

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

UNIVERSITE BADJI MOKHTAR - ANNABA
BADJI MOKHTAR – ANNABA UNIVERSITY



جامعة باجي مختار – عنابة

Faculté : TECHNOLOGIE
Département : ELECTRONIQUE
Domaine : SCIENCES ET
TECHNIQUES
Filière : télécommunication
Spécialité : systèmes de telecommunication

Mémoire

Présenté en vue de l'obtention du Diplôme de Master

Thème:

Segmentation sémantique multi-échelles pour les images de drones

Présenté par : MAHMOUDI SAMEH

Encadrant : BOUKARI Karima

UBM Annaba

Jury de Soutenance :

ZADAM Mohammed	M.C.B	ANNABA	Président
BOUKARI Karima	M.C.A	ANNABA	Encadrant
/	/	/	/
AMARA Fethi	M.C.B	ANNABA	Examineur
/	/	/	/

Année Universitaire : 2021/2022

Remerciements :

Tout d'abord, nous remercions Dieu le tout-puissant de nous avoir donné le courage et la volonté de faire ce travail humble.

Avec tout le respect, nous tenons à remercier vivement tout le corps enseignant de **l'université BADJI MOKHTAR –ANNABA-** qui nous a suivis tout au long de nos études.

Le travail présenté dans ce mémoire a été effectué sous la direction de **MME BOUKARI** à qui je tiens à adresser mes plus vifs remerciements, pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à alimenter ma réflexion et ses encouragements lors de la réalisation de cette mémoire.

Mes sincères remerciements au président et aux membres de jury pour m'avoir fait l'honneur en acceptant d'examiner et de juger ce travail.

Je tiens aussi à remercier tous ceux qui ont participé de près ou de loin dans la réalisation de ce travail ne serait-ce qu'en me prodiguant un conseil ou autre.

Dédicace

Du profond de mon cœur, je dédie ce travail à tous ceux qui me sont cher,

A MA CHERE MERE

Aucune dédicace ne saurait exprimer mon respect, mon amour éternel et ma considération pour les sacrifices que vous avez consenti pour mon instruction et mon bien être.

Je vous remercie pour tout le soutien et l'amour que vous me portez depuis mon enfance et j'espère que votre bénédiction m'accompagne toujours.

Que ce humble travail soit l'exaucement de vos vœux tant formulés, le fruit de vos innombrables sacrifices.

Puisse dieu, le Très Haut, vous accorder santé, bonheur et long vie

A LA MEMOIRE DE MON PERE

Ce travail est dédié à mon père, décédé trop tôt qui m'a toujours poussé dans mes études.

J'espère que, du monde qui est sien maintenant, il apprécié cet humble geste comme preuve

De reconnaissance de la part d'une fille qui a toujours prié pour le salut de son âme.

Puisse Dieu, le tout puissant, l'avoir en sa sainte miséricorde

A tous mes professeurs Leur générosité et leur soutien m'oblige de leurs témoigner mon profond respect et ma loyale considération.

A Mes sœurs, mes frères, mes nièces et mes neveux

A Toutes mes amies, ' et surtout **MOUNIA et SABAH** en leurs souhaitant le succès dans leur vie.

A Toute ma famille, proche soit-elle ou lointaine.

ملخص :

تستفيد الصور الجوية على ارتفاعات منخفضة ومتوسطة من مجموعة واسعة من تطبيقات الطائرات بدون طيار مثل مراقبة حركة المرور والمراقبة وتخطيط المسار والتفتيش تُستخدم صور الأقمار الصناعية للزحف على الويب ، خاصةً لاكتساب وتحديث نظام تحديد المواقع العالمي GPS ومع ذلك ، فإن بعض التطبيقات ، مثل مراقبة الكوارث ، ومراقبة حركة المرور ، وتخطيط الإخلاء في حالات الطوارئ ، تتطلب معالجة في الوقت الفعلي للبيانات المكتشفة ، وصور الأقمار الصناعية غير مناسبة مثل البيانات المجمعة في الموقع بواسطة الطائرات بدون طيار. تستخدم الطائرات بدون طيار على نطاق واسع كمنصات فعالة من حيث التكلفة لجمع البيانات الميدانية بكفاءة لتطوير الأجهزة. تصنيف الصور هي أحد المجالات الشعبية للباحثين في رؤية الكمبيوتر. تسلط هذه الدراسة الضوء على استخدام البيكسل الفائقة التكراري الخطي البسيط (SLIC) جنباً إلى جنب مع مصنف SVM لتجزئة الصور بالإضافة إلى تقنيات التعلم العميق لتوصيف وحدات البيكسل الفائقة. الصور. تصبح التجزئة بعد ذلك مشكلة تصنيف فائقة البيكسل

الكلمات الرئيسية: خوارزمية SVM ، صور SLIC فائقة البيكسل ، ميزة CNN

Résumé :

L'imagerie aérienne à basse et moyenne altitude bénéficie d'un large éventail d'applications de drones telles que la surveillance et la surveillance du trafic, la planification et l'inspection d'itinéraires.

L'imagerie satellitaire est utilisée pour l'exploration du Web, en particulier pour l'acquisition et la mise à jour du GPS.

Cependant, certaines applications, telles que la surveillance des catastrophes, la surveillance du trafic et la planification des évacuations d'urgence, nécessitent un traitement en temps réel des données détectées, et l'imagerie satellitaire n'est pas adaptée, comme les données collectées sur site par des drones.

Les drones sont largement utilisés comme plates-formes rentables pour collecter efficacement des données terrain pour le développement de matériel.

La classification d'images est l'un des domaines populaires pour les chercheurs en vision par ordinateur. Cette étude met en évidence l'utilisation d'un superpixel de regroupement itératif linéaire simple (SLIC) en combinaison avec un classifieur SVM pour la segmentation d'image ainsi que les techniques d'apprentissage en profondeur pour la caractérisation des superpixels. la superpixelisation est une étape de preprocessing ou pré-segmentation pour réduire le volume des pixels dans les images aériennes de grande dimension. La segmentation devient alors un problème de classification des superpixels.

Mots clés : L'algorithme SVM, les images superpixelisées SLIC, CNN feature

Abstract:

Low and medium altitude aerial imagery benefits from a wide range of drone applications such as traffic monitoring and surveillance, route planning and inspection.

Satellite imagery is used for web crawling, especially for GPS acquisition and updating.

However, some applications, such as disaster monitoring, traffic monitoring, and emergency evacuation planning, require real-time processing of detected data, and satellite imagery is not suitable, such as collected data. on site by drones.

Drones are widely used as cost-effective platforms to efficiently collect field data for hardware development

Image classification is one of the popular areas for computer vision researchers. This study highlights the use of a simple linear iterative clustering (SLIC) superpixel in combination with an SVM classifier for image segmentation as well as deep learning techniques for superpixel characterization. Is a preprocessing or pré-segmentation step to reduce the volume of pixels in high-dimensional aerial images Segmentation Then becomes a super pixel classification problem.

Keywords: SVM algorithm, superpixelated SLIC images, CNN feature.

Liste des figures

Chapitre I : Segmentation par classification des images aériennes

Figure I.1 Exemple de segmentation	p 3, p4
Figure I.2 Exemple de segmentation sémantique	p5
Figure I.3 Cas SVM non linéairement	p9

Chapitre II: deep learning

Figure. II.1 Différence entre CNN et Machine Learning	p12
Figure II.2 - Différents types de convolution	p14
Figure. II.3 - Cartes des caractéristiques.	p16
Figure II.4 - Architecture alexnet	p 17

Chapitre III: Expérimentation et résultats

Figure III.1 L'organigramme des étapes du programme.	p21
Figure III.2 résultats de classification.	p22
Figure III.3 - Architecture alexnet	p23
Figure III.4 -résultat de la segmentation avec L=200	p23
Figure III.5 -résultat de la segmentation avec L=300.	p24
Figure III.6 -résultat de la segmentation avec L=400.	p24
Figure III.7 -résultat de la segmentation avec L=500.	p25
Figure III.8 -résultat de la segmentation de l'image L=600.	p25

Liste des tableaux :

TABLEAU III.1 les couches d'Alex-Net. p20

TABLEAU III.2 Les scores pour l'image test 010 p26

Liste des Formues :

Formues chapitre I :

Formule (I.1) : $FSLIC(si, pj) = \sqrt{\mathcal{F}c(si, pj)^2 + \mathcal{F}p(si, pj)^2 \left(\frac{\omega cNI}{\sqrt{\omega s}}\right)^2}$ p6

Formule (I 2): $\mathcal{F}c(si , pj) = \sqrt{(I i , Ij)^2 + (ai - aj)^2 + (bi - bj)^2}$ p6

Formule (I.3) : $\mathcal{F}p(si , pj) = \sqrt{(xi , xj)^2 + (yi - yj)^2}$ P6

n

Formues chapitre II :

Formule (II .1) : $S(x) = (f * g)(x) \sum_{a=-\infty}^{+\infty} f(t)g(x-a)$p12

Formule (II. 2) : $S(I, j) = (K * I)(I, j) \sum_m \sum_n I(i-m, j-n)K(m, n)$ p13

Formule (II.3) : $S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i+m, j+n)K(m, n)$ p13

Formule (II .4) : $O = \frac{h-k+2p}{s} + 1$ p14

Formule (II .5) : $s(i, j) = (K * I')(i, j) = \sum_{m=0}^{k-1} \sum_{n=0}^{k-1} I'(s.i+m, s.jn)K(m, n)$ p14

Formule (II .6) : $O = \frac{h-k-(k-1)(d-1)+2p}{s} + 1$ P15

Formule (II .7) : $O = (h-1) s+k-2p$ p15

Formule (II .8) : $S = \frac{1}{\xi} (\xi \in \mathbb{N}^*)$ p15

Formules de chapitre III :

Formule (III.1) : Accuracy = (TP+TN) / (FN+FP+TP+TN) p21

Formule (III.2) : Précision = TP / (TP+FP) ; p21

Formule (III.3) : Fmeasure = 2*TP / (2*TP+FP+FN);p21

Formule (III.4) : Spécificity = TN / (TN+FP)...p21

Formule (III.5) : Sensitivity (la sensibilité)= TP / (TP+FN) ;p21

Formule (III .6) : MCC= (TP*TN-FP*FN)/sqrt ((TP+FP)*(TP+FN)*(TN+FP)*(TN+FN))p21

Formule (III. 7) : Jacquard Similarity Coefficient =TP / (TP + FP ++ FN)p21

Liste des acronymes :

CHAPITRE I :

(GPS): Unité de traitement graphique (Global Positioning System).

(SLIC): Simple linéaire iterative clustering

(SVM) : Les machines à vecteurs de support

(CNN) : Convolutional Neural Networks

(K): Le nombre des superpixels

(NI) : Le nombre de pixels de l'image

(NS) : ensembles de pixels

(ω NS): un paramètre indiquant le nombre de superpixels

(ωc) : Le paramètre de compacité de l'algorithme SLIC

(b) : la marge maximale.

CHAPITRE II :

(ANN) : Artificiel Neural Network

(IA) : L'intelligence artificielle.

(k) : la taille du noyau de convolution

(s) : le pas

(p) : le padding.

CHAPITRE III :

(TP): True Positive.

(TN): True Negative.

(FN): False Negative.

(FP): False Positive.

(MCC): Matthews's correlation coefficient.

(UAV): Unmanned Aerial Vehicle

Sommaire :

Remerciements	
Dédicace .	
Résumé .	
Liste des figures .	
Liste des tableaux .	
Liste des Formues	
Liste des acronymes .	

Table des matières :

Introduction générale	p1
Chapitre I : Segmentation par classification des images aériennes	p2
I.1 Introduction.....	p3
I.2 définition de la segmentation.....	p3
I.3 Méthode de classification	p4
I.3.1 Méthodes de calcification supervisées.....	.p4
I.3.2 Méthodes de calcification non supervisées.....	.p4
I .4 Segmentation sémantique :.....	p4
I .5 Superpexaliation	p5
I.5.1 Introduction	p5
I .5.2 Définition de la superpixelisation	p5
I.5.3 L'algorithme SLIC	p 5
I.5.3.1Principe général	p5
I.5.3.2Méthode d'Achanta et al. (SLIC)	p6
I .5.3.3Classification par SVM.....	p7
Chapitre II: deep learning	p11
II.1 Introduction	P12
II.2 Définition	P12
II.3Application du deep Learning au contexte routier	P12
II .4 Réseaux de neurones convolutifs	p12
II .4 .1 Les type de convolution	P14
II .4 .1.1 Convolution classique ou convolution à pas	p14
II .4 .1.2 Convolution à trous ou convolution dilatée.....	.p14
II .4 .1.3 Convolution transposée ou dé convolution.....	p15
II .4 .1.3.1 Les interactions convolutives.....	.p15
II .4 .1.3.2 Le partage des paramètres	p15
II .4 .1.3.3 Les représentations équivariantesp16
II .4 .2 Couche d'échantillonnage (Pooling).....	p17
II .4 .2 Réseau Alex Net.....	p17
Chapitre III: Expérimentation et résultats	p18
III.1 Introduction	p19
III.2Plan de travail	p19
III.3 Matériel utilisé	p20
III.4 Logiciel utilisé	p20

III .5 Segmentation par superpixelisation	p21
III .6 Classification supervisée par SVM	p22
III.7 Les résultats de segmentation	p23
III.8 Résultats expérimentaux	p26
III.9 Conclusion.....	p26
III.10 Conclusion Générale	p27
Bibliographie.....	p28

Introduction générale:

L'un des objectifs à long terme de la vision par ordinateur est la compréhension automatisée de la scène à partir d'une variété d'images. L'exposition des spécifications d'image est utile pour des applications telles que l'édition d'images, la recherche d'images et la perception de l'environnement pour les véhicules autonomes. La Détection des objets comme des routes dans L'imagerie aérienne à basse et moyenne altitude a bénéficié à toute une gamme de types d'applications d'UAV telles que la surveillance des catastrophes, la surveillance du trafic et les plans d'évacuation d'urgence à cette fin. En raison de la variabilité des différents facteurs tels que la couleur, la forme, l'éclairage et les ombres ou des obstacles sur la surface de la route, la détection de route est un problème difficile. Les techniques pour résoudre ce problème sont principalement basées sur l'apprentissage en profondeur et les réseaux de neurones convolutifs (CNN). Ces méthodes permettent une meilleure compréhension visuelle en appliquant un processus de segmentation sémantique dans lequel chaque pixel est affecté à une catégorie d'objet. Le résultat de la segmentation fournit des informations significatives pour prendre en charge les tâches de compréhension de scène de haut niveau.

Les réseaux de neurones convolutifs (CNN) ont apporté des améliorations considérables pour les tâches de segmentation d'images dans le domaine de la vision par ordinateur. Malgré leur succès, un défi inhérent est le compromis entre la précision et le coût de calcul. Les efforts de calcul élevés pour les grands réseaux fonctionnant sur la grille de pixels de l'image les rendent inéligibles pour de nombreuses applications en temps réel telles que divers systèmes avancés d'assistance à la conduite (ADAS). Dans ce travail, nous proposons une nouvelle approche CNN, basée sur la combinaison de super-pixels et de caractéristiques extraits de CNN appliqués pour la segmentation des routes. L'idée centrale est de réduire le temps de calcul et la complexité en segmentant l'image en régions homogènes (superpixels) et en alimentant les descripteurs d'images extraites de ces régions dans un CNN plutôt que de travailler directement sur la grille de pixels. Pour activer les opérations de convolution sur les superpixels irréguliers, nous introduisons un schéma de projection en réseau comme partie de la méthode de création de superpixels, qui compose les relations de voisinage et force la topologie à rester fixe pendant le processus de segmentation. La réduction de l'entrée dans le domaine des superpixels permet aux CNN structure pour rester petite et efficace à calculer tout en gardant l'avantage des couches convolutives.

Chapitre 01

Segmentation par classification des images aériennes

Chapitre I : La segmentation par classification des images aérienne

I.1 introduction :

La segmentation d'images est un processus visant à décomposer une image en un ensemble de régions ou classes ou sous-ensembles homogènes au sens d'un ou plusieurs critères. En imagerie aérienne, la segmentation est très importante, que ce soit pour l'extraction de paramètres ou de mesures sur les images, ainsi que pour la représentation et la visualisation. Dans notre application la segmentation sera exploitée pour détecter les zones routières.

La segmentation fait appel à plusieurs branches des mathématiques et de l'informatique. Comprendre les enjeux réels et déterminer les performances, les avantages et les inconvénients de chaque approche est donc une tâche difficile. Le choix d'une méthode sera toutefois guidé par des critères comme le type d'image, le bruit, les artéfacts d'acquisition, l'information a priori disponible, le temps de calcul, le niveau d'interaction acceptable ou encore la possibilité de corriger efficacement le résultat.

La segmentation est l'un des principaux thèmes du domaine de l'analyse d'images

I.2 Définition de la segmentation :

La segmentation est une étape clé pour un grand nombre de méthodes d'analyse d'images. Son but consiste à produire une partition des pixels de l'image en composantes connexes, selon des critères prédéfinis (homogénéité en termes de couleurs, de texture, etc.). Ces composantes connexes sont nommées régions et correspondent à des objets ou à des parties d'objets, formant un pavage de l'image en primitives visuelles de haut niveau. La figure -1- contient un exemple de segmentation.

Les pixels de l'image originale (figure A) sont groupés en une dizaine de régions (figure B)

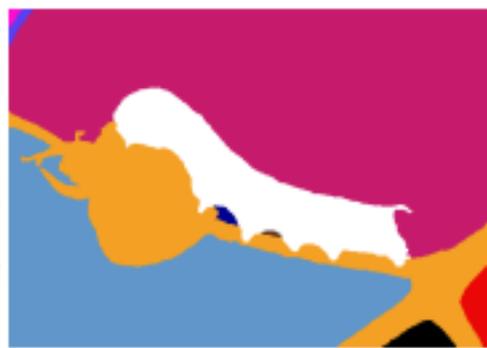
(La figure C) représente les contours de ces régions, où apparaissent en blanc les pixels ayant un de leurs voisins qui appartient à une autre région.

Trois hypothèses guident les travaux menés en segmentation :

- l'homogénéité à l'intérieur de chaque région en termes de couleur ou de texture doit être maximisée
- les similitudes entre les différentes régions doivent être minimisées
- les pixels voisins appartenant à des régions différentes doivent être aussi dissemblables que possible. [1]



(A) Image originale



(B) Segmentation d' Image : une couleur par région



(C) Contours des régions de cette segmentation.

Figure – I.1 - Exemple de segmentation [1]

I.3 Méthode de classification :

De nombreuses méthodes classiques ont été consacrées, elles peuvent être séparées en deux grandes catégories : les méthodes de classification supervisée et les méthodes de classification non supervisée.

I.3.1 Méthodes de Classification supervisée :

Le but de la classification supervisée est avant tout de définir des règles permettant de classer des objets selon les variables qualitatives ou quantitatives qui les caractérisent.

Au départ, il existe un échantillon dit d'apprentissage dont la classification est connue, et cet échantillon est utilisé pour apprendre des règles de classification.

Il est nécessaire d'étudier la fiabilité de ces règles afin de les comparer et de les appliquer et d'évaluer les situations de sous- ou de sur apprentissage. Habituellement, un deuxième échantillon indépendant est utilisé, appelé échantillon de validation ou d'essai.

I.3.2 Méthodes de Classification non supervisée :

Faites-le dans le sens inverse, c'est-à-dire qu'ils ne nécessitent aucun apprentissage ni aucune tâche d'étiquetage manuel préalable. Il comprend des nuages de points qui représentent n'importe quel espace dans un ensemble de groupes appelés clusters. Cela implique généralement le domaine de l'analyse de données, comme l'ACP. . Un «Cluster» est une collection d'objets qui sont similaires entre eux et qui sont dissemblables par rapport aux objets appartenant à d'autres groupes

I.4 Segmentation sémantique :

Alors que la segmentation divise simplement l'image en régions,

La segmentation sémantique divise les pixels en ensembles correspondant à

Classe d'objet. Comme un même objet peut correspondre à plusieurs zones ou plusieurs

Il y a des objets de la même catégorie dans l'image, la collection n'est pas forcément composants connexes. Si le problème de segmentation peut être considéré comme une recherche Frontières des objets dans les photos, frontières de la segmentation sémantique réside dans le repérage et l'identification des objets qui le composent, C'est-à-dire comme une résolution Problème de classification, une classe par classe d'objet (La figure b) contient une segmentation sémantique associée au graphe a. Les pixels sont divisés en trois groupes :

Le premier correspond à une table en bois (jaune), le second à des chenilles (rose) et Troisième branche de tomate (bleu).



A : image originale

(b) Segmentation sémantique de cette image
Trois classes sont identifiées : le bois (en jaune),
la chenille (en rose) et la tomate (en bleu).

Figure - I.2 - Exemple de segmentation sémantique [1]

I.5 Superpexaliation :

I.5.1 Introduction :

De nombreuses applications de traitement d'image peuvent bénéficier des étapes suivantes :

La pré-segmentation regroupe a priori les pixels d'une image. Un exemple d'analyse d'image implique généralement l'identification et la localisation de différents objets présents dans une image. Cette identification peut être réalisée par pré-segmentation

objets, puis les classer. Par conséquent, la segmentation est une étape essentielle qui peut avoir un impact important sur la précision et le temps de calcul des méthodes de classification. Pour réduire le temps de calcul, une méthode de décomposition de superpixel est introduite. Ces méthodes regroupent les pixels d'une image en régions homogènes tout en Essaie de respecter les contours des objets dans l'image, réduisant ainsi le nombre L'élément à traiter, d'où la charge de calcul.

I.5.2 Définition de la superpixelisation :

La superpixelisation est une sur-segmentation, généralement basée sur la couleur ou niveau de gris des pixels de l'image .Un superpixel est alors défini comme un groupe de pixels ayant des caractéristiques similaires. Par conséquent, la superpixelisation est une simplification d'une image. Comparé aux méthodes basées sur les pixels, il réduit la quantité de données à traiter, réduisant ainsi le coût de calcul des différentes opérations et améliorant les performances. Les méthodes de superpixelisation peuvent être divisées en deux catégories, produisant des décompositions régulières ou irrégulières.

I.5.3 L'algorithme SLIC :

I.5.3.1 Principe général :

L'algorithme SLIC et choix des paramètres Proposé par Achanta et al. SLIC est une version modifiée de l'algorithme des k-moyennes.

L'algorithme SLIC permet de segmenter une image en K régions appelées superpixels en fonction du contenu de l'image. La segmentation en superpixels est très utile pour faire de la détection d'objets, pour découper des images et pour alléger la quantité de données à traiter

Les superpixels initiaux sont générés en regroupant les pixels dans une grille régulière. Choisir La taille de la case est un paramètre qu'il convient de régler, soit en précisant le nombre souhaité Superpixels (qui divisent ensuite les pixels de l'image en autant de cases que nécessaire), ou Spécifiez la surface moyenne requise pour le superpixel (la hauteur et la largeur de la boîte sont déterminé à se rapprocher le plus possible de cette zone) L'algorithme SLIC répète alors dix fois les deux étapes suivantes :

1. Calculez la couleur moyenne et la position de chaque superpixel ;
2. Chaque pixel est attaché à un superpixel, maximisant la fonction de similarité [2]

I.5.3.2 Méthode d'Achanta et al. (SLIC) :

La méthode SLIC proposée par Achanta et al. [1], correspondant à la version modifiée

Algorithme des k-moyennes. Il se compose de trois étapes :

Initialisation, où la première sursegmentation est donnée ;

Regroupez les pixels en ensembles de sorte que chaque pixel soit ajouté à L'ensemble des caractéristiques visuelles les plus proches de lui-même ;

Le post-traitement garantit que la collection obtenue à l'étape précédente forme un divisé en éléments connexes.

A l'étape d'initialisation, les pixels sont regroupés en superpixels correspondant au découpage régulier de l'image sous forme de grille. D'autres configurations sont également possibles considérons, par exemple, les superpixels hexagonaux

La deuxième étape est basée sur un processus itératif, répétant l'action dix fois Ce qui suit :

1- Calculer la couleur et la position moyenne de ces pixels du groupe NS correspondant au superpixel initial

2 -Chaque pixel, décrit par sa couleur et sa position dans l'image, est affecté à un superpixel il est le plus proche au sens de la mesure de similarité

3 -mises à jour de fonctionnalités superpixel

L'une des clés du succès de SLIC est que chaque pixel est comparé de manière unique à l'ensemble le plus proche, permettant à la méthode de produire une sursegmentation Complexité presque linéaire par rapport au nombre de pixels dans l'image.

Soit un superpixel, le pixel p_j et l_i , a_i et b_i la couleur moyenne représentée dans l'espace CIELab s_i , x_i et y_i coordonnées barycentrique de s_i , l_j , a_j et b_j la couleur p_j , x_j et y_j leurs coordonnées

La similarité entre s_i et p_j est évalué par la fonction suivante :

$$FSLIC(s_i, p_j) = \sqrt{\mathcal{F}c(s_i, p_j)^2 + \mathcal{F}p(s_i, p_j)^2 \left(\frac{\omega_{cNI}}{\sqrt{\omega_s}}\right)^2} \quad (I.1)$$

Avec :

$$\mathcal{F}c(s_i, p_j) = \sqrt{(l_i, l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \quad (I.2)$$

La distance euclidienne entre la couleur moyenne du super pixel et celle du pixel.

$$\mathcal{F}p(s_i, p_j) = \sqrt{(x_i, x_j)^2 + (y_i - y_j)^2} \quad (I.3)$$

La distance euclidienne entre le barycentre du super-pixel et la position du pixel.

NI Le nombre de pixels de l'image

ωc Un paramètre de compacité, pondérant l'influence de la position du pixel par rapport à sa couleur ;

Ns un paramètre indiquant le nombre de super-pixels souhaités.

Soit \hat{S} l'ensemble des super-pixels ainsi obtenus. Les pixels étant regroupés aussi en fonction de leur couleur, il n'est pas garanti que S_0 forme une partition de l'image en composantes connexes. Afin d'assurer le respect de la propriété 1 (validité), les composantes connexes de \hat{S} sont extraites. Celles dont le nombre de pixels est en dessous d'une taille minimale sont fusionnées avec une composante connexe voisine, donnant S, la sur-segmentation finale.

I.5. 3.3 Classification par SVM :

Définition :

Les machines à vecteurs de support sont un ensemble de méthodes d'apprentissage supervisé pour la classification, la régression et la détection des valeurs aberrantes. Toutes ces tâches sont courantes dans l'apprentissage automatique.

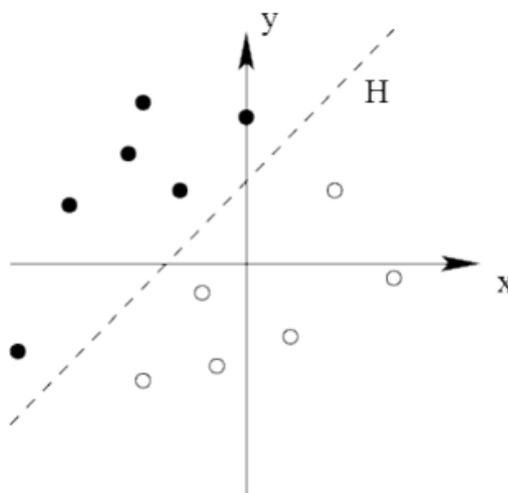
Les SVM sont différents des autres algorithmes de classification en raison de la façon dont ils choisissent la limite de décision qui maximise la distance par rapport aux points de données les plus proches de toutes les classes. La frontière de décision créée par les SVM est appelée classificateur de marge maximale ou hyperplan de marge maximale.

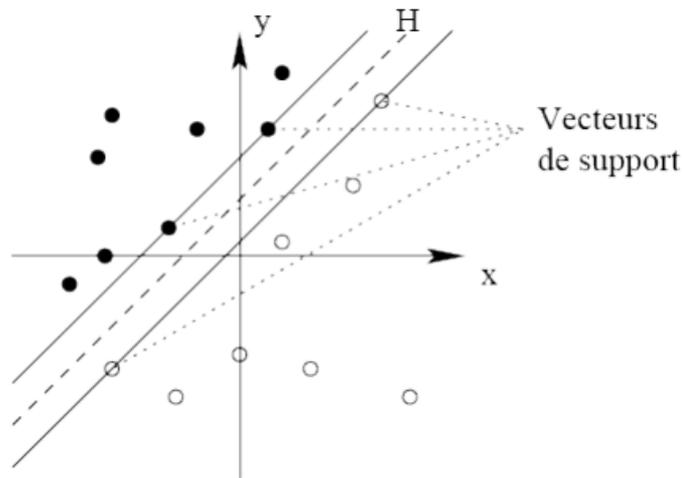
Le classificateur SVM linéaire simple fonctionne en créant une ligne droite entre deux classes. Cela signifie que tous les points de données d'un côté de la ligne représenteront une catégorie, tandis que les points de données de l'autre côté de la ligne seront placés dans une catégorie différente. Cela signifie qu'il peut y avoir un nombre illimité de lignes parmi lesquelles choisir. Ce qui rend l'algorithme SVM linéaire meilleur que certains autres algorithmes comme k voisins les plus proches, c'est qu'il choisit la meilleure ligne pour classer les points de données, il choisit la ligne qui sépare les données et est la plus éloignée du point de données le plus proche.[4]

SVM principe de fonctionnement général :

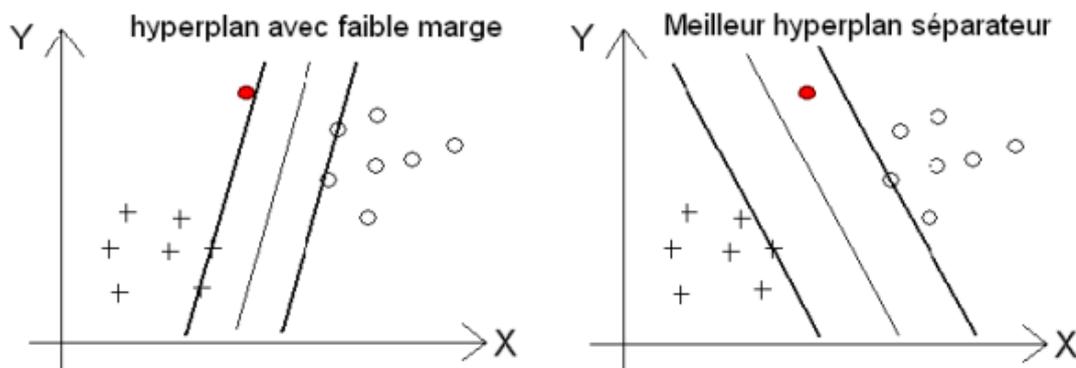
Notions de base : Hyperplan, marge et support vecteur :

Le but du SVM est de trouver un classificateur pour deux classes d'exemples donnés Cela séparera les données et maximisera la distance entre ces deux classes. En utilisant SVM, cela Le classificateur est un classificateur linéaire appelé hyperplan. Dans l'image ci-dessous, nous avons identifié un hyperplan qui sépare deux ensembles de points.

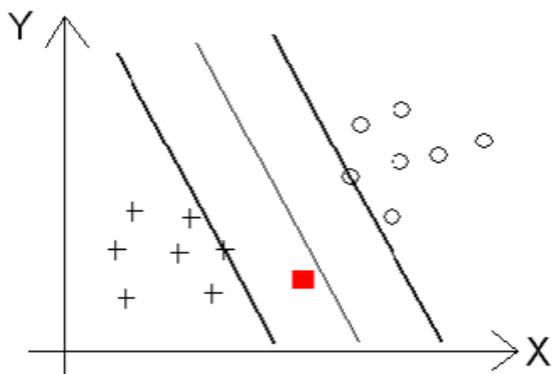




Intuitivement, le fait d'avoir une marge plus large procure plus de sécurité lorsque l'on classe un nouvel exemple. De plus, si l'on trouve le classificateur qui se comporte le mieux vis-à-vis des données d'apprentissage, il est clair qu'il sera aussi celui qui permettra au mieux de classer les nouveaux exemples. Dans le schéma qui suit, la partie droite nous montre qu'avec un hyperplan optimal, un nouvel exemple reste bien classé alors qu'il tombe dans la marge. On constate sur la partie gauche qu'avec une plus petite marge, l'exemple se voit mal classé.



La classification d'un nouvel exemple inconnu est donnée par sa position par rapport à l'hyperplan optimal. Dans le schéma suivant, le nouvel élément sera classé dans la catégorie des « + ».



On note les cas séparables et les cas séparables non linéaire.

Les premiers sont les plus simples en SVM car ils permettent de trouver facilement le classifieur linéaire. Dans la plupart des problèmes réels où il n'y a pas de possibilité de séparabilité linéaire entre les données, le classifieur à marge maximale ne peut pas être utilisé car c'est seulement si les classes des données d'apprentissage sont linéaires

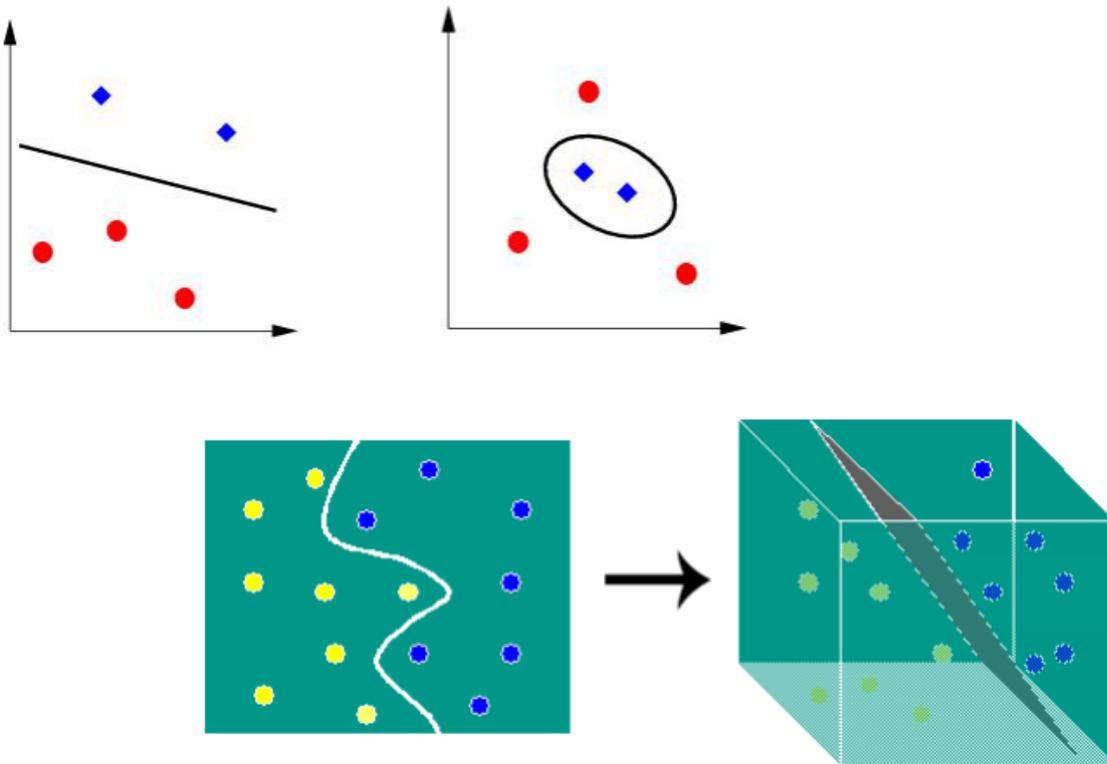


Figure I.3 Cas SVM non linéairement

Par conséquent, nous transformons le problème de séparation non linéaire dans l'espace de représentation en un problème de séparation linéaire dans un espace de ré-description de dimension supérieure. Cette transformation non linéaire est obtenue par une fonction, en effet, il existe plusieurs familles de fonctions noyaux paramétrables, et c'est à l'utilisateur de la SVM de décider quelle fonction est la meilleure pour les exemples de noyaux suivants : Gaussien, polynomiale, Sigmoidale, et Laplacien. [4]

Le but de l'algorithme SVM est de trouver un hyperplan dans un espace à N dimensions (N - le nombre de caractéristiques) qui classe sans ambiguïté les points de données.

Pour séparer ces deux classes de points de données, de nombreux hyperplans possibles peuvent être choisis. Notre objectif est de trouver un plan avec la plus grande marge, c'est-à-dire la plus grande distance entre les points de données des deux classes. La maximisation de la distance de marge fournit un certain renforcement afin que les futurs points de données puissent être classés avec plus de confiance. [4]

Chapitre II

Deep Learning

Chapitre II: Deep Learning

II .1 Introduction:

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

II .2 Définition:

Le Deep Learning, ou apprentissage profond, (également connu sous le nom de deep structured Learning, hierarchical Learning ou deep Machine Learning) est une branche du Machine Learning. Ce sous-ensemble du deep Machine Learning repose sur le principe des réseaux de neurones artificiels (Artificiel Neural Network - ANN), utilisés cependant à une échelle beaucoup plus importante, reposant sur le fait qu'une augmentation du nombre de couches et de neurones d'un ANN augmente les performances de classification. Théorisé à l'origine en 1986 [2], le concept du Deep Learning n'est devenu populaire que récemment, exigeant un très grand nombre de données annotées et une grande puissance de calculs. [2]

-Artificiel Neural Network (ANN) : Un réseau de neurones artificiels est un réseau composé de neurones artificiels interconnectés, où chaque neurone représente une unité de traitement d'information.

II .3 Application du Deep Learning au contexte routier :

Le Deep Learning a montré ses performances sur différents domaines tels que la vision par ordinateur et la médecine. Vu ses résultats encourageants, beaucoup de chercheurs ont appliqué le Deep Learning dans le contexte routier afin d'assurer la mobilité et la sécurité routière:

II .4 Réseaux de neurones convolutifs :

L'architecture de CNN est spécifique au traitement d'images, et Les informations de l'image sont traitées par convolution en différents points, avec plusieurs filtres dans chaque couche.

L'efficacité de l'apprentissage en profondeur est due à ces réseaux, qui incluent les avancées récentes en matière de conduite autonome, d'analyse d'images complexes, etc. Dédié à l'analyse d'images, (CNN) embarque toute la chaîne de traitement représentée sur la figure est à l'opposé de l'apprentissage automatique classique comme indiqué, ces CNN peuvent être considérés comme une boîte noire, en utilisant un ensemble d'images de formation de la même taille pour ajuster plusieurs paramètres réseau pour se concentrer sur une tâche.

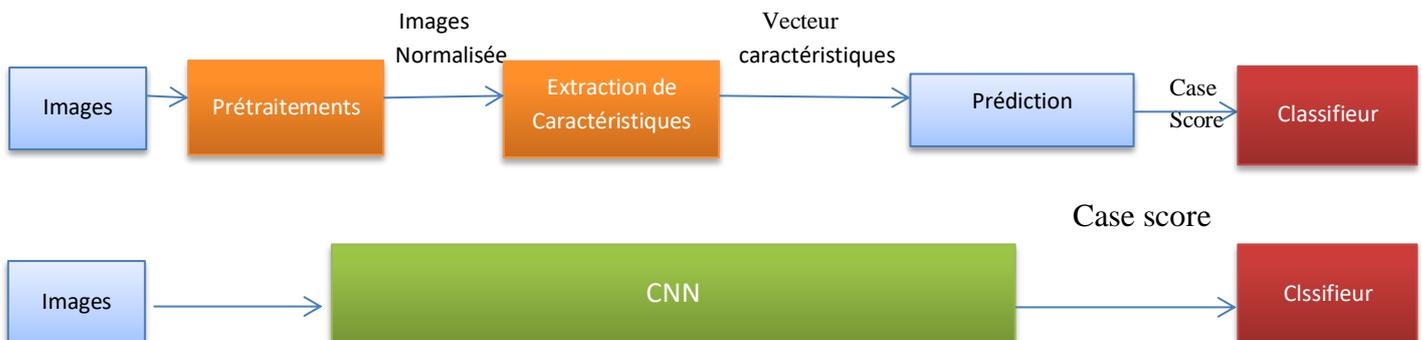


Figure- II .1- Différence entre CNN et Machine Learning Traditionnel pour l'analyse d'image.[7]

Un réseau neuronal convolutif est une opération mathématique appelée convolution ou produit de convolution. Il s'agit d'une opération linéaire. Chaque réseau de neurones convolutifs contient au moins une couche convolutive. Soient f et g deux fonctions définies sur \mathbb{R} . Dans un réseau neuronal convolutif, le premier paramètre f est égal à l'entrée, le deuxième paramètre g est égal au noyau de convolution et la sortie est égale à la carte de caractéristiques. En effet, un réseau de neurones convolutifs utilisant des convolutions discrètes est défini comme suit :

$$S(x) = (f * g)(x) = \sum_{a=-\infty}^{+\infty} f(t)g(x-a) \quad (\text{II}.1)$$

Dans l'apprentissage automatique, l'entrée est toujours un tableau de données multidimensionnel et le noyau est toujours un tableau de paramètres multidimensionnel sera ajusté par l'algorithme d'apprentissage. Ces tableaux sont référencés tenseur. La convolution est toujours utilisée avec des dimensions supérieures à 1. La convolution la plus couramment utilisée est la convolution 2D. Dans ce cas, pour l'entrée d'image I et A noyau K , la convolution discrète s'écrit :

$$S(I, j) = (K * I)(I, j) = \sum_m \sum_n I(i-m, j-n)K(m, n) \quad (\text{II}.2)$$

D'après L'image I et le noyau K sont parcourus dans des sens opposés, en effet, lorsque m augmente, l'indice de I diminue et l'indice de K augmente. La seule raison d'utiliser la convolution est sa commutativité.

Cependant, cette propriété n'est pas importante dans la mise en œuvre du réseau neurones, car les autres fonctions utilisant la convolution ne sont pas non échangeables. C'est pour cette raison que la plupart des bibliothèques réseau.

Les neurones implémentent une fonction de corrélation croisée, mais ils l'appellent convolution. Cet opérateur est le même que la convolution, sauf qu'il perd sa commutativité

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i+m, j+n)K(m, n) \quad (\text{II}.3)$$

Une couche convolutive est caractérisée par :

- la dimension du noyau de convolution, généralement une convolution un à un la dimension est égale à 2, avec un noyau carré.
- le nombre de filtres convolutionnels, C : c'est le nombre de cartes caractéristiques de la sortie de la couche. Ces cartes sont représentées sous forme de tenseurs tridimensionnels H est la hauteur de la carte, W est la largeur et C est le nombre de canaux.
- le pas de convolution (stride), c'est le décalage du noyau de convolution à chaque calcul. Par exemple dans le cas de deux dimensions s =(le noyau de convolution sera déplacé d'un seul pixel vers la droite pour chaque mouvement horizontal du noyau et d'un seul pixel vers le bas pour chaque mouvement vertical du noyau lors de la création de la carte de caractéristiques.
- le padding p . C'est un paramètre qui permet de dépasser la taille de pour appliquer la convolution en ajoutant des pixels autour de l'image.

Dans ce qui suit, et pour faciliter l'analyse et la visualisation, nous utilisons la convolution 2D, avec comme entrée le carré $h = w$, un multiplicateur de convolution k -dimensionnel, un pas s , et un padding p [3]

II .4.1 Les types de convolution :

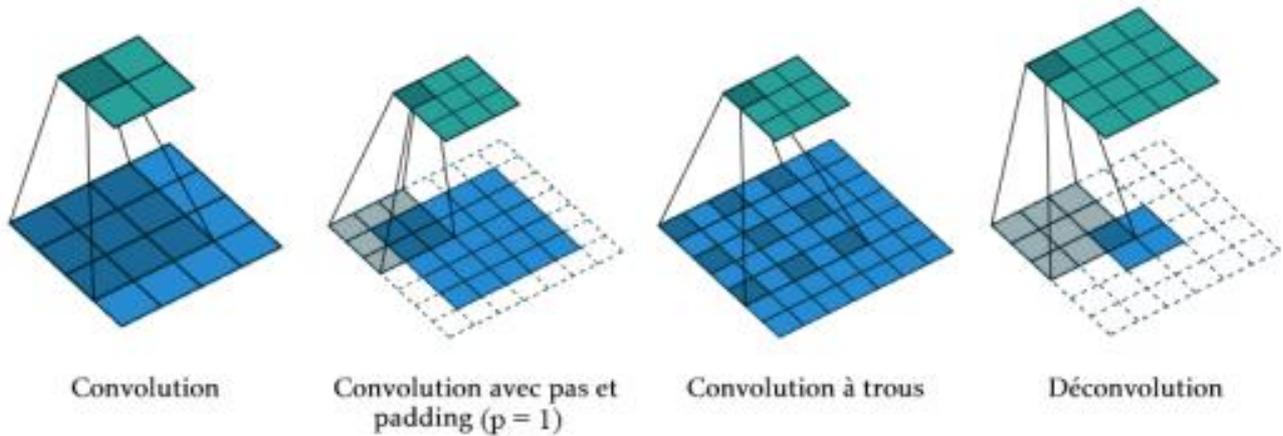


Figure – II.2 - Différents types de convolution

Il existe différents types de convolutions :

II .4.1.1 Convolution classique ou convolution à pas :

Elle est définie par la taille du noyau de convolution k , le pas s et le padding p . Il s’agit de réduire la dimension des cartes d’activation produites par un facteur de s selon l’équation (). Ce type de convolution consiste à faire glisser le noyau de convolution de gauche à droite et du haut au bas avec le pas s jusqu’à ce qu’il arrive à parcourir toute l’image. $O = h$

$$O = \frac{h-k+2p}{s} + 1s \tag{II .4}$$

Une image I de taille $h \times h$ et I_0 de taille $(h + 2p) \times (h + 2p)$ construite à partir de I en ajoutant p pixels aux alentours de l’image, dans le cas où $p = 0$, l’image I_0 sera identique à l’image I .

L’équation de la convolution :

$$s(i, j) = (K * I')(i, j) = \sum_{m=0}^{k-1} \sum_{n=0}^{k-1} I'(s \cdot i + m, s \cdot j + n) K(m, n) \tag{II .5}$$

Avec: $1 \leq h - k + 2p + 1$ et $j \leq h - k + 2p + 1$

II .4.1.2 Convolution à trous ou convolution dilatée :

Elle se base sur un noyau dit éclaté ou d’un facteur de dilatation d .

Il s’agit de parcourir l’image en sautant certains de pixels Les pixels manquants dans le noyau sont remplacés par 0. La taille du pixel manquant dans le noyau est remplacée par 0

$$O = \frac{h - k - (k - 1)(d - 1) + 2p}{s} + 1$$

(II .6)

II .4.1.3 Convolution transposée ou dé convolution :

Il s'oppose à la convolution classique en construisant I à partir de cartes d'activation dont la taille est calculée comme suit :

$$O = (h-1) s + k - 2p \quad (\text{II .7})$$

A partir de () la convolution de déplacement peut être une convolution en escalier avec :

$$S = \frac{1}{\xi} (\xi \in \mathbb{N}^*) \quad (\text{II .8})$$

Ce type de convolution est utilisé dans la phase de l'auto-encodeur à convolution

La convolution exploite trois idées importantes qui peuvent améliorer apprentissage automatique :

II .4.1.3.1 Les interactions convolutives :

Dans un réseau de neurones traditionnel, les couches utilisent une matrice de paramètres multipliée par un seul paramètre pour décrire chaque unité d'entrée et l'interaction entre chaque unité libération. Cela signifie que chaque unité de sortie interagit avec chaque unité d'entrée. Cependant, les réseaux convolutifs ont généralement peu d'interactions. Ceci est réalisé en rendant le noyau plus petit que l'entrée. Par exemple, L'image d'entrée peut contenir des milliers ou des millions de pixels, mais merci

Au niveau de la couche convolutive, nous pouvons détecter de petites caractéristiques utiles, comme les contours où le noyau n'occupe que des dizaines ou des centaines de pixels. Cela permet de stocker moins de paramètres, ce qui réduit le besoin de mémoire pour améliorer l'efficacité de l'algorithme. C'est aussi le calcul de la sortie nécessite moins d'opérations.

II .4.1.3.2 Le partage des paramètres :

Cela inclut l'utilisation du même paramètre pour plusieurs fonctions dans le modèle.

Dans un réseau de neurones traditionnel, chaque élément n'utilise la matrice de pondération qu'une seule fois lors du calcul de la sortie 's couche. Il multiplie simplement un élément de l'entrée. Est part le paramètre permet des pondérations liées, c'est-à-dire la valeur de pondération appliquée une entrée est liée à la valeur d'un autre poids appliqué ailleurs. Est part les paramètres utilisés par l'opération de convolution sont destinés à ne pas être appris chaque position a un ensemble de paramètres distinct, nous apprenons un seuls ensemble. Cela n'affecte pas le temps d'exécution de la propagation vers l'avant mais cela réduit encore les besoins de stockage du modèle.

II .1.3.3 Les représentations équivariantes :

La couche de convolution équivariante à la translation. Une fonction équivalente signifie que si l'entrée change, la sortie change de la même manière. Autrement, une fonction f est équivalente à une fonction g si $f \circ g = g \circ f$.

Dans le cas d'une couche convolutive, une translation de l'image d'entrée provoque une translation des cartes d'activation.

Yann LeCun a introduit les couches convolutives au début du réseau

Les neurones extraient les caractéristiques des images de manière corrélée

Le noyau de convolution prête attention à la corrélation des caractéristiques proportionnelles à l'intervention tardive de la convolution. Autrement dit, plus la convolution est tardive, plus le noyau est complexe et plus les formes détaillées peuvent être détectées. En effet, dans la première couche convolutive, la valeur d'un noyau est initialisée aléatoirement. Ces valeurs sont ensuite mises à jour au fur et à mesure de l'avancement du processus entraînant le réseau pour améliorer les résultats d'extraction de caractéristiques.

La couche convolutionnelle est généralement suivie d'une couche ReLU pour modifier la carte d'entités en sortie, d'où le nom de carte d'entités rectifiée.

L'application de la fonction ReLU à la sortie d'une couche convolutive présente plusieurs avantages.

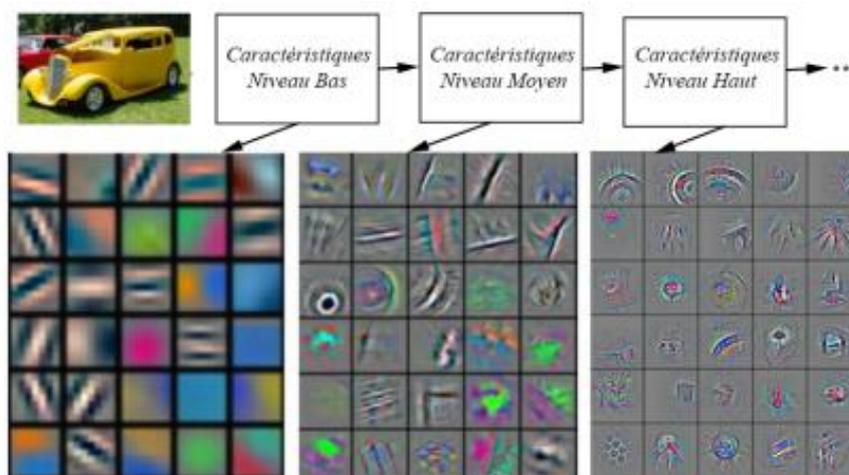


Figure – II.3 - Cartes des caractéristiques. [3]

Les CNN sont similaires aux réseaux de neurones profonds ordinaires, mais leur architecture est spécifique au traitement d'images. En effet, les informations de l'image sont traitées par convolution en différents points, avec plusieurs filtres dans chaque couche. Cette

La popularité de l'apprentissage en profondeur est due à ces réseaux, qui incluent les avancées récentes en matière de conduite autonome, d'analyse d'images complexes, etc. Dédié à l'analyse d'images, (CNN) embarque tous

La chaîne de traitement représentée sur la figure est à l'opposé de l'apprentissage automatique classique Comme indiqué, ces CNN peuvent être considérés comme une boîte noire, en utilisant un ensemble d'images de formation de la même taille pour ajuster plusieurs paramètres réseau pour se concentrer sur une tâche. Nous recommandons par la présente

Rendez rapidement les éléments qui composent un CNN. Pour faciliter la compréhension

Les opérations de ces CNN.

Conception architectural pour les CNN

II .4.2 Couche d'échantillonnage (Pooling) :

La couche d'échantillonnage est responsable de la réduction de la taille spatiale de la carte d'entités, mais elle conserve les informations les plus importantes. Il existe différents types d'échantillonnage, y compris l'échantillonnage Max - ou Max Pooling-, Average Sampling - ou Average Pooling, etc.

L'échantillonnage consiste à appliquer un noyau de taille $n \times n$ sur la carte d'activation en le faisant glisser avec une taille de pas préalablement définie (la taille de pas est généralement égale à Taille du noyau n). Evite les phénomènes de superposition Max Pooling Renvoie la valeur maximale de la partie de l'image couverte par le noyau. au lieu de prendre la valeur maximale, on peut prendre la moyenne de tous les éléments couverts par le noyau, Ceci est assuré par la mutualisation moyenne. Max Pooling supprime le bruit. [5]

D'autre part, Average Pooling effectue simplement une réduction de dimensionnalité comme suit ce mécanisme de réduction du bruit

II .4.2 Réseau Alex Net :

Alex Net a été conçu par Hinton, lauréat du concours Image Net 2012, et son élève Alex Krizhevsky. Aussi après cette année, des réseaux de neurones plus nombreux et plus profonds ont été proposés, tels que l'excellent vgg et Google Le Net. Son taux de précision officiel du modèle de données est de 57,1 %, et le top 1-5 atteint 80,2 %. Ceci est tout à fait remarquable pour les algorithmes de classification d'apprentissage automatique traditionnels. []

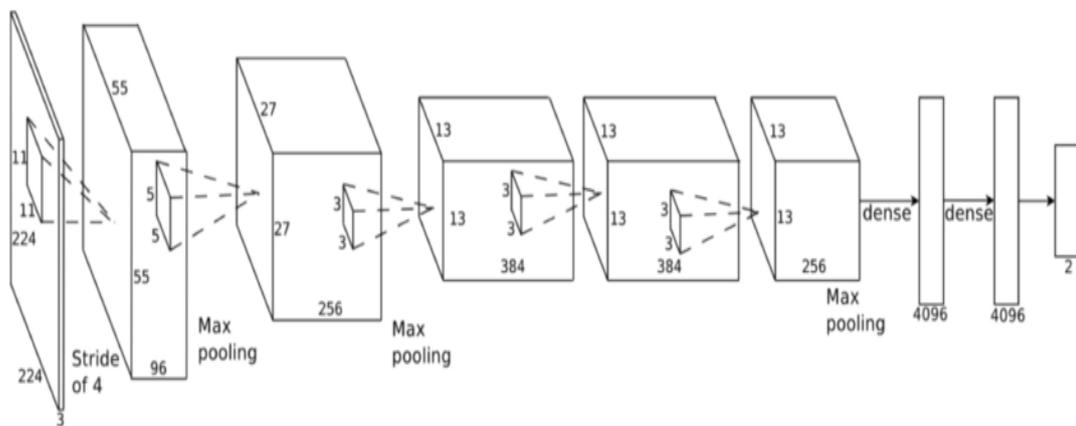


Figure – II.4 - Architecture alexnet [6]

Alexnet a 8 couches. Les 5 premières sont des couches convolutionnelles et les 3 dernières sont des couches entièrement connectées. Entre les deux, nous avons également des "couches" appelées regroupement et activation.

Le schéma du réseau est tiré de l'article original. L'image ci-dessus est l'ordre des couches dans Alexnet. Vous pouvez voir que l'ensemble de problèmes est divisé en deux parties, la moitié fonctionnant sur le GPU 1 et l'autre moitié fonctionnant sur le GPU 2. Les frais généraux de communication sont maintenus à un faible niveau et les performances globales sont bonnes [5]

Chapitre III

Expérimentation et résultats

Chapitre III :

Expérimentation et résultats

III.1 Introduction :

Ce chapitre nous permet de voir d'une manière générale le fonctionnement du programme et l'exécution de notre application les résultats obtenus ainsi que les différentes technologies utilisées pour y parvenir

III.2 Plan de travail:

Pour un grand nombre d'applications civiles et militaires, les routes constituent une partie importante de l'infrastructure urbaine. Par conséquent, leur détection et leur modélisation représentent une étape sémantiquement importante.

Ce travail établit une stratégie de classification d'images de drones à haute résolution spatiale à l'aide de réseaux de neurones convolutifs. Les étapes sont les suivantes :

- Pour diminuer le temps de traitement des pixels sur l'image aérienne ; on procèdera à une superpixelisation. Les superpixels seront les images à l'entrée du réseau de CNN alexnet.
- Pendant la phase d'apprentissage on calculera le vecteur caractéristique FC 7 issu de la couche FC 7 du modèle pré entraîné alexnet (voir tableau) de chaque superpixel (image) ; pour enrichir la base d'apprentissage on a extrait les superpixels sur l'image avec 3 échelles différentes à savoir taille de l'image x1 ; ou x0.5 ;ou x0.75.
- Tous les superpixels ont été redimensionnés à la taille de l'entrée de Alexnet 227*227*3 ;
- on récupérera tous les FC7 (caractéristiques) qui seront les entrées du classifieur SVM pour son apprentissage.
- Pendant la phase de test on calculera les superpixels de l'image ; ils seront les images tests dont on calculera les FC7, et enfin le Classifieur se chargera de les classer.

Type of Layer	No. of Filter	Feature Map Size (height x width x channel)	Kernel Size	No. Of Stride	No. of Padding
Image input layer		227x227x3			
(1st convolutional layer) Relu-1 Cross-channel normalization Max pooling1	96 1	55x55x96 27x27x96	11x11 3x3	4x4 2x2	0x00x0
(2nd convolutional layer) Relu-2 Cross-channel normalization Max pooling2	256 1	27x27x256 13x13x256	5x5 3x3	1x1 2x2	2x20x0
(3rd convolutional layer) Relu-3	384	13x13x384	3x3	1x1	1x1
(4th convolutional layer) Relu-4	384	13x13x384	3x3	1x1	1x1
(5th convolutional layer) Relu-5 Max pooling5	256 1	13x13x256 6x6x256	3x3 3x3	1x1 2x2	1x10x0
Fully connected layer-6(fc6) Relu-6		4096x1 4096x1			
Fully connected layer-7 (fc7) Relu-7		4096 x1 4096 x1			
Fully connected layer-8 (fc8) Softmax layer Output layer		1000x1 1000 class			

Tableau -III.1- les couches d'Alex-Net

III.3 Matériel utilisé :

Processeur : Intel ® CORE™ i5

Capacité Mémoire(RAM) : 8Go

Vitesse d'horloge : 2.70Ghz

Capacité disque dur : 1To

III .4 Logiciel utilisé :

System d'exploitation Windows 10.64 bit

Logiciel de calcul mathématique matlab99R2020b_win64

III .5 Segmentation par superpixelisation :

L’organigramme des étapes de traitement du programme :

Pour l’apprentissage :



Pour le test :



Figure – III.1 L’organigramme des étapes du programme

Pour évaluer le réseau de segmentation sémantique, nous avons besoin d’un ensemble d’images et la collection correspondante d’images étiquetées par pixel (vérité terrain).

Nous avons testé notre programme sur une base de données collectées sur internet. Images collectées drone, composé d’images semi-urbaines et non urbaines. Ils se composent de plusieurs routes avec différentes résolutions, largeurs et formes. Image de la route actuelle Changements de couleur, d’éclairage et de contraste. Nous avons téléchargé l’ensemble de données à partir de [4]

Indices de scores de la segmentation :

Les notes quantitatives sont calculées en comparant les résultats à la vérité terrain au pixel près. Nous nous référons à la vérité terrain comme la zone de route, où chaque image test.

Les indices sont exprimés en vrais positifs (TP), faux positifs Positif (FP) et faux négatif (FN) et vrai négatif (TN) sont définis comme suit :

1- Accuracy = $(TP+TN) / (FN+FP+TP+TN)$ (III.1)

2- Précision = $TP / (TP+FP)$; (III .2)

Le taux de positifs prédits (positive predictive value)

3-Fmeasure = $2*TP / (2*TP+FP+FN)$; (III. 3)

Le score de correspondance des contours de la frontière F1 (BF) indique dans quelle mesure la frontière prédite de chaque classe s'aligne sur la frontière réelle.

4- Specitivity = $TN / (TN+FP)$; (III. 4)

Taux de vrais négatifs (True Négative Rate)

5- Sensitivity (la sensibilité)= $TP / (TP+FN)$; (III. 5)

Correspondant au taux de vrais positifs (true positive rate)

6- MCC (Matthews corrélation coefficient)=
 $(TP*TN-FP*FN) / \sqrt{((TP+FP)*(TP+FN)*(TN+FP)*(TN+FN))}$ (III .6)

7- Jacquard Similarity Coefficient = $TP / (TP + FP ++ FN)$;(III. 7)

valeur prédite	Valeur observée	
	Positif	Négatif
Positif	TP	FP
Négatif	FN	TN

Indicateurs de performance en classification :

Prenons l'exemple d'un classifieur binaire, c'est-à-dire, qui prédit 2 classes notées classe 0 et classe 1.

Pour mesurer les performances de ce classifié, il est d'usage de distinguer 4 types d'éléments classés pour la classe voulue :

- Vrai positif VP : Elément de la classe 1 correctement prédit
- Vrai négatif VN : Elément de la classe 0 correctement prédit
- Faux positif FP : Elément de la classe 1 mal prédit
- Faux négatif FN : Elément de la classe 0 mal prédit

III.6 Classification supervisée par SVM :

Voici quelques images des superpixels obtenues à différentes échelles de la base d'apprentissage :



Superpixels No road (32*32*3)



Superpixels road (32*32*3)

Figure- III.2 résultats de classification

Dans notre travail en utilise 12 images et une seule image test

Nous avons d'abord crée le réseau alexnet sur matlab (figure) puis extrait la couche FC7 pour la caractérisation .Le classifieur SVM prendra en charge l'étape de classification.

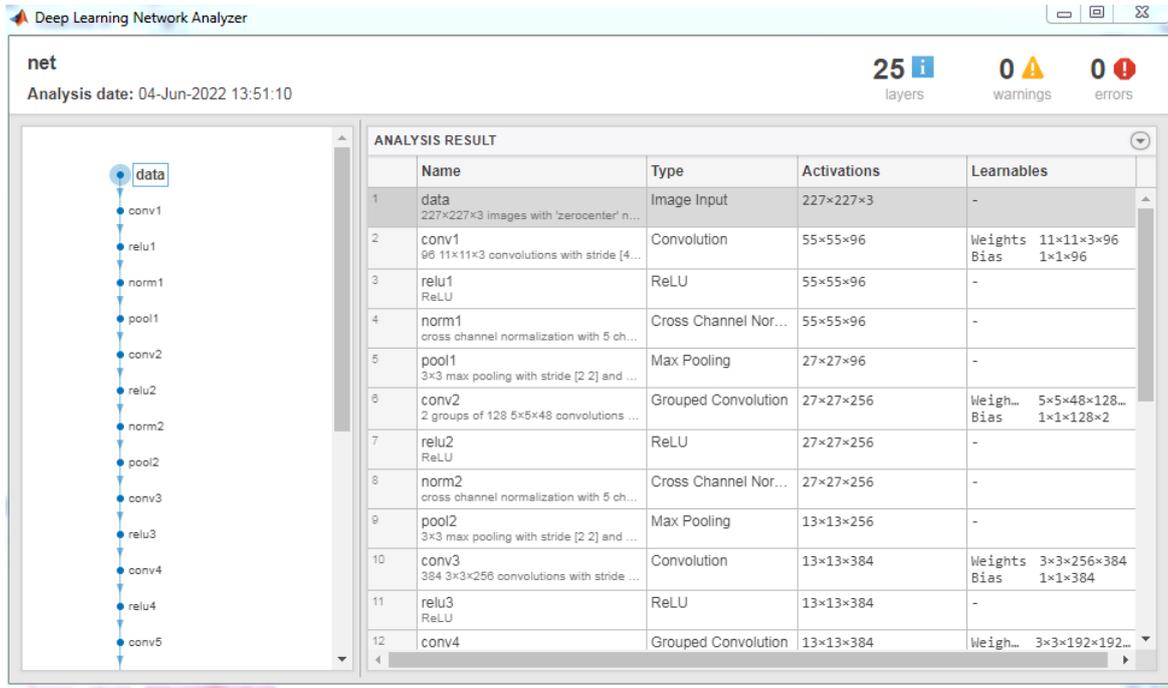


Figure – III.3- Architecture alexnet

III.7 Les résultats de segmentation

Nous avons choisi de tester 5 valeurs pour le nombre de superpixels sur l'image noté L: 200, 300, 400, 500, 600. les résultats sont les suivants:

L=200

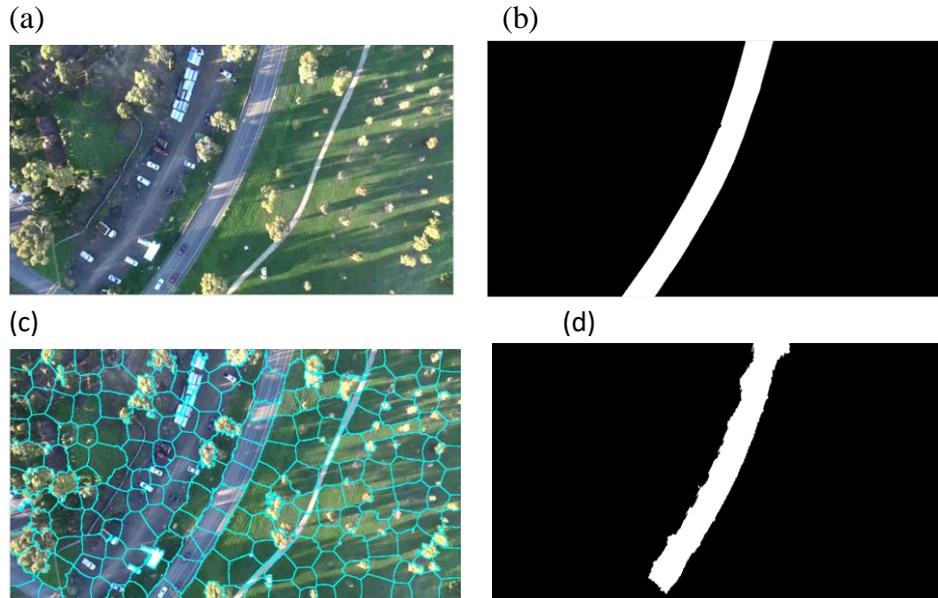


Figure III.4-résultat de la segmentation de l'image 0.10 avec L=200

- (a) image originale.
- (b) image vérité.
- (c) image superpixelisée.
- (d) image résultat avec classification SVM et FC CNN

L=300

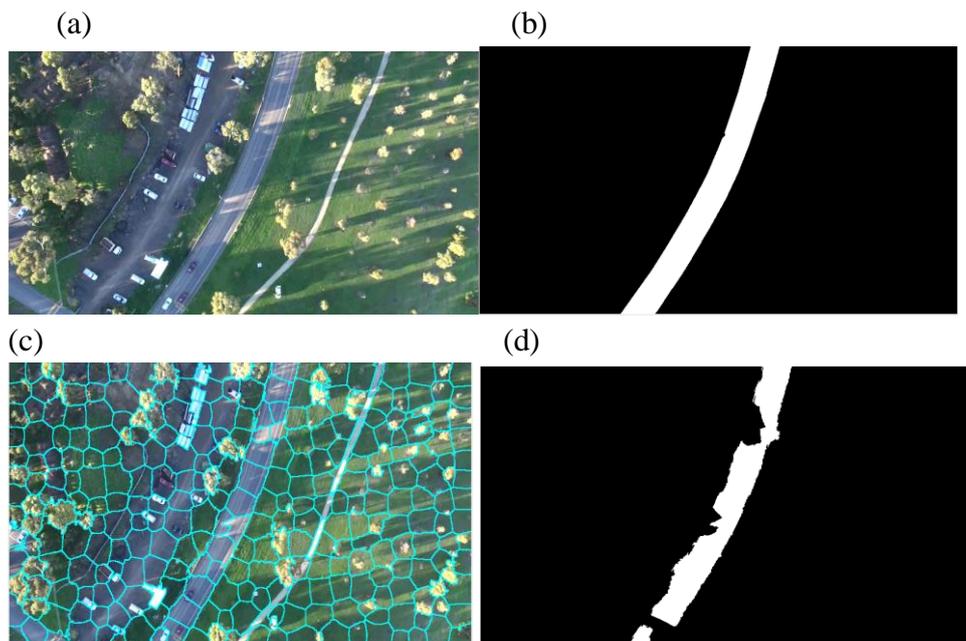


Figure III.5 -résultat de la segmentation l'image 0.10 avec L= 300

- (a) image originale.
- (b) image vérité.
- (c) image superpixilisée.
- (d) image résultat avec classification SVM et FC CNN.

L=400

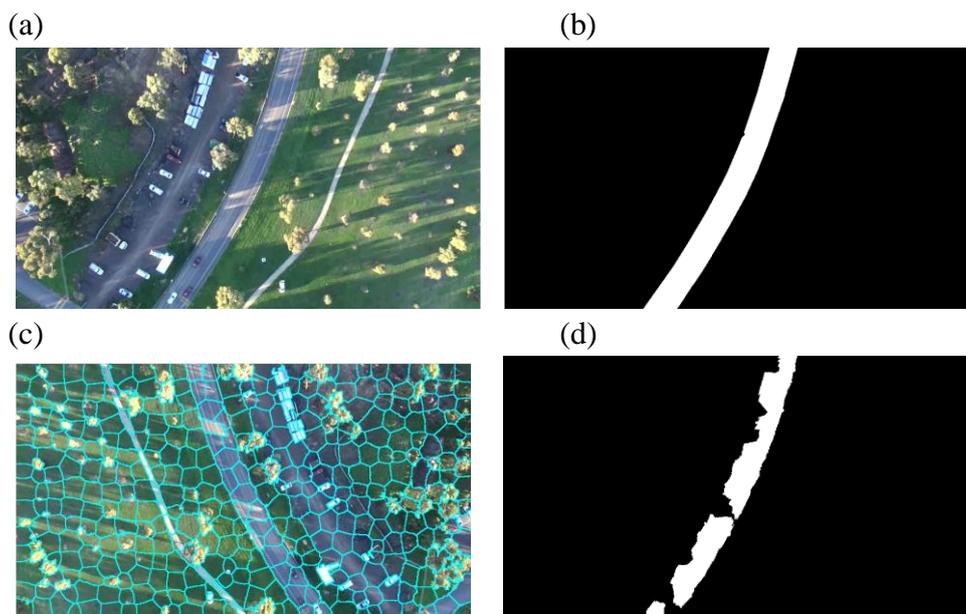


Figure III.6 -résultat de la segmentation de l'image 0.10 avec L=400

- (a) image originale.
- (b) image vérité.
- (c) image superpixilisée.
- (d) image résultat avec classification SVM et FC CNN.

L=500

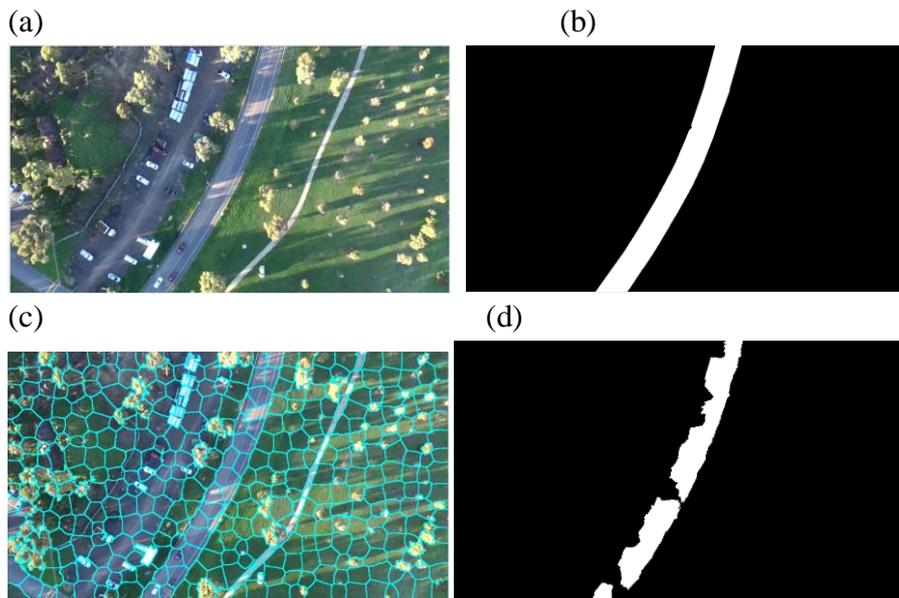


Figure III.7 -résultat de la segmentation de l'image 0.10 avec L=500

- (a) image originale.
- (b) image vérité.
- (c) image superpixelisée.
- (d) image résultat avec classification svm et fc cnn.

L=600

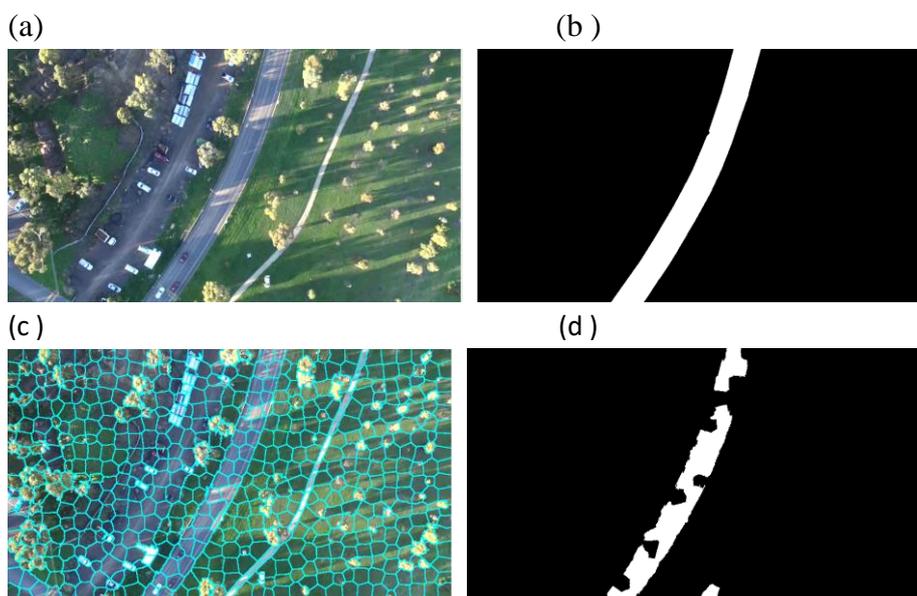


Figure III.8 -résultat de la segmentation de l'image 0.10 avec L=600

- (a) image originale.
- (b) image vérité.
- (c) image superpixelisée.
- (d) image résultat avec classification SVM et FC CNN.

Résultats expérimentaux :

Une fois le processus d'apprentissage terminé, nous allons discuter les résultats obtenus :

L \ Paramètres Superpixels	Accracy	Sensitivity	F mesures	Dice	Jaccard	Specivity
200	0.9401	0	0	0	0	1
300	0.9321	0	0	0	0	1
400	0.9321	0	0	0	0	1
500	0.9321	0	0	0	0	1
600	0.9320	0	0	0	0	1

Tableau III.2 : les scores pour l'image test 010

En comparant les résultats de la segmentation des trois images le programme SVM a pu extraire la zone de la route pour l'image_010. Par contre le choix, les scores montrent une légère amélioration pour L=200.

III.6.Conclusion :

Dans ce chapitre, nous avons présenté les différents résultats obtenus par la technique de segmentation proposée. La méthode de superpixelisation d'image sur des images aériennes .on crée alors des imgettes de taille raisonnable à classer par svm.

Conclusion Générale :

Dans ce mémoire nous avons abordé le problème de la segmentation d'images.

Dans le premier temps, nous effectuons une pré segmentation par la méthode SLIC qui aboutit à une segmentation de l'image en régions homogènes telle que une région est composé de l'ensemble des pixels connexes possèdent des mêmes propriétés au sens d'un prédicat d'homogénéité donné. la méthode SLIC nécessite le réglage du paramètre L comme nombre de superpixels qui jouera sur la taille de ces derniers. Nous obtenons ainsi des images de taille réduite pour la classification.

Les réseaux neuronaux profonds apprennent des caractéristiques de haut niveau dans les couches cachées. C'est l'une des plus grandes forces des CNN et cela permet de réduire les besoins en ingénierie des caractéristiques. En fait, les images subissent plusieurs transformations. Tout d'abord elles passent par de nombreuses couches convolutionnelles où le réseau apprend de nouvelles caractéristiques de plus en plus complexes. Ensuite, l'information de l'image transformée passe par les couches entièrement connectées et est transformée en une classification. Les caractéristiques de haut niveau peuvent être récupérées à partir des dernières couches convolutives FC7.

Pour éviter un temps d'apprentissage très long en utilisant le CNN, et puisque toutes les couches sont responsables de l'apprentissage de certaines caractéristiques des images, les caractéristiques peuvent être extraites du réseau à tout moment pendant le processus d'apprentissage. Nous utilisons ces caractéristiques extraites comme données d'entrée pour un modèle de classification avec des Machines à Vecteur de Support (SVM).

Les résultats obtenus sur l'image UAV, montrent que la méthode arrive à détecter la route à condition de choisir une riche base d'apprentissage et les bons paramètres.

Bibliographie:

- [1] HAL Id: tel-01949597 <https://tel.archives-ouvertes.fr/tel-01949597> Submitted on 10 Dec 2018
- [2] Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua ET S. Susstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(11):2274–2282, 2012
- [3] HAL Id: tel-03022934 <https://tel.archives-ouvertes.fr/tel-03022934> Submitted on 25 Nov 2020
- [4] http://georges.gardarin.free.fr/Surveys_M/Survey_SVM.pdf
- [6] https://www.researchgate.net/figure/AlexNet-CNN-architecture-layers_fig1_318168077
- [5] https://www.google.com/search?sxsrf=ALiCzsaB_8_QmiSaHAeuCp_Op9YIdIEwvw:1654586579354&q=Alex+Net+architecture&sa=X&ved=2ahUKEwjFvb
- [7] Svm <https://zestedesavoir.com/tutoriels/1760/un-peu-de-machine-learning-avec-les-svm/>