d'Alger

4.4.2 Images de la baie d'Alger traitées Via Mask RCNN

En visualisant les résultats obtenus, on constate que la détection est médiocre à cause des données d'apprentissage qui ne sont pas adaptés aux images testées, les résultats sur la Figure 4.13.

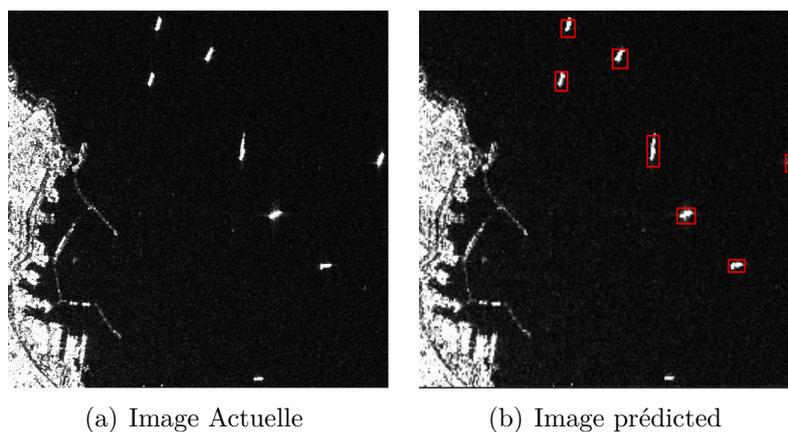
FIGURE 4.13 – *Detection de la baie d'Alger via Mask RCNN*

Le contenu de la base de donnée, qu'on a généré avec les images du satellite Sentinel1, sur la région de la baie d'Alger et ses environ :

Les images acquises pendant trois date différentes (14/06/2020, 21/06/2020 et 31/08/219). Elle contient 60 images (51 images de taille 200x200 et 9 images de taille 400 x 400).

Pour l'apprentissage on a utilisé 51 images et pour le teste les 9 images restantes.

En comparant les deux détections, on remarque que la précision de la détection a été améliorée grâce aux images de l'apprentissage qui sont parfaitement adaptées aux images testées, les résultats sur la Figure 4.14

FIGURE 4.14 – *Detection de la baie d'Alger via Mask RCNN*

4.5 La comparaison entre CA-CFAR et Mask RCNN

En tenant compte des résultats obtenus, on a pu constater que l'un des problèmes majeurs dans la méthode CA-CFAR réside dans la fausse alarme qui consiste à confondre entre les navires et la surface terrestre et pour minimiser l'erreur, il faut appliquer un masque pour éliminer la surface indésirable, par contre dans le cas du modèle RCNN, le problème se pose à la détection des navires adjacents à la surface terrestre. La Figure 4.15 donne des images de deux résultats de détection des navires avec CA-CFAR et Mask RCNN. On remarque que le Mask RCNN est plus robuste au bruit de la surface terrestre et donne de meilleurs résultats par rapport à CA-CFAR.

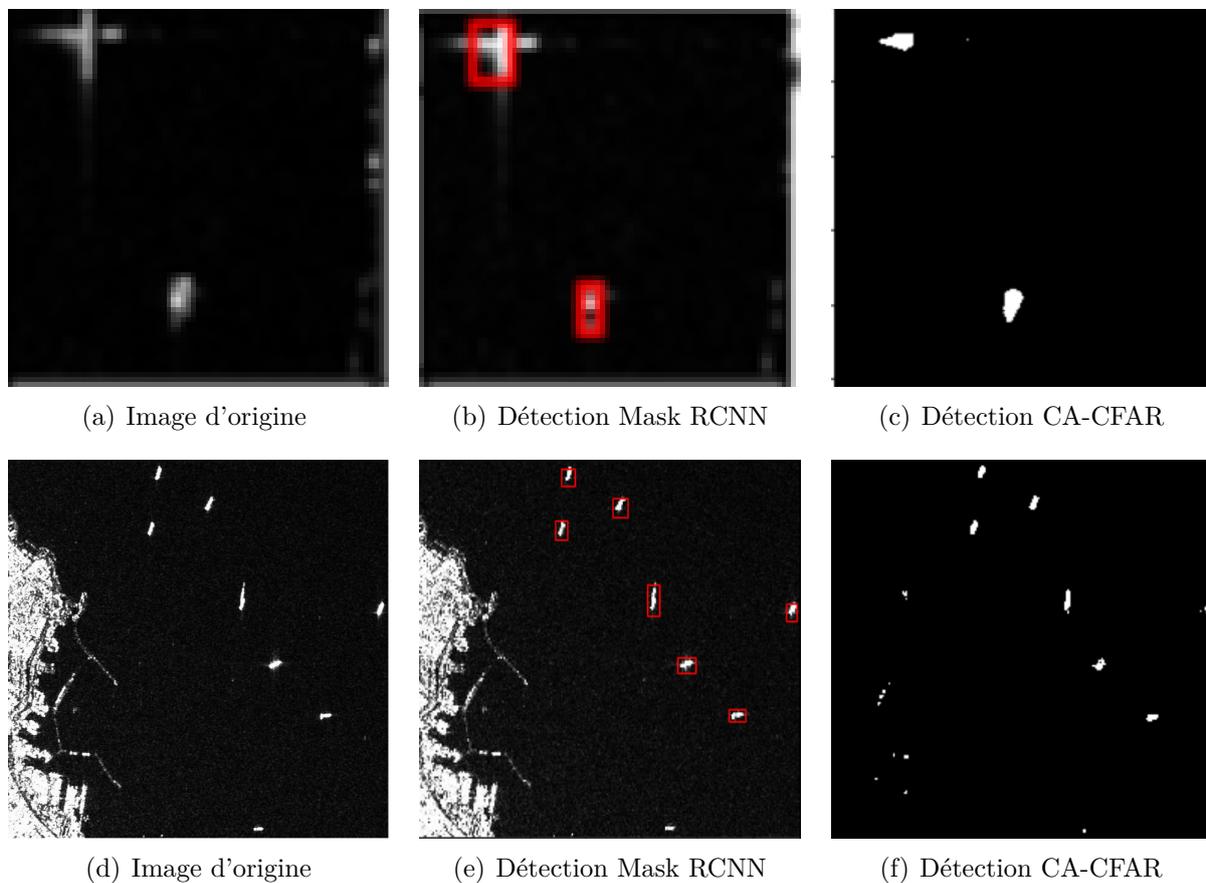


FIGURE 4.15 – La comparaison entre CA-CFAR et Mask RCNN.

Avantages Mask RCNN

- Une adaptation automatique par rapport aux cibles et au milieu.
- Fournit une très grande précision.
- Pouvoir faire la détection et la classification au même temps.

- Une fois l'apprentissage est terminé avec succès, la prédiction se fera rapidement .

Inconvénients Mask RCNN

- Nécessite une très grande quantité de données.
- La base de données doit être conçue par des gens expérimentés dans le domaine de la détection.
- Nécessite des machines à caractéristique et performance très puissante.
- Le changement de milieu ou de la zone de détection implique un changement de données durant l'apprentissage.
- Le temps de traitement durant l'apprentissage est long.

Avantages CA-CFAR

- Ne nécessite pas un apprentissage ni une base de données.

Inconvénients CA-CFAR

- Le changement des paramètres se fait manuellement.
- Problème de précision à cause de la confusion entre le bruit et la cible, ce qui induit une grande difficulté concernant les choix des paramètres optimaux.
- Mauvaise détection quand il s'agit d'un milieu hétérogène.
- La prise de quelque partie des grands navires en background en cas de différence de taille du navire.

Conclusion et Perspective

Conclusion

L'objectif de cette thèse était d'étudier la possibilité d'automatiser, partiellement ou totalement, la détection de bateaux à partir d'images satellitaires.

Ce domaine assez récent reste encore très peu étudié et il n'existe pas encore d'outils de détection automatique sur ce type d'images.

Le but est de fournir des méthodes efficaces, rapides et robustes, tout en étant justifiées théoriquement.

Cette thèse s'insérant dans un contexte de recherche et développement, les données utilisées étaient uniquement des images SAR du satellite Sentinel1.

L'ensemble de notre étude consistait à utiliser deux méthodes de détection (CA-CFAR et Mask RCNN) et les rendre assez robustes afin d'obtenir des résultats satisfaisants.

Durant notre travail on a utilisé une base de données des images SAR sentinel1 de la Chine, et on a construit une base de données des images SAR Sentinel 1 de la région d'Alger et ses environs pour tester nos algorithmes. Elle constitue aujourd'hui pour nous une base de référence pour tester et comparer, facilement et rapidement, les performances d'autres algorithmes.

On a appliqué le détecteur CA-CFAR sur les images de la baie d'Alger, on devait varier différents paramètres dont le facteur de seuil α et la taille de fenêtre afin d'améliorer les résultats, mais tout cela ne permet pas d'éliminer définitivement le bruit.

Par la suite, nous avons traité les mêmes images avec l'algorithme Mask RCNN, ce qui nous a permis de détecter les navires à très haute précision automatiquement.

D'après les résultats obtenus, le Mask RCNN a fait preuve de supériorité en ce qui concerne la précision et la robustesse face au CA-CFAR.

Perspective

On encourage les étudiants qui veulent poursuivre notre parcours à s'investir dans le domaine de détection en temps réel ainsi que l'utilisation des modèles de Deep Learning hybride afin d'améliorer les caractéristiques et les performances du système de détection.

Bibliographie

- [1] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN : Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
- [2] J. Hui, “mAP (mean Average Precision) for Object Detection,” Apr. 2019. [Online]. Available : https://medium.com/@jonathan_hui/map-mean-average-precision-for-object-detection-45c121a31173
- [3] F. Pierre, “Outils statistiques et géométriques pour la classification des images sar polarimétriques hautement texturées.” *PhD thesis, RENNES 1*, Dec. 2013.
- [4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once : Unified, Real-Time Object Detection,” 2016, pp. 779–788. [Online]. Available : https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html
- [5] J. Redmon and A. Farhadi, “YOLOv3 : An Incremental Improvement,” *arXiv :1804.02767 [cs]*, Apr. 2018, arXiv : 1804.02767. [Online]. Available : <http://arxiv.org/abs/1804.02767>
- [6] Y. Wang, C. Wang, H. Zhang, Y. Dong, and S. Wei, “A SAR dataset of ship detection for deep learning under complex backgrounds,” vol. 11, no. 7, p. 765, publisher : Multidisciplinary Digital Publishing Institute.
- [7] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” *arXiv :1703.06870 [cs]*, Jan. 2018, arXiv : 1703.06870. [Online]. Available : <http://arxiv.org/abs/1703.06870>
- [8] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár, “Microsoft COCO : Common Objects in Context,” *arXiv :1405.0312 [cs]*, Feb. 2015, arXiv : 1405.0312. [Online]. Available : <http://arxiv.org/abs/1405.0312>
- [9] R. n. Canada. Détection de navires. Last Modified : 2014-11-24 Library Catalog : www.rncan.gc.ca Publisher : Ressources naturelles Canada. [Online]. Available : <https://www.rncan.gc.ca/sciences-terre/geomatique/imagerie-satellitaire-photos-aeriennes/imagerie-satellitaire-produits/ressources-educatives/9520>
- [10] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, “SSD : Single Shot MultiBox Detector,” in *Computer Vision – ECCV 2016*, ser. Lecture Notes in Computer Science, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham : Springer International Publishing, 2016, pp. 21–37.

-
- [11] ESA. Copernicus open access hub. [Online]. Available : <https://scihub.copernicus.eu/dhus/#/home>