

# Fiche de lecture

*Les réseaux de neurones convolutifs dans le monde des cylindrées*

Auteur  
GHOUIBI Ghassen

Encadrée  
Mr. Jean Jacques Mariage

## Table des matières

<b>I</b>	<b>Introduction</b>	<b>2</b>
I .1	Problématique . . . . .	2
I .2	Contexte . . . . .	2
<b>II</b>	<b>Réseaux de neurones convolutifs</b>	<b>2</b>
Couche de convolution	. . . . .	3
Couche de pooling	. . . . .	3
Couche de correction	. . . . .	4
Couche FC	. . . . .	4
Couche de perte	. . . . .	4
<b>III</b>	<b>Solution CNN</b>	<b>4</b>
III .1	Reconnaissance de signalisation en temps réel . . . . .	4
Implémentation proposer	. . . . .	4
Conclusion	. . . . .	5
III .2	Détection de collision . . . . .	6
Implémentation proposer	. . . . .	6
Conclusion	. . . . .	7
<b>IV</b>	<b>Mon avis</b>	<b>8</b>
<b>V</b>	<b>Conclusion</b>	<b>8</b>

## I Introduction

Le mixage entre l'informatique et la mécanique se fait en toute douceur si on prend comme exemple les voitures.

Dans certaines tâches les humaines sont plus rapides puisqu'ils ont acquis un point de vue humain qu'une machine n'a pas. C'est pour cela par exemple qu'on vous demande de prouvez votre humanité à travers un formulaire facile, CAPTCHA ou autre où le but est par exemple de trouver les images qui contiennent des voitures.

Pour les ordinateurs c'est un peu plus compliqué d'identifier dans des images qui contiennent des panneaux par exemple vu qu'il doivent absorber une quantité de données importantes et se doter de performance de calcul.

Les entreprises spécialisées comme **GAFA** et **Tesla** investissent des millions de dollars dans le domaine du Big Data qui est en explosion continue. Dans notre rapport on va se focaliser sur les voitures autonomes.

En effet, la reconnaissance de millions d'images ou de vidéos pour identifier un contenu précis est coûteux et demande une démarche particulière et pourtant il existe très peu de techniques pour analyser ces données de manière automatisée et efficace.

### I.1 Problématique

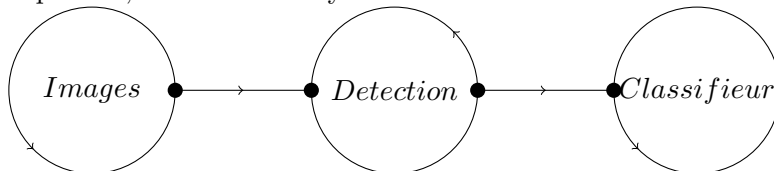
Les solutions de classification d'image présentent actuellement dans les voitures intelligentes vont-elles permettre de réduire les accidents de la route ?

### I.2 Contexte

Dans ce rapport on peut définir notre contexte comme un problème de classification, d'où on va présenter quelque variante des réseaux de neurones convolutifs qui est une approche très populaire pour résoudre des problèmes de reconnaissance de formes.

## II Réseaux de neurones convolutifs

Les réseaux de neurones convolutifs sont des méthodes d'apprentissage supervisé, ils suivent ce cycle ci-dessous :



Plus en détails la réception des images en entrées ensuite les réseaux de neurones convolutifs détectent les features de chaque image puis entraînent un classifieur dessus.

Comme détaillé auparavant on s'intéresse au problème de classification, Le réseau va donc calculer à partir de nos entrées un score pour chaque classe, Finalement la classe sera affectée à l'objet d'entrée qui à le meilleur score.

Une architecture de réseau de neurones se composent à travers un empilement de couches de traitement :

1. couche de convolution qui traite les données d'un champ récepteur.
2. couche de pooling qui permet de compresser l'information.
3. couche de correction en référence à la fonction d'activation.
4. couche entièrement connectée qui est une couche de type perceptron.
5. couche de perte (LOSS).

### Couche de convolution

La couche convolution est un bloc de construction indispensable d'un CNN.

Elle contient trois paramètres qui sont :

- La Profondeur
- Le Pas
- La marge

L'entrée d'une couche convolution est soit une image soit la sortie d'une autre couche de convolution. Pour calculer le nombre de neurones de sortie voici la formule :

$$W_0 = \frac{W - K + 2P}{S} + 1 \quad (1)$$

- $W$  volume d'entrée
- $K$  nombre de champs récepteurs
- $S$  le pas
- $P$  taille de la marge

### Couche de pooling

Cette partie permet de découper une image en entrée en petite série de rectangles de  $n$  pixels.

### Couche de correction

Pour améliorer le traitement entre les couches une fonction mathématique appelée fonction d'activation de neurones sur les sorties par exemple la correction `Unité Linéaire Rectifiée`, la correction par `tangente hyperbolique`, fonction `sigmoïde`.

### Couche FC

Les neurones d'une couche entièrement connectée sont connectés avec toutes les sorties de la couche précédente.

### Couche de perte

La couche de perte montre la manière où l'entraînement du réseau pénalise l'écart entre le signal prévu et réel, en effet plusieurs fonctions de pertes peuvent être utilisées par exemple `Softmax`, `Euclidienne` et `Entropie croisée sigmoïde`.

## III Solution CNN

### III .1 Reconnaissance de signalisation en temps réel

Dans cette partie on se base sur l'article de Alexander Shustanov et Pavel Yakimov, qui expliquent l'implémentation de reconnaissance de signalisation à travers un réseau de neurones convolutifs.

La reconnaissance d'une signalisation se fait sur deux étapes, d'abord la localisation ensuite la classification.

La solution proposée dans l'article montre que l'utilisation de l'algorithme GHT a donné un résultat final de 97.3% de précision testé sur le dataset GTSRB<sup>1</sup>

### Implémentation proposer

Comme détaillé auparavant l'article utilise la bibliothèque Deep Learning de TensorFlow sur le dataset GTSRB avec l'architecture finale suivante :

Convolutional, stride 2, kernel 3x3x16
Convolutional, stride 2, kernel 3x3x32
Convolutional, stride 2, kernel 3x3x64
Fully connected-512
SoftMax

Le test sur d'autres architectures a donné un résultat de 0.9 d'occurrence de classification, en revanche on a décidé de réduire le nombre de layers. Cette

---

1. <http://benchmark.ini.rub.de/?section=gtsrb&subsection=news>

architecture montre qu'un seul layer n'est pas suffisant pour avoir l'occurrence souhaitée.

Méthode	Occurance	FPS
Sliding window + SVN	100%	1
Modified GHT with preprocessing + CNN	99.94%	50
ConvNet	99.55%	38
Modified GHT with preprocessing	97.3%	43
Modified GHT without preprocessing	89.3%	25
Viola-Jones	90.81%	15
HOG	70.33%	20

Les algorithmes développés dans cet article ont été testé sur des vidéos prises avec un Anroid Nvidia Shield Tablet installé dans une voiture voici un exemple :



Finalement pour l'évaluation du temps d'exécution pour la classification voici les ressources utilisées :

Hardware	Training	Clasification d'une image (64x64)
Nvidia GeForce GTX 650	7 min	0.05 ms
Nvidia GeForce GTX 650M	12 min	0.14 ms
Intel Core i7	16 min	0.37 ms

## Conclusion

Dans cet article l'implémentation d'algorithme de classification pour la reconnaissance de trafic montre un très bon resultat avec 99.94% d'image classée correctement.

### III .2 Détection de collision

Dans cette partie on va se baser sur l'article écrit par XIAOHUI HUANG,PAN HE de l'université de Floride aux états-unis.

De nos jours les caméras de surveillance submergent les rues des grandes villes pour des raisons de sécurité par exemple. En effet, utiliser ces caméras pour détecter en temps réel les collisions présentent un défi vu que les scénarios de conduite sont trop variés. Par exemple chaque conducteur aura un réflexe différent par rapport à une situation de danger.

#### Implémentation proposer

En général un outil de surveillance vidéo devrait répondre à plusieurs exigences :

- Détecter tous les véhicules (arrêtés ou en mouvement)
- Classifier les véhicules détectés en catégories (bus,voitures,moto,vélo)
- Extraire des caractéristiques spatiales (mouvement,vitesse, trajectoire),détection d'anomalies
- Détection de condition de circulation (vitesses variables entre les voies,trafic léger) , condition d'éclairage (ensoleillé,couvert,nuit,pluvieux)

L'article utilise TNAD<sup>2</sup> qui contient trois types de données vidéos des intersections de trafic qui sont utilisés pour détection d'accident et aussi pour des tâches de surveillance.

Les méthodes de détection d'objets sont généralement des approches basées sur l'apprentissage automatique ou l'apprentissage approfondi sur des propositions régionales (zones ou partitions) ou des approches d'apprentissage approfondi de bout en bout.

Au début à l'aide R-CNN on va résoudre le problème de sélection d'un grand nombre de régions pour générer un certain nombre de propositions régionales et ensuite classer ces régions à l'aide de SVM.

L'algorithme de recherche sélective est ensuite utilisé pour les propositions de région et pour générer une sous-segmentation initiale ainsi que de nombreuses régions candidates. Ensuite il va permettre de combiner récursivement les régions similaires en plus grandes, et finalement utiliser les régions générées pour produire la version finale de proposition de région candidate.

---

2. Traffic Near Accident Dataset

Ces propositions de régions candidates sont transformées en carré et alimentées en réseau de neurones convolutionnels qui prennent le rôle d'un extracteur de fonctionnalités.

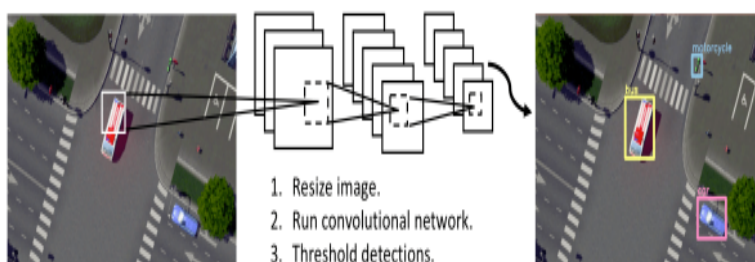
La couche en sortie sert à alimenter un SVM pour classer la présence de l'objet dans cette proposition de région candidate.

L'apprentissage approfondi de bout en bout YOLO<sup>3</sup> utilise un seul réseau convolutionnel pour prédire ce qu'on appelle bounding-box, des boîtes englobantes, ainsi que les probabilités correspondantes.

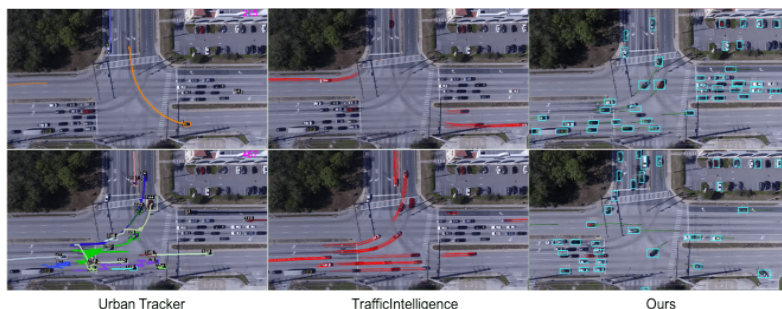
Il prend d'abord l'image et la divise en une grille  $S \times S$ , dans chacune des parties de la grille, nous prenons  $m$  boîtes englobantes. Pour chacune des boîtes englobantes à échelle multiple. Le réseau génère une probabilité de classe et des valeurs de décalage pour le cadre de sélection.

Ensuite, il sélectionne les boîtes englobantes qui ont la probabilité de classe au-dessus d'une valeur seuil et les utilise pour localiser l'objet dans l'image.

Voici une image qui nous permet de mieux comprendre :



## Conclusion



Les expériences dans cet article ont montré l'avantage du modèle proposé par rapport à d'autres modèles avec des performances quantitatives à des fréquences d'images élevées.

---

3. un algorithme de détection d'objet qui considère la détection comme un problème de régression de bout en bout



## IV Mon avis

Pour minimiser les accidents de la route il sera indispensable d'utiliser la technologie. En effet ceci pourra être résolu avec des voitures autonomes sachant qu'il existe quelques prototypes de voitures intelligentes. Cependant la passation du modèle classic de voitures aux voitures électriques se fait de manière très lente, les recherches actuelles ne font que prouver la stabilité et l'efficacité des algorithmes comme CNN. Au final la transition vers l'automatisation des voitures peut donc être longue. De plus, ces voitures permettront peut être de protéger les humains et réduire de beaucoup le nombre d'accidents mais il faudra s'attendre à d'autres problèmes de la part des voitures autonomes.

## V Conclusion

Les travaux effectués dans les articles sont très intéressants, en apprenant dans ces articles on constate que chaque problème de classification ou autre demande certain pré-requis, il faudrait donc étudier d'autres algorithmes afin d'avoir une connaissance large et pouvoir mettre en concurrence ces derniers.

Tesla dépassent de loin toutes les concurrences dans le monde des cylindrées mais est-ce-qu'on arrivera à avoir des voitures complètement autonomes dans le futur proche ouvert au public avec des prix abordables.