

Université de Sfax Faculté des Sciences de Sfax Département d'Informatique et des Communications		Filière I3 AU : 2025-2026
Enseignante : Sahar DAMMAK		

Projet : Système de reconnaissance automatique des plaques d'immatriculation

Objectif du projet :

Concevoir et implémenter un système capable de détecter, segmenter et reconnaître automatiquement les numéros d'immatriculation des véhicules à partir d'images ou de vidéos.

Ce projet combine amélioration d'image, détection d'objets et reconnaissance de caractères, couvrant plusieurs aspects fondamentaux de la vision par ordinateur.

Description générale

Le système peut se reposer sur une approche en plusieurs étapes :

1. Amélioration de la qualité de l'image

- Réduction de bruit,
- Correction de contraste,
- Amélioration de luminance ...

2. Détection de véhicules :

- Dans un premier temps, un modèle de détection d'objets identifie les véhicules présents dans l'image ou la vidéo.
- Les véhicules détectés sont classés (voiture, camion, moto...) et encadrés dans des boîtes englobantes.
- Cette étape permet de réduire la zone de recherche et d'éviter les détections erronées sur d'autres objets.

3. Détection de la plaque d'immatriculation

- Sur chaque région correspondant à un véhicule détecté, une deuxième détection est appliquée pour localiser précisément la plaque d'immatriculation.
- Cette étape peut utiliser un modèle spécifique entraîné pour la détection de plaques.

4. Prétraitement de la plaque

- La plaque extraite est redressée, mise à l'échelle et améliorée (augmentation du contraste, suppression du bruit, binarisation).
- Ces transformations facilitent la lecture des caractères, même en conditions difficiles (éclairage, angle, saleté).

5. Reconnaissance des caractères (OCR)

- Un module OCR lit les caractères de la plaque et extrait le texte correspondant.

- Le résultat est ensuite validé et corrigé à l'aide de règles syntaxiques selon le format des plaques tunisiennes.

Jeux de données

1. Détection de véhicules

Datasets recommandés :

Nom du dataset	Description	Taille / Format	Lien
UA-DETRAC	~140 000 images et 100 vidéos annotées pour la détection et le suivi de véhicules (voitures, bus, camions).	Très complet, annotations de bounding boxes.	http://detrac-db.rit.albany.edu/
Stanford Cars	16 185 images de voitures, 196 catégories (modèles, angles, etc.).	Très bonne diversité d'angles et de couleurs.	https://ai.stanford.edu/~jkrause/cars/car_dataset.html
KITTI Vision Benchmark	Données réelles d'environnement routier (voitures, routes, piétons).	Images + LiDAR + annotations 2D/3D.	http://www.cvlibs.net/datasets/kitti/

2. Détection de plaques d'immatriculation

Datasets recommandés :

Nom du dataset	Description	Pays	Taille / Format	Lien
OpenALPR Benchmark Dataset	Contient des milliers d'images de plaques de différents pays.	International	Annoté pour la détection et OCR.	https://github.com/openalpr/benchmarks
Tunisian Licensed Plates Dataset	709 images, chacune avec une plaque annotée, donc 709 objets « plaque »	Tunisie	images + annotations (bounding boxes) de plaques d'immatriculation tunisiennes	https://www.kaggle.com/datasets/achrafkhazri/labeled-licence-plates-dataset?resource=download
SSIG SegPlate Dataset	2 000 images de plaques brésiliennes avec annotations	Brésil	Segmenté et labellisé.	https://github.com/rodrigoberriel/ssig-dataset

	précises des caractères.			
AOLP (Academic Open License Plate Dataset)	2 049 images de plaques taïwanaises sous trois scénarios : accès, trafic, patrouille.	Taïwan	Données réelles, variées.	https://mmlab.ie.cuhk.edu.hk/datasets/aolp/
Chinese City Parking Dataset (CCPD)	+250 000 images de plaques chinoises dans des conditions variées.	Chine	Très grand dataset, annotations YOLO disponibles.	https://github.com/detectR_ecog/CCPD

3. Reconnaissance des caractères (OCR de plaques)

Datasets recommandées

Nom du dataset	Description	Contenu	Lien
SSIG Dataset (caractères segmentés)	Plaques avec caractères isolés et étiquetés.	Lettres/chiffres annotés individuellement.	https://github.com/rodrigoberriel/ssig-dataset
CCPD-OCR Subset	Sous-ensemble de CCPD pour OCR direct.	Plaques complètes + texte.	https://github.com/detectRecog/CCPD
Synthetic License Plate Dataset (SLPD)	Jeu de données synthétique de plaques générées automatiquement (plusieurs polices, angles, éclairages).	Plaques artificielles + texte.	https://github.com/matthewearl/deep-anpr
OCR Datasets génériques (EMNIST, SynthText)	Pour préentraîner un modèle OCR général.	Lettres/chiffres imprimés.	https://www.nist.gov/itl/products-and-services/emnist-dataset