

RAPPORT

Chef d'Oeuvre

Spécifications

COURDY-BAHSOUN CLÉMENCE - DEKER SYLVAIN
KOTTATH SANDEEP - MOUSSA NAHOR - YAZI RYMA

Groupe : Ré-id

M2 IGAI

1 Introduction

Le projet proposé par le client, Alain Crouzil, est le développement d'une interface permettant de faire de la ré-identification de véhicule dans un corpus de vidéo. Il s'agit donc de détecter les véhicules dans les vidéos et de les comparer au véhicule recherché afin de fournir une liste pertinente des vidéos pouvant contenir le véhicule recherché.

Pour la ré-identification de véhicule la méthode Fusion of Attributes and colors features (FACT) produit de bons résultats en temps et en précision. Elle repose sur le calcul de trois descripteurs de l'image (texture, couleur et sémantique)

La méthode BOW-SIFT permet de calculer les descripteurs SIFT des points d'intérêts d'une image, puis de les classifier sous la forme d'un dictionnaire par apprentissage non supervisé hiérarchique avec la méthode des k-moyenne, avec un nombre de classe fixé à 10000. L'histogramme du nombre de descripteurs par classe devient alors le descripteur de l'image.

La méthode BOW-CN permet de calculer un descripteur de couleur sous la forme d'un vecteur obtenu par le calcul d'un histogramme sur 250 classe obtenue par un algorithme de k-moyenne. Les descripteurs permettant de calculer l'histogramme sont les valeurs des canaux des pixels.

Le réseau de neurones convolutifs GoogleNet prend une image en entrée et en lui appliquant des convolutions il permet d'obtenir différents descripteurs qui permettent une caractérisation de l'image propre au réseau : les poids de la couche finale du réseau vont être utilisé comme descripteur sémantique du véhicule.

2 Cahier des charges

Lors de la recette, le livrable devra impérativement contenir les fonctionnalités suivantes :

1. Détection de véhicules
2. Ré-identification
3. Une interface utilisateur

2.1 Exigences fonctionnelles

Détection de véhicules

Il s'agit d'une fonctionnalité prévu en pré traitement, en préalable à la ré-identification. On choisit arbitrairement de ne pas l'intégrer en temps réel dans

l'application.

1. Détection des véhicules dans un corpus de vidéo fournit.
2. Calcul des descripteurs de texture, de couleur et sémantique de chaque véhicule détecté
3. Écrire un fichier contenant pour chaque image de véhicule détecté son nom, la vidéo dans laquelle elle apparaît et le résultat obtenu par la méthode FACT.

Ré-identification de véhicule

1. Calcul des descripteurs de texture, de couleur et sémantique du véhicule recherché.
2. Calcul du descripteur de l'image par la méthode FACT.
3. Comparaison par distance euclidienne de ce descripteur aux descripteurs préalablement calculés lors de la détection. Trier les résultats par ordre croissant.

Interface utilisateur

1. Charger l'image du véhicule rechercher.
2. Lancer la ré-identification
3. Afficher une liste triée des véhicules ré-identifié, de leur vidéo et de leur score de ressemblance.

3 Vue générale du système

Pour la réalisation du logiciel de ré-identification les fonctionnalités peuvent être découpées en une liste de tâches à faire en suivant un ordre pertinent de priorité comme détaillé par la suite dans le planning prévisionnel.

3.1 Découpage en tâches

La liste des tâches à réaliser :

1. Spécialisation du réseau de neurone (Yolo3 ou Single shot multibox detector : SSD) pour détecter les véhicules avec un système de boîte englobante.
2. Découpage des vidéo du corpus en image.
3. Utilisation du réseau de neurone pour détecter les véhicules dans un ensemble d'images et récupération des boîtes englobantes

4. Création d'un fichier listant les images (nommées avec pertinence) de véhicules et leur vidéo d'appartenance.
5. Méthode d'apprentissage non supervisée hiérarchique par k-moyenne.
6. Entraînement/Spécialisation du réseau de neurone convolutif GoogleNet avec les bases d'apprentissages en libre accès (Veri/CompsCar).
7. Méthode FACT.
 - (a) Calcul des descripteurs de texture par la méthode SIFT sur les pixels d'intérêts d'une image.
 - (b) Calcul des descripteurs de couleur par la méthode Color Name (CN).
 - (c) Réalisation des Bag-of-Word par k-moyenne, $k=10000$ pour SIFT et $k=250$ pour CN.
 - (d) Calcul du descripteur (texture ou couleur) représentant l'image par réalisation de l'histogramme du Bag-Of-Word associé.
 - (e) Récupération du descripteur sémantique de l'image en sortie de la dernière couche du réseau convolutif googleNet.
 - (f) Calcul du descripteur de l'image par somme pondérée des descripteurs de texture, de couleur et sémantique de l'image.
8. Calcul des descripteurs par FACT de l'ensemble des images obtenues lors de la détection et les concaténer dans le fichier.
9. Calcul de FACT pour le véhicule recherché.
10. Lire le fichier des véhicules détectés et récupérer les informations nécessaires.
11. Comparer par distance euclidienne le descripteur de l'image recherchée aux descripteurs FACT de tous les véhicules détectés.
12. Trier les véhicules par ordre croissant de ressemblance avec le véhicule recherché.
13. Réalisation d'une interface
 - (a) Charger l'image d'un véhicule recherché (redimensionner si nécessaire)
 - (b) Afficher la liste triée par pertinence des véhicules ré-identifiés avec le nom de leur vidéo de provenance et le score calculée (c'est à dire la mesure de ressemblance).

3.2 Découpage en modules

Les différentes tâches énumérées dans la section précédente vont être réparties dans des modules comme décrit ci-après et dans le schéma 1 :

1. Module **Apprentissage** : regroupe les différentes méthodes d'apprentissage à mettre en oeuvre
2. Module **FACT** : module qui permet de calculer le score d'une image selon trois descripteurs.
3. Module **Détection de véhicules - Pré traitement** : pour détecter l'ensemble des véhicules d'un corpus de vidéo.
4. Module **IHM** : interface utilisateur pour ré-identifier une véhicule dans un ensemble de vidéo.
 - (a) Sous-Module **Ré-identification** : comparaison du descripteur FACT de l'image recherchée avec ceux de l'ensemble des images détectées pour trier les véhicules par pertinence de reconnaissance.

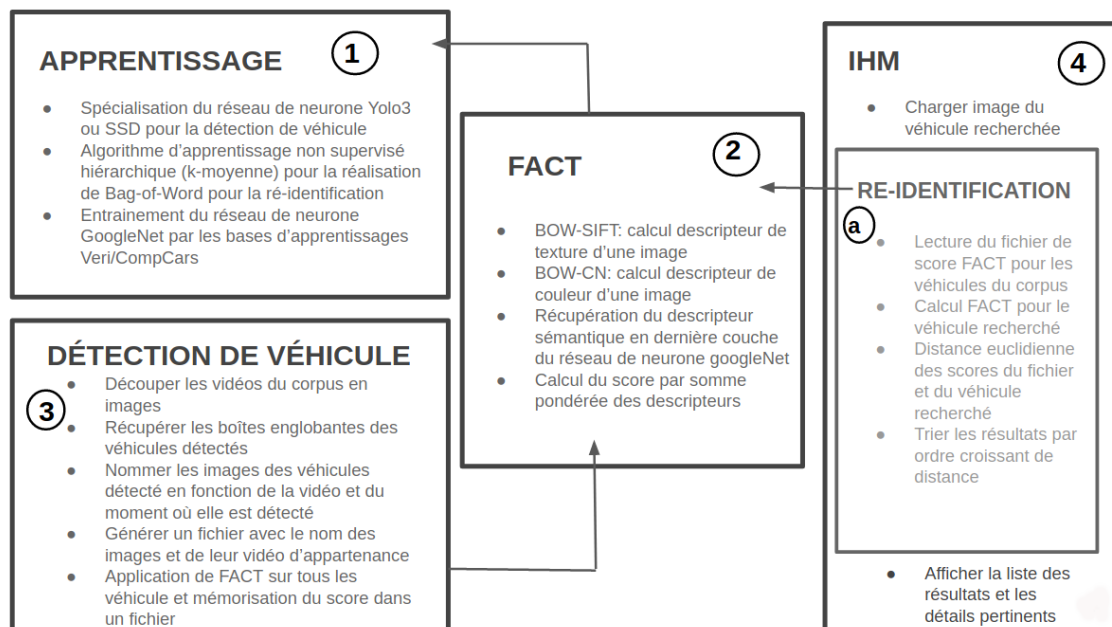


FIGURE 1 – Découpage en module des tâches

4 Fonctionnalités et tests par module

4.1 Ré-identification

Pour le programme dans son ensemble les tests consisteront à regarder si le véhicule sélectionné en entrée fait bien partit des 5 premiers résultats. Si c'est le cas alors la ré-identification est un succès sinon elle a échouée.

4.2 Apprentissage

Une fois les modèles entraînés il sera possible de programmer un script pour la partie BOW-CN en pré-traitant certaines images, ce script nous permettra d'avoir une idée de la précision de BOW-CN.

Les tests du réseau de neurone convolutif consisteront à prendre des images de véhicules et à les passer au réseau, si le véhicule détecté est ressemblant alors le test est un succès. Cette ressemblance se traduit par : "Avec un rapide visionnage est-on capable de voir des différences majeures ?"

4.3 Détection

Les tests concernant le module de détection seront effectués sur des images que nous aurons choisies. En entrée du module se trouve l'image et en sortie une liste d'images correspondant aux boîtes englobantes contenant les véhicules. Ces tests doivent être vérifiés manuellement et doivent comporter des images variées pour prendre en compte les changements de luminosité ou de résolution.

4.4 IHM

L'IHM devra afficher dans le bon ordre les résultats, c'est à dire que les plus probables doivent être les premiers. Nous utiliserons de faux résultats pour vérifier son bon fonctionnement. On entend par faux résultats une liste de valeurs factices qui devront être affichées par ordre croissant pour passer le test.

5 Planning prévisionnel et analyse des risques

5.1 Planning

Le planning prévisionnel représente, en ordonnée, les différentes tâches à effectuer et en abscisse le numéro des différentes semaines jusqu'à échéance. Voir figure 5.1

C = Clémence Courdy-Bahsoun

SD = Sylvain DEKER

SK = Sandeep Kottath

N = Nahor Moussa

R = Ryma Yazı

	Janvier				Fevrier			
	1	2	3	4	5	6	7	8
1. Détection de véhicules	SK	SK						
2. Découpage des video en images			SK					
3. Récupération des boites englobant associés au images				SK				
4. Association/Indexation Véhicules-Video (Fichier)					SD			
5. K-Moyen hiérarchique des descripteurs						C	C	
6. Entraînement de GoogleNET	N	N	N					
7.a Calcul des descripteurs Sift (Texture)	C	C						
7.b Calcul des descripteurs CN (Couleur)	SD	SD						
7.c Réalisation d'un Bag-Of-Word pour Sift et CN			C+SD					
7.d Calcul de deux descripteurs (texture et couleurs) à partir de 7.a et 7.b avec les Bow en 7.c				C+SD				
7.e Récupération du descripteur de GoogleNET				N				
7.f Calcul du score FACT à partir des 3 descripteurs 7.d et 7.e					C			
8. Calcul du score FACT sur l'ensemble des véhicules contenues dans les boites englobante en 3.						SD		
9. Calcul de score FACT sur l'image du véhicule recherché						SD		
10. Associer le score FACT aux véhicules indexés dans le fichier en 4.						N		
11. Comparer le score de l'image du véhicule de référence aux autres (distance SSD)							SD	
12. Trier les véhicules par ordre croissant							SD	
13.a Charger l'image du véhicule recherché							SD	
13.b Affichage des véhicules réidentifiés avec le nom de leur video et la score de distance calculée							N	

FIGURE 2 – Planning prévisionnel

5.2 Risques

Risques :

Un certain nombres de risques ont été répertoriés et classés à l’aide de la matrice de risque 5.2.

1. Facteur humain : mésentente et problème de communication au sein de l’équipe. Disparition d’un membre.
2. Mauvaise évaluation du temps alloué aux tâches.
3. Répartition des tâches.
4. Problèmes d’assemblage des différents modules entre eux.
5. Relation avec le client.

GRAVITE \ PROBABILITE	Mineure	Significative	Grave	Très grave
Fréquent		2, 5		1
Peu fréquent	3			
Rare			4	
Très rare				

FIGURE 3 – Matrice des risques

Solution :

1. Anticiper en partant sur des attentes minimales avec le client.
2. Réunion d'équipe régulière pour voir l'avancement et mettre à jour le planning.
3. Réunion d'équipe régulière pour parler des difficultés et des réussites rencontrés afin d'avancer le plus efficacement possible.
4. Suivre les spécifications et les mêmes normes afin de faciliter l'intégration des modules.
5. S'assurer avec des rencontres régulières de la convergence entre le travail réalisé et les attentes du client.

Bibliographie

- [1] Xinchun Liu, Wu Liu, Huadong Ma, Huiyuan Fu. *Large-scale vehicle re-identification in urban surveillance videos*. ICME, 1–6, 2016.
- [2] Joseph Redmon, Ali Farhadi. *YOLOv3 : An Incremental Improvement*. arXiv/CoRR, 2018.
- [3] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. *SSD : Single Shot MultiBox Detector*. ECCV, 21–37, 2016.
- [4] Linjie Yang, Ping Luo, Chen Change Loy, Xiaoou Tang. *A Large-Scale Car Dataset for Fine-Grained Categorization and Verification*. arXiv/CoRR, 2015.
- [5] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich. *Going Deeper with Convolutions*.
<https://arxiv.org/abs/1409.4842>