



UNIVERSITE NATIONALE DU VIETNAM
INSTITUT FRANCOPHONE INTERNATIONALE



Option: Systèmes Intelligents et Multimédia (SIM)

VISION PAR ORDINATEUR

RAPPORT DE TP2

THEME : Sliding Window

Rédigée et présentée par:

FANANG NDONG Claude Ericka: étudiante

Sous la supervision de : NGUYEN Thi Oanh

coordinateur du module CV

Année Académique : 2024-2025

Introduction

La détection d'objets est une composante cruciale de la vision par ordinateur, avec des implications significatives dans de nombreux domaines tels que la sécurité, la surveillance, et l'automobile autonome. Parmi les approches de détection d'objets, la méthode de sliding window est largement utilisée pour sa simplicité et son efficacité. Ce rapport vise à fournir une évaluation détaillée de cette technique, en se concentrant sur deux aspects majeurs : la détection de visages en utilisant la méthode de Viola-Jones et la construction d'un modèle de détection d'objets spécifique, en l'occurrence, des voitures.

1- Détection de Visages avec la Méthode de Viola-Jones

La méthode de Viola-Jones est une technique classique de détection de visages largement utilisée en raison de sa rapidité et de sa précision. Dans cette section, nous avons entrepris une évaluation approfondie de sa performance en utilisant la technique de sliding windows.

1.1.Méthodologie

Nous avons commencé par charger un modèle pré-entraîné « **haarcascade_frontalface_default.xml** » pour la détection de visages fourni par OpenCV. Ensuite, nous avons appliqué la technique de sliding windows sur une image spécifique, en variant la taille de la fenêtre et le taux de recouvrement. L'objectif était d'analyser l'impact de ces paramètres sur la qualité de la détection.

1.2. Paramètres et Expérimentations

- **La taille de la fenêtre** : Définie par les variables ``min_face_size`` et ``max_face_size``, cette taille influence la sensibilité de la détection, une fenêtre trop petite peut manquer certains visages, tandis qu'une fenêtre trop grande peut inclure des régions non pertinentes.
- **Le pas de déplacement (overlap)** : Contrôlé par la variable ``overlap_percent``, ce paramètre détermine la quantité de chevauchement entre les fenêtres successives. Un

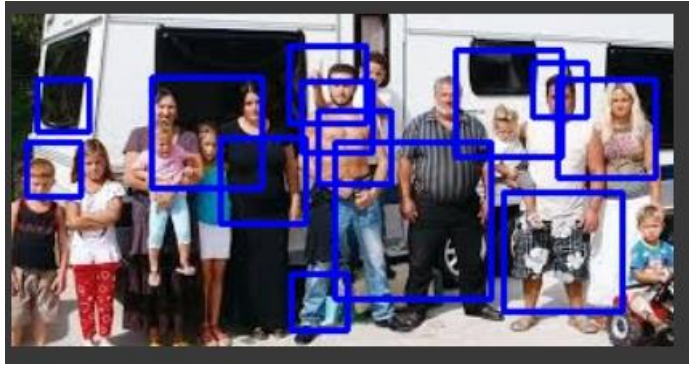
chevauchement plus élevé peut conduire à une meilleure détection mais nécessite également plus de calculs.

- **Le modèle de détection** : Utilisé via la méthode ``detectMultiScale()`` avec le modèle pré-entraîné de Viola-Jones, la qualité de la détection dépend de la performance de ce modèle. Des modèles différents peuvent avoir des performances variées sur différents types d'images.

1.3. Résultat1



1.4. Résultat 2 :



1.5. Interprétation des résultats:

Les résultats obtenus ont mis en lumière l'importance critique de la taille de la fenêtre et de l'overlap dans le processus de détection. En faisant varier ces paramètres, nous avons observé que des fenêtres trop petites pouvaient entraîner des détections plus précises, tandis que des fenêtres trop grandes pouvaient générer des faux positifs. De plus, un taux de recouvrement plus élevé a généralement conduit à une meilleure précision de détection, bien que cela se traduise par une augmentation du coût computationnel.

2- Construction d'un Modèle de Détection de Voitures

Dans la deuxième partie de l'expérience, nous avons abordé la construction d'un modèle de détection d'objets en utilisant la méthode de sliding windows, avec un accent particulier sur la détection de voitures.

2.1. Méthodologie

Pour ce faire, nous avons utilisé le modèle DETR (DEtection TRansformer) avec l'architecture ResNet-50 de Facebook, spécialement conçu pour la détection d'objets, et l'avons entraîné sur un ensemble de données contenant des images de voitures. Ensuite, nous avons appliqué la technique de sliding windows pour détecter les voitures dans une image spécifique.

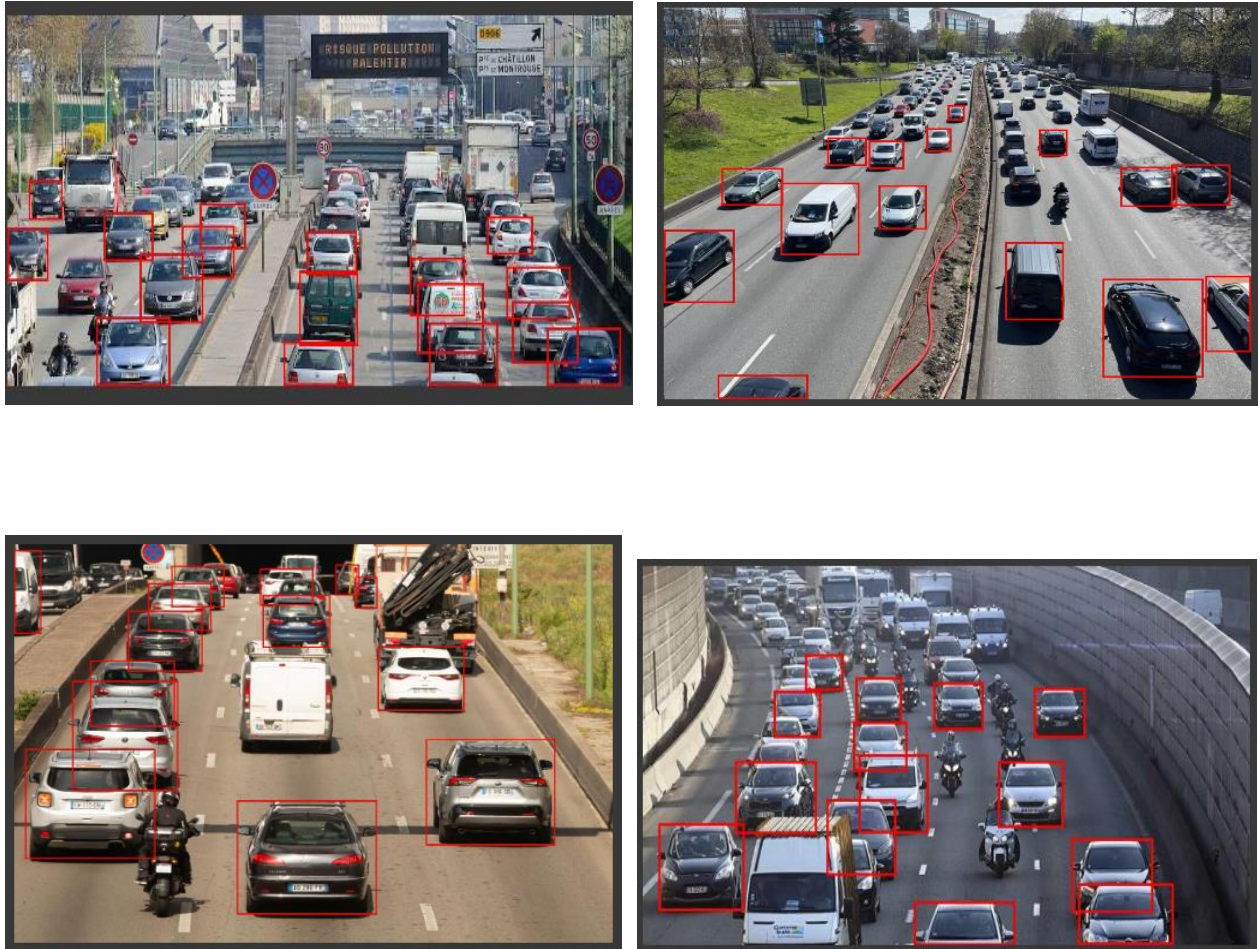
2.2. Paramètres et Expérimentations :

Plusieurs paramètres peuvent impacter le résultat en sortie de ce code :

- **Seuil de confiance (threshold) :** Le seuil de confiance détermine le niveau de certitude requis pour qu'un objet soit considéré comme détecté. Dans notre code, le seuil est défini à 0.9 avec `threshold=0.9`, ce qui signifie que seuls les objets avec une confiance supérieure à 90% seront considérés comme détectés.
- **Modèle de détection (DETR) :** Le choix du modèle de détection a un impact significatif sur la précision et la capacité à détecter différents types d'objets. Dans notre cas, le modèle utilisé est DETR avec l'architecture ResNet-50.
- **Taille de l'image d'entrée :** La taille de l'image d'entrée peut influencer la qualité de la détection. Des images de résolutions différentes peuvent nécessiter des ajustements dans le traitement et la détection.
- **Taille de la boîte englobante:** Le modèle DETR peut prédire des boîtes englobantes de différentes tailles pour les objets détectés. Des paramètres internes au modèle, tels que la taille des filtres convolutifs utilisés lors de la détection, peuvent influencer la taille des boîtes englobantes prédites.
- **Taille des boîtes englobantes prédites :** La taille des boîtes englobantes prédites par le modèle peut également être influencée par des facteurs tels que la résolution de l'image, la complexité de la scène et la variabilité des objets présents.
- **Pré-traitement de l'image :** Les étapes de pré-traitement de l'image, effectuées par le processeur de l'image DETR, peuvent également affecter les résultats de la détection. Cela peut inclure des transformations telles que le redimensionnement, la normalisation et la conversion des canaux de couleur.

En ajustant ces paramètres, il est possible d'optimiser la performance de la détection d'objets pour différentes tâches et conditions d'application.

2.3. Résultats Obtenus



2.4. Interprétation des résultats :

Les résultats ont démontré que le modèle DETR est capable de détecter efficacement les voitures dans une variété de contextes, y compris des scènes complexes avec plusieurs voitures. Cependant, nous avons également constaté que des ajustements fins des paramètres tels que la taille de la fenêtre et le seuil de confiance étaient nécessaires pour obtenir les meilleurs résultats, en fonction des caractéristiques spécifiques de chaque scène.

Conclusion

En conclusion, cette expérience a fourni un aperçu détaillé de l'efficacité et des défis associés à l'utilisation de la technique de sliding windows pour la détection d'objets. La méthode de Viola-Jones s'est avérée efficace pour la détection de visages, tandis que le modèle DETR a démontré sa capacité à détecter avec précision les voitures. Toutefois, il est essentiel de souligner que le réglage précis des paramètres est crucial pour optimiser les performances dans chaque cas d'application. Cette étude souligne l'importance de comprendre en profondeur les différentes approches de détection d'objets pour relever efficacement les défis de la vision par ordinateur dans divers domaines d'application.

Lien Colab :

<https://colab.research.google.com/drive/17taQR19jiBPiIwrgzBGMwllp6zAFDkPh#scrollTo=sGZN4ouTYRv4>