

Mise en Cohérence des Objectifs du TIPE :

Vision par ordinateur appliquée à la détection de panneaux

Antoine Groudiev

1 Titre de votre sujet TIPE (20 mots)

Vision par ordinateur appliquée à la détection de panneaux

2 Quelle est votre motivation pour le choix du sujet ? (50 mots)

La vision par ordinateur est une problématique centrale dans la construction d'un véhicule autonome. Véritable lien entre le monde extérieur et l'ordinateur, la détection d'objet permet à une voiture sans conducteur de se situer dans son environnement direct.

3 En quoi votre étude s'inscrit-elle dans le thème de l'année ? (50 mots)

La prédiction de l'évolution future des villes est impossible sans analyser le développement des véhicules autonomes, qui définiront peut-être les moyens de transport de demain. L'étude du domaine de la vision par ordinateur permet ainsi de s'intéresser à une problématique décisive de l'avenir de la ville.

4 Bibliographie commentée (650 mots)

La vision par ordinateur est la branche de l'informatique visant à permettre à un ordinateur de transformer une image, sous la forme d'un tableau de nombres, en informations permettant des décisions. [1]

Les applications de ce domaine d'étude sont nombreuses et déjà présentes de nos jours, entre autres pour l'imagerie médicale, la détection faciale et la reconnaissance de l'écriture. [1][2] Une application notable en lien avec la ville est le développement des véhicules autonomes, qui dépendent de la vision par ordinateur pour détecter les autres voitures, les panneaux, les piétons et potentiels obstacles à éviter. [1]

La détection d'objet est par essence une discipline difficile, à cause de deux contraintes opposées. À cause de l'immense quantité d'information reçue, le traitement d'une image est un processus très long. Cependant, l'analyse d'un flux vidéo en temps réel nécessite une fréquence de rafraîchissement élevée, pour pouvoir, dans le cas d'un panneau par exemple, freiner à temps. Couplé à l'importance d'une détection juste, pour ne pas freiner brusquement en cas de fausse alerte, la détection de panneaux requiert donc des algorithmes efficaces et précis, sans marge d'erreur.

L'algorithme le plus célèbre [2] de détection de visages est celui développé par Paul Viola et Michael Jones [3] en 2001. Il présente deux idées novatrices pour accélérer la vitesse de détection, les concepts d'image intégrale et de boosting[4]. Il est ainsi théoriquement possible d'implémenter une telle méthode, puis d'entraîner le modèle avec un jeu d'images bien choisi, pour être capable de détecter n'importe quel objet dans une image, en alliant vitesse et précision. La même méthode peut être ainsi appliquée à la détection de panneaux, s'inscrivant dans notre problématique.

Si le fonctionnement théorique de l'algorithme est décrit dans le papier original de Viola et Jones [3], son implémentation en pratique n'est pas développée ; la lecture de d'autres articles [5][6] permet de mieux appréhender la démarche concrète d'implémentation.

L'implémentation de l'algorithme est suivie de la question de son entraînement : le programme, pour fonctionner, nécessite une phase d'apprentissage de plusieurs heures. Des milliers d'images, dites positives, de l'objet

à détecter, ainsi que négatives, ne comprenant pas l'objet à détecter, doivent être fournies à l'algorithme pour que ce dernier apprenne à reconnaître l'objet en question. Il faut ainsi récupérer, recadrer, trier à la main, ajuster l'éclairage des images [3][5] pour garantir le bon fonctionnement du programme, ce qui constitue une longue et importante étape de la construction du détecteur.

Enfin, la dernière phase de la démarche scientifique de la création d'un tel outil est une étape de tests, visant à déterminer l'exactitude de l'algorithme, à l'aide de méthodes de calcul propres à l'apprentissage automatique [7].

5 Problématique retenue (50 mots)

La problématique de mon travail est d'implémenter, entraîner, et tester un algorithme de vision par ordinateur, capable de détecter un panneau STOP dans une vidéo en respectant des contraintes de vitesse et d'exactitude.

6 Objectifs du TIPE (100 mots)

1. Implémentation en Python de la méthode de Viola et Jones
2. Constitution d'un jeu d'images d'entraînement, pré-traitement de ces données, et entraînement de l'algorithme
3. Application du programme à un flux vidéo, mesure de l'exactitude de l'algorithme et confrontation aux contraintes réelles

7 Mots clés (5 mots-clés, français & anglais)

Vision par ordinateur	Apprentissage automatique	Intelligence artificielle	Pré-traitement d'images	Classification supervisée
Computer vision	Automatic learning	Artificial Intelligence	Image pre-processing	Supervised classification

Références

- [1] IBM, *What is computer vision ?*, <https://www.ibm.com/topics/computer-vision>
- [2] Richard Szeliski, *Computer Vision : Algorithms and Applications, 2nd ed. (2022)*, <https://szeliski.org/Book/>
- [3] Paul Viola, Michael Jones, *Rapid Object Detection using a Boosted Cascade of Simple Features*, Conference on Computer Vision and Pattern Recognition, <https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf>
- [4] Michael Pound, Sean Riley, Computerphile, *Detecting Faces (Viola Jones Algorithm)*, <https://www.youtube.com/watch?v=uEJ71VlUmMQ&t=15s>
- [5] Yi-Qing Wang, *An Analysis of the Viola-Jones Face Detection Algorithm*, IPOL, https://www.ipol.im/pub/art/2014/104/?utm_source=doi
- [6] Anmol Parande, *Understanding and Implementing the Viola-Jones Image Classification Algorithm*, Medium, DataDrivenInvestors, <https://medium.datadriveninvestor.com/understanding-and-implementing-the-viola-jones-image-classification-algorithm-85621f7fe2>
- [7] David M W Powers, *Evaluation : From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation*, Journal of Machine Learning Technologies, https://web.archive.org/web/20191114213255/https://www.flinders.edu.au/science_engineering/fms/School-CSEM/publications/tech_reps-research_artfcts/TRRA_2007.pdf