

———— CHAPITRE 3 ————

ÉTAT DE L'ART DE LA RECONNAISSANCE
DES CHIFFRES

[Mycours.com](https://www.mycours.com)

3.1 INTRODUCTION

La reconnaissance des caractères est souvent considérée comme l'un des principaux problèmes d'apprentissage du réseau neuronal. Ce chapitre sera totalement consacré à la présentation de la reconnaissance des chiffres qui est la base de notre travail.

Comme nous n'avons pas pu trouver des travaux identiques à notre problématique, de reconnaissance des chiffres des compteurs électroniques de tours présents dans les administrations algériennes, nous nous focaliserons sur des travaux similaires comme la reconnaissance des chiffres des plaques minéralogiques.

3.2 DÉTECTION ET RECONNAISSANCE DES PLAQUES MINÉRALOGIQUES

La lecture automatique de plaques minéralogiques ou Lecture automatisée de plaques d'immatriculation, consiste à mettre en place une combinaison de moyens électroniques et informatiques, en utilisant des techniques de traitement d'images et de vision par ordinateur, pour extraire le numéro d'identification d'après l'image de la plaque, sous format de caractères. La figure (3.1) illustre un exemple d'un système de reconnaissance¹.

1. <https://www.alphanumeric-vision.com>

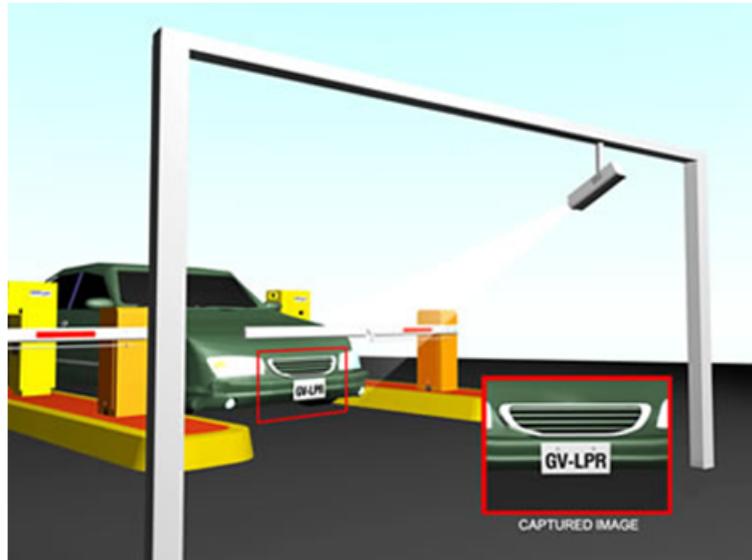


FIGURE 3.1: Schéma du système de reconnaissance

3.3 ÉLÉMENTS D'UN SYSTÈME TYPIQUE DE LECTURE AUTOMATIQUE DE PLAQUES MINÉRALOGIQUES

3.3.1 Matériel

Camera

Elle sert à prendre une image du véhicule [29].

Éclairage

C'est une lumière projetée sur la scène, pour éclairer la plaque. Ce système permet une utilisation du jour comme de la nuit du système de reconnaissance [29].

Ordinateur

Un ordinateur qui exécute le système de reconnaissance, lit l'image, identifie et détecte la plaque dans l'image [29].

3.3.2 Logiciel

Frame Grabber

C'est une carte d'interface entre le pc et la camera, qui permet le transfert des images vers le Pc [29].

Base de données

Elle inclut les données et les résultats de la reconnaissance [29].

3.4 DOMAINES D'APPLICATION DU SYSTÈME DE LECTURE AUTOMATIQUE DE PLAQUES MINÉRALOGIQUES

Il existe de nombreuses applications qui utilisent de tels systèmes de reconnaissance, dont certaines sont décrites ci-dessous :

3.4.1 Contrôle de vitesse

Le numéro d'immatriculation est utilisé pour imposer une amende aux véhicules qui roulent à grande vitesse, à l'utilisation illégale des voies et à la détection de véhicules volées ou recherchées [30]. La figure (3.2) montre un exemple de voiture capturée



FIGURE 3.2: Application du système dans le contrôle de vitesse [30]

3.4.2 Contrôle d'accès au parking

Le système est utilisé pour l'entrée automatique au parking, pour les membres déjà enregistrés, ou dans certains cas, pour le calcul des frais de stationnement des voitures. Lorsque une voiture entre, la plaque d'immatriculation est reconnue et marquée. A la sortie, elle est relue et la durée de stationnement est facturée pour le conducteur [30]. La figure (3.3) illustre un parking contrôlé par une application.



FIGURE 3.3: Système de contrôle d'accès au parking [30]

3.4.3 Contrôle Passage aux frontières

Cette application aide le registre des entrées ou des sorties dans un pays et peut être utilisée, pour surveiller les postes frontières [30]. Cette installation est illustrée à la figure (3.4). Chaque information de véhicule est enregistrée dans une base de données.



FIGURE 3.4: Frontière franco-allemande [30]

3.5 ÉTAPES DU PROCESSUS DE LECTURE AUTOMATIQUE DE PLAQUES MINÉRALOGIQUES

Chaque système de reconnaissance automatique de plaques d'immatriculation est constitué des étapes suivantes :

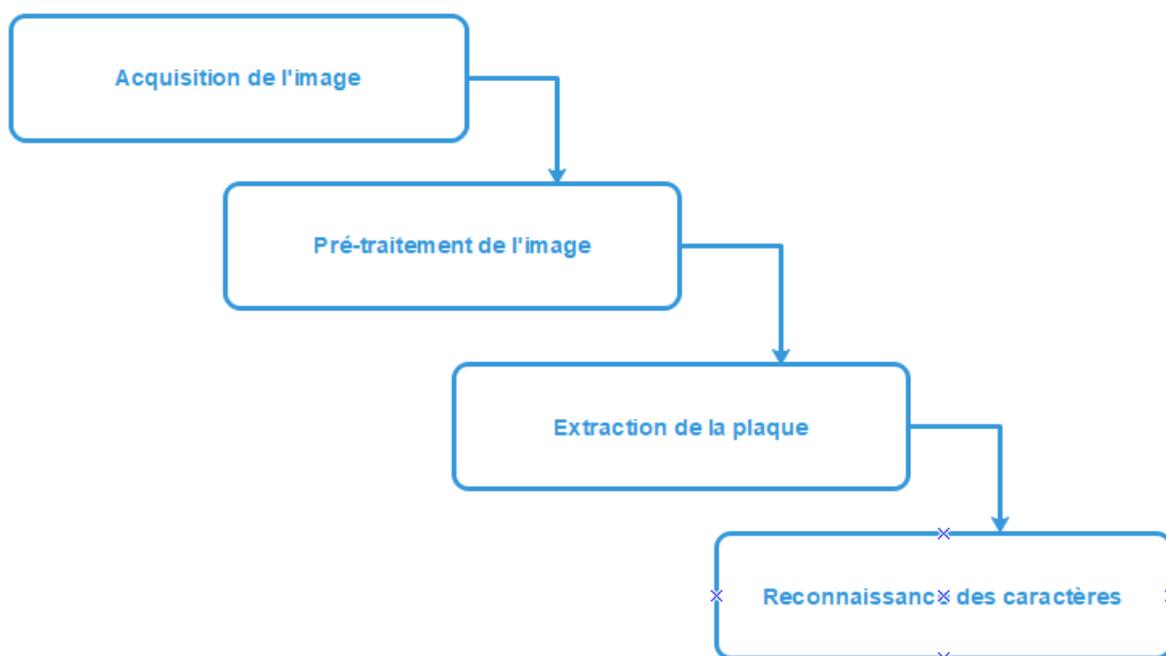


FIGURE 3.5: Étapes du processus de lecture de plaques

3.5.1 Acquisition de l'image

La première étape consiste à acquérir l'image ou tout simplement, à capturer des images et des vidéos du véhicule lors de son passage.

3.5.2 Pré-traitement d'image

Cette étape consiste à traiter l'image, principalement, la convertir en niveau de gris pour que la phase du Deep Learning soit plus rapide. Une image en niveaux de gris de M pixels de hauteur et de N pixels de largeur est représentée par une matrice de type de

données double de taille $M * N$, dont les valeurs ont été mises à l'échelle pour représenter les intensités. Chaque valeur de pixel des images en niveaux de gris, est comprise entre 0 (noir) et 255 (blanc), Voici deux méthodes pour convertir une image en niveau de gris :

Méthode 1 : moyenne

$$\text{Grayscale} = (B + G + R)/3$$

Méthode 2 :

$$\text{Grayscale} = (B * 0.3) + (G * 0.59) + (R * 0.11)$$

La figure (3.6) représente une image avant et après conversion



(a) Image Originale



(b) Image en niveau de gris

FIGURE 3.6: Image d'une voiture avant et après conversion en niveau de gris [31]

3.5.3 Extraction de la plaque d'immatriculation

L'étape suivante consiste à extraire la plaque d'immatriculation de l'image. Pour cela, il existe différentes méthodes, qui se basent sur la recherche des caractéristiques de la plaque dans l'image comme la couleur, la forme etc [32]. Pour détecter la position de la plaque dans l'image, il existe différentes méthodes pour l'extraction de la plaque qui peuvent être classées en différentes catégories :

3.5.3.1 Extraction basée sur le détecteur YOLO

YOLO (You Only Look Once) est l'un des algorithmes les plus rapides pour la détection d'objet qui est basé sur un réseau de neurone profond [33], son principe est simple :

il faut entraîner ce réseau sur un ensemble d'images annotées à la main, c'est-à-dire déterminer la position de l'objet à extraire, puis après l'entraînement l'algorithme va détecter automatiquement la position de l'objet. Un exemple est illustré dans la figure (3.7) :



FIGURE 3.7: Plaque détectée par YOLO [34]

3.5.3.2 *Extraction basée sur le contour*

Étant donné, que la plaque d'immatriculation a normalement une forme rectangulaire avec un rapport d'aspect connu, elle peut être extraite en trouvant tous les rectangles possibles dans l'image.

Les méthodes de détection des contours sont couramment utilisées pour trouver ces rectangles. Le filtre Sobel est utilisé pour détecter les bords. En raison de la transition de couleur entre la plaque d'immatriculation et la carrosserie, la limite de la plaque d'immatriculation est représentée par des bords dans l'image. Les bords des plaques d'immatriculation doivent être continus [35].

3.5.3.3 *Extraction basée sur les caractéristiques globales de l'image*

L'analyse des composants connectés (CCA) est une technique importante dans le traitement d'images binaires. Elle scanne une image binaire et étiquette ses pixels en composants, en fonction de la connectivité des pixels (Mesures spatiales), telles que l'aire et le rapport d'aspect, sont l'utilisant couramment pour l'extraction de plaques d'immatriculation

Les objets connectés, qui ont les mêmes caractéristiques géométriques que la plaque, sont choisis pour être candidats.

Cet algorithme peut échouer dans le cas d'images de mauvaise qualité, ce qui entraîne des contours déformés.

La corrélation croisée 2D est utilisée pour trouver les plaques d'immatriculation.

La corrélation croisée 2D avec un modèle de plaque d'immatriculation pré-enregistrée est effectuée sur toute l'image pour localiser la zone de plaque d'immatriculation la plus probable.

L'extraction de plaques d'immatriculation à l'aide de la corrélation avec un modèle est indépendante de la position de la plaque d'immatriculation dans l'image. Cependant, la corrélation croisée 2D prend du temps [35].

3.5.3.4 *Extraction basée sur la texture*

Ce type de méthode dépend de la présence des caractères dans la plaque d'immatriculation, ce qui entraîne un changement dans le niveau de gris entre la couleur des caractères et la couleur de fond de plaque. Il en résulte également, une zone de densité de bord élevée en raison de la transition de couleur.

Les filtres de Gabor sont l'un des principaux outils de l'analyse de texture. Cette technique a l'avantage d'analyser la texture dans des orientations et des échelles illimitées. Cependant, cette méthode prend du temps.

La fréquence spatiale est identifiée en utilisant la transformée de Fourier discrète (DFT) et la transformée en ondelettes (DWT), car elles produisent des harmoniques qui sont détectées dans l'analyse du spectre. La DFT et la DWT sont utilisées, de manière rangée, pour détecter la position horizontale de la plaque et de manière colonne pour détecter la position verticale.

Toutes les méthodes basées sur la texture ont l'avantage de détecter la plaque d'immatriculation même si sa limite est déformée. Cependant, ces méthodes sont complexes sur le plan informatique, en particulier lorsqu'il y a de nombreux bords, comme dans le cas d'un fond complexe ou dans des conditions d'éclairage différentes [35].

3.5.3.5 *Extraction basée sur les caractéristiques de la couleur*

Étant donné que certains pays ont des couleurs spécifiques pour leurs plaques d'immatriculation, certains travaux impliquent l'extraction de plaques d'immatriculation en localisant leurs couleurs dans l'image.

L'idée de base est que la combinaison de couleurs d'une plaque est unique, et cette combinaison se produit presque uniquement dans une région de plaque [35].

Comparaison entre ces méthodes

Méthode	Avantage	Inconvénient
Extraction basée sur le contour	Simple, rapide et direct	Mise en oeuvre difficile pour les images complexe car elle est trop sensible aux bords indésirables
Extraction basée sur les caractéristiques globales de l'image	Direct, indépendant de la position de la plaque d'immatriculation.	Peut générer des objets déformés.
Extraction basée sur la texture	Être capable de détecter même si les bornes sont déformées.	
Extraction basée sur les caractéristiques de la couleur	Être capable de détecter les plaques d'immatriculation inclinées et déformées.	RGB est limité aux conditions d'éclairage, HLS est sensible aux bruit

TABLE 3.1: Comparaison entre les méthodes extraction de la plaque d'immatriculation [35]

3.5.4 Reconnaissance des caractères

Après l'extraction de la plaque d'immatriculation, la dernière étape du processus de la lecture automatique des plaques minéralogiques est la reconnaissance des chiffres de la plaque minéralogique. Pour cela, nous allons présenter quelques modèles de la littérature concernant cette problématique.

3.5.4.1 *Modèle de Selmi et al*

Configuration du modèle

Le tableau (3.2) illustre l'architecture du modèle proposé [36].

Couche	Paramètre
Entré	Image en niveau de gris 32X32 Pixel
Couche de Convolution (ReLu)	Profondeur : 32 Fenêtre : 5X5 Pas : 1
Couche de Pooling	Fenêtre : 5X5
Couche de Convolution (ReLu)	Profondeur : 64 Fenêtre : 3X3 Pas : 1
Couche de Pooling	Fenêtre : 5X5
Couche de Convolution (ReLu)	Profondeur : 128 Fenêtre : 3X3 Pas : 1
Couche de Convolution (ReLu)	Profondeur : 256 Fenêtre : 3X3 Pas : 1
Couche de Pooling	Fenêtre : 5X5
Couche Entièrement Connecté FC (ReLu)	Taille : 1024 neurones
Couche de Dropout	0.5
Couche Entièrement Connecté FC (Softmax)	Taille : 37 neurones

TABLE 3.2: Configuration du Modèle Selmi [36]

Description du modèle

Le modèle proposé par Selmi et al Dans [36] en 2017 est un réseau de neurone convolutionnel (CNN). L'entrée du modèle est une image de taille 32X32 pixel. Il est constitué principalement de dix couches, huit couches de convolution activées par une fonction d'activation ReLu, trois couches de pooling, une couche Dropout et deux couches entièrement connectées. Une d'entre elles est la sortie composée de 37 neurones c'est à dire 37 classes (10 chiffres de 0 à 9 et 26 lettres et la catégorie négative). Cette couche est activée par la fonction Softmax.

Selmi a testé son modèle sur deux datasets, le premier est le dataset AOLP qui contient 2049 images de plaques minéralogiques originaires de Taiwan. Elle est divisée en trois parties : Patrouille de route (RP) avec 611 images, Contrôle d'accès (AC) avec 681 images et Law Enforcement (LE) avec 757 images.

Le deuxième dataset est issu de PKU (Pekin University). Il contient 3977 images de plaques chinoises [36].

3.5.4.2 *Modèle de Wu et al*

Configuration du modèle

Le tableau (3.3) illustre l'architecture du modèle proposé [37].

Couche	Paramètre
Entré	Image 136X36 Pixel
Bloc de convolution (Couche de convolution Relu +Batch Normalisation Max Pooling)	Profondeur : 32 Fenêtre : 3X3 Pas 1 Fenêtre : 2X2
8 Dense Block (Couche de convolution Relu +Batch Normalisation)	Profondeur : 128 Fenêtre : 3X3
Bloc de transition (Couche de convolution Relu +Batch Normalisation Moyenne Pooling)	Profondeur : 128 Fenêtre : 1X1
8 Dense Block (Couche de convolution Relu +Batch Normalisation)	Profondeur : 192 Fenêtre : 3X3
Bloc de transition (Couche de convolution Relu +Batch Normalisation Moyenne Pooling)	Profondeur : 128 Fenêtre : 1X1
8 Dense Block (Couche de convolution Relu +Batch Normalisation)	Profondeur : 192 Fenêtre : 3X3
Couche Entièrement Connecté FC (Softmax)	Taille : 68 neurones

TABLE 3.3: Configuration du Modèle Wu [37]

Description du modèle

Ce modèle est proposé par l'équipe de chercheurs Wu et al en 2018 [37]. Il est basé sur DenseNet avec une sortie 68 neurones donc 68 classes (31 caractères chinois, dix chiffres et 26 lettres).

Principalement, il est constitué d'un bloc de convolution activé avec une fonction ReLu, 3 dense Block, 2 bloc de transition et une couche entièrement connectée (sortie) activée avec la fonction d'activation Softmax.L'entré du modèle est une image de taille 136X36.

Le modèle est testé sur deux ensembles d'images :le premier Dataset-1[ref] qui est constitué de 203774 images et le second est le dataset AOLP utilisé dans le modèle de Selmi [37].

3.5.4.3 *Modèle de Spahel et al*

Configuration du modèle

Le tableau (3.4) illustre l'architecture du modèle proposé [38].

Couche	Paramètre
Entré	Image en niveau de gris 200X40 Pixel
3 Bloc de convolution (Couche de convolution Relu +Batch Normalisation)	Profondeur : 32 Fenêtre : 5X5 Pas : 1
Max Pooling	Fenêtre : 2X2
3 Bloc de convolution (Couche de convolution Relu +Batch Normalisation)	Profondeur : 64 Fenêtre : 3X3 Pas : 1
Max Pooling	Fenêtre : 2X2
3 Bloc de convolution (Couche de convolution Relu +Batch Normalisation)	Profondeur : 128 Fenêtre : 3X3 Pas : 1
Max Pooling	Fenêtre : 2X2
8 Branche de sortie (2 X Couche Entièrement Connecté (Softmax))	Taille : 37 neurones

TABLE 3.4: Configuration du Modèle Spahel [38]

Description du modèle

C'est un modèle basé sur un réseau de neurone convolutionnel (CNN) proposé par Spahle et al en 2017 [38]. Le modèle proposé a une architecture spéciale au niveau des couches entièrement connectées, qui sont représentées par plusieurs branches.

Principalement, le modèle est constitué de trois couches de convolution avec des fonctions d'activation ReLu, trois couches de pooling puis huit branches. Pour chaque branche, elle contient deux couches entièrement connectées (FC) avec la fonction Softmax, l'entrée du modèle est une image de dimension 200X40 Pixels [38].

Ce modèle est testé sur trois Datasets différents : ReId, HDR et Svoboda [39].

3.6 CONCLUSION

Dans ce chapitre, nous avons introduit la reconnaissance des caractères, précisément la reconnaissance des chiffres. Puis nous avons présenté la problématique de détection et reconnaissance des plaques minéralogique, qui est une problématique assez proche de la nôtre, en détaillant toutes les étapes du processus de lecture automatique de plaques minéralogiques. Enfin, nous avons présenté certains modèles de la littérature en ce qui concerne la reconnaissance des plaques minéralogiques. Dans le chapitre suivant, nous allons présenter notre modèle qui permettra la reconnaissance des numéros de tours sur un afficheur sept segments.

MCours.com