Rapport de Projet

Détection d'Objets en Temps Réel

Théophile Lahaussois

28 mai 2025

Objectif

L'objectif de ce travail est la conception et l'implémentation d'un système de détection d'objets en temps réel à partir d'un flux vidéo à l'aide de MediaPipe.

Algorithmes et Modèles

MediaPipe ObjectDetector

MediaPipe Tasks est un framework open source de Google pour construire des pipelines de traitement multimédia en temps réel. Pour la vision par ordinateur, il propose l'API `ObjectDetector` qui :

- Charge un modèle TensorFlow Lite optimisé (TFLite).
- Gère le cycle de vie du modèle (chargement, inférence, libération).
- Réalise le prétraitement (normalisation, redimensionnement).
- Optimise la performance sur CPU.

EfficientDet-Lite0

EfficientDet est une architecture de détection d'objets reposant sur **EfficientNet-Lite0**, un CNN offrant un bon compromis précision/complexité.
Le modèle utilisé a été récupéré depuis le dépôt officiel MediaPipe : `https://storage.googleapis.com/mediapipe-assets/efficientdet_lite0_int8.tflite`

Prétraitement des Données

Avant chaque inférence, les frames OpenCV (BGR) sont transformées :

1. Conversion BGR → RGB via `cv2.cvtColor`, car le modèle attend un format sRGB.
2. Redimensionnement à 320×320 et normalisation des pixels entre 0 et 1.

Conversion BGR → RGB

```
rgb = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)
mp_image = mp.Image(image_format=mp.ImageFormat.SRGB,
    data=rgb)
```

Paramètres de Détection

— `max_results=10` : limite le nombre de boîtes pour simplifier le post-traitement.
— `score_threshold=0.3` : seuil minimal de confiance pour détecter des objets partiellement masqués.
— `category_allowlist` : restriction aux classes « person, bottle, cell phone, toothbrush, cup ».

Pipeline de Traitement

1. Initialisation du détecteur et chargement du modèle TFLite.
2. Capture vidéo continue avec OpenCV.
3. Pour chaque frame :
   — Prétraitement (BGR → RGB, normalisation).
   — Inférence : `detector.detect_for_video(mp_image, timestamp)`.
   — Dessin des boîtes et calcul du FPS.

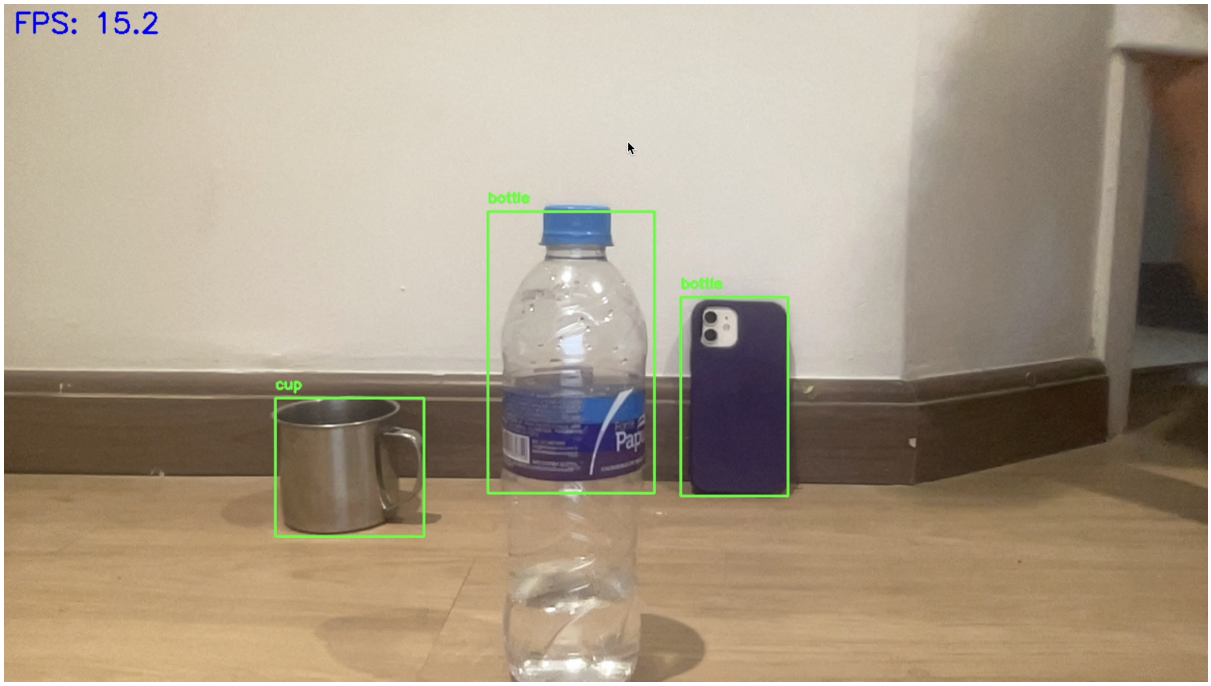
Configuration du détecteur

```
options = vision.ObjectDetectorOptions(
base_options=mp.tasks.BaseOptions(model_asset_path="
    models/efficientdet_lite0.tflite"),
running_mode=vision.RunningMode.VIDEO,
max_results=10,
score_threshold=0.3,
category_allowlist=["person","bottle","cell␣phone","
    toothbrush","cup"]
)
```

Conclusion

Ce projet illustre l'intégration d'un modèle TFLite dans un pipeline MediaPipe pour réaliser une détection d'objets du quotidien, en temps réel.

Les choix de prétraitement et de paramètres sont fait de tel sorte à avoir un bon compromis entre performance et précision.


FPS: 15.2
bottle
bottle
cup