



## Partie #3

La troisième partie de votre projet consiste à utiliser un détecteur d'objets basé sur les réseaux convolutionnels afin de détecter les véhicules et les piétons présents dans la scène.

Ceci veut dire que le fichier Ground-truth ne sera plus du tout utilisé afin de réaliser le suivi des acteurs de la scène. Ce fichier sera seulement utilisé afin d'évaluer la qualité des résultats obtenus.

Ajouter donc le détecteur d'objets conçu à la partie 1. La sortie de votre tracker devrait être les images de la séquence montrant les boîtes de détection avec les couleurs montrant les associations temporelles. De plus votre tracker devrait produire un fichier texte de détection ayant un format similaire au fichier ground-truth, c'est-à-dire une détection par ligne avec:

No\_de\_l'image ID\_objet X Y Largeur Hauteur Classe

Le ID\_objet est simplement la valeur RGB de la couleur de la boîte. Le numéro de classe pour les piétons est 1 et celui des voitures est 3.

Vous devez ensuite calculer le MOTA de votre solution. Ceci se fait en comparant votre fichier de détection avec le fichier ground-truth en procédant ainsi:

Pour chaque objet détecté, vérifier si cet objet a un objet correspondant dans le fichier ground-truth ( $\text{IoU} > 0.4$ ). Si oui, noter le ID de cet objet dans le GT (si il y en a plusieurs, considérer l'objet avec le plus grand IoU). Sinon compter cette détection comme un faux positif (FP).

Pour chaque objet du ground-truth, vérifier si cet objet a un objet correspondant dans le fichier de détection ( $\text{IoU} > 0.4$ ). Sinon compter cet objet comme un faux négatif.

Finalement, pour chaque groupe d'objet ayant la même étiquette dans vos détections, compter le nombre d'étiquettes différentes provenant du ground-truth. Ce nombre-1 donne le nombre de changement d'identités.

A partir de ces valeurs, il est possible de calculer la valeur MOTA.