

Projet PI² : Modèle de Deep Learning et Knowledge Graph pour la Réidentification d'assets graphiques dans des vidéos

Le projet proposé s'intéresse à la réidentification d'assets graphiques, connu sous le terme de **ReId**. La **ReId**, pour des personnes par exemple, est une tâche de vision par ordinateur dont l'objectif est de faire correspondre l'identité d'une personne à plusieurs instances tirées d'une vidéo ou d'une séquence d'images. Des techniques de Deep Learning existent mais fonctionnent sur des vidéos de qualité et lorsque les personnes sont aisément identifiables [1].

L'enjeu de ce projet est de pouvoir détecter et de suivre une personne, en utilisant des caractéristiques telles que l'apparence (sous forme de descripteurs appelés « *feature vectors* »), la forme ou la position du corps, les vêtements, l'âge ou encore le groupe ethnique pour faire correspondre son identité dans différentes images capturées par une, voire plusieurs caméras.

L'objectif est donc de comparer les méthodes de Deep Learning classiques à une méthode combinant des techniques de vision par ordinateur et de théorie des graphes (Knowledge Graph, Graph Embedding, Graph Mining). Les étudiants qui choisiront ce projet auront l'opportunité de prendre part à toutes les étapes constitutives d'un projet de recherche scientifique allant de la réalisation d'un état de l'art à la mise en œuvre d'expériences puis la validation des résultats sous l'encadrement d'un doctorant et de professeurs de l'ESILV.

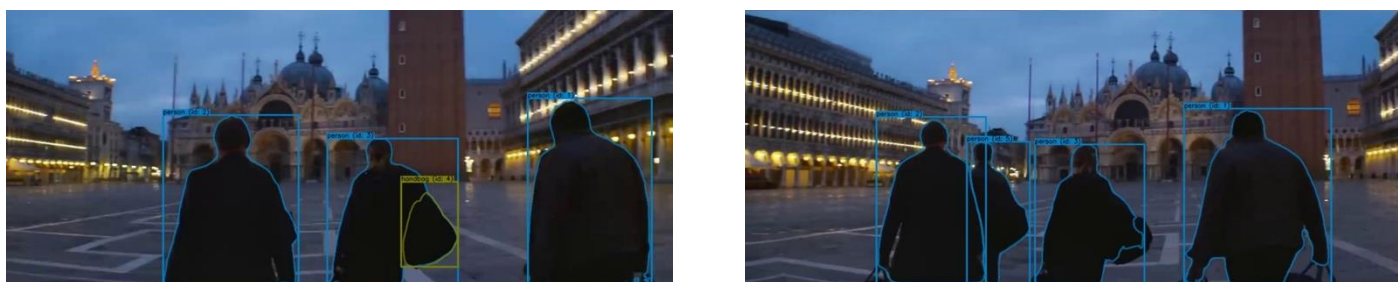


Illustration de la réidentification de personnes¹

Ce projet s'inscrit dans le cadre d'une thèse intitulée « Approche Multicouche pour la Détection Automatique d'Événements Complexes de Multisources Vidéo » financée par l'Agence de l'Innovation de Défense (AID) qui est un organisme nationale rattaché à la Direction Générale de l'Armement et au Ministère des Armées en charge de la mise en œuvre les politiques ministérielles en matière d'innovation et de recherche. Elle oriente les stratégies élaborées par les états-majors sur les thématiques relevant de la défense nationale et coordonne la mise en œuvre des travaux scientifiques et techniques. Elle est également chargée de développer les partenariats et coopérations internationales nécessaires avec des acteurs du domaine de la défense.

Projet proposé par : le laboratoire DVRC de l'ESILV.

Encadrement par : Pierre Lefebvre, Nicolas Travers & Ahmed Azough

¹ Paramount Pictures. (2023, 17 mai). *Mission : Impossible – Dead Reckoning Part One*. YouTube. <https://www.youtube.com/watch?v=avz06PDqDbM>

Vos principales missions (si vous l'acceptez) seront les suivantes :

1. Revue de la littérature sur le thème de la réidentification de personnes
 - a. Recensement des principaux modèles de réidentification de l'état de l'art
 - b. Recensement des principaux jeux de données (datasets) utilisés
 - c. Recensement des principales métriques pour l'évaluation des modèles
 - d. Rédaction d'une synthèse de vos recherches
2. Construction d'un dataset pour la réidentification de personnes [2]
 - a. Choix et regroupement de vidéos libres de droits issues de systèmes de vidéosurveillance (bonus si sur la thématique de la violence)
 - b. Identification des différents formalismes et types de fichiers pour l'étiquetage de données et choix de ces derniers pour les étapes suivantes
 - c. Data cleaning (détection et correction des données incomplètes, incorrectes ou illisibles comme des fichiers dont le format est tronqué, des vidéos floues ou dont le contenu ne correspond pas à la thématique initiale)
 - d. Data labeling (étiquetage) à l'aide d'outils manuels et/ou automatiques tels que Labelstudio, CVAT, V7 Lab, Yolo [3], etc.
3. Expérimentations
 - a. Ecriture de scripts pour le preprocessing des données en python (importation du dataset, standardisation, split, data augmentation, etc.)
 - b. Participation à la conception d'un modèle de réidentification et au protocole de validation puis entraînement de ce dernier
 - c. Réalisation d'un nouveau modèle
 - d. Validation des modèles et comparaison des résultats avec la littérature

Profils recherchés :

Bon niveau d'anglais et non effrayé par la lecture d'articles scientifiques.

Forte appétence pour les domaines du Deep Learning et du Computer Vision.

La connaissance du langage Python et des principales librairies du domaine du Deep Learning (numpy, pandas, TensorFlow, PyTorch) serait un plus non négligeable.

Références :

- [1] Ye, M., Shen, J., Lin, G., Xiang, T., Shao, L., & Hoi, S. C. H. (2022). Deep Learning for Person Re-Identification: A Survey and outlook. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(6), 2872–2893. <https://doi.org/10.1109/tpami.2021.3054775>
- [2] Zheng, L., Bie, Z., Sun, Y., Wang, J., Su, C., Wang, S., & Tian, Q. (2016). MARS: A video benchmark for Large-Scale Person Re-Identification. In *Lecture Notes in Computer Science* (pp. 868–884). https://doi.org/10.1007/978-3-319-46466-4_52
- [3] Bochkovskiy, A., Wang, C., & Liao, H. M. (2020). YOLOV4: Optimal speed and accuracy of object detection. *arXiv (Cornell University)*. <https://arxiv.org/pdf/2004.10934v1>