

Apprendre sans données, une approche d'apprentissage automatique guidée par simulation en 3D pour l'extraction robuste de texte de cartes nationales d'identité

Edouard Bertrand*, Anaïs Druart*, Axel Thevenot*, Christophe Rodrigues*

* Léonard De Vinci Pôle Universitaire, Research Center, 92 916 Paris La Défense, France

Nous nous intéressons à l'extraction de texte à partir de cartes nationales d'identité (CNI) françaises. Les techniques de reconnaissance optique de caractères (OCR) sont performantes, mais les résultats peuvent être mitigés pour les numérisations à partir de smartphone en raison d'une grande variabilité des angles de vue, de l'éclairage, de la qualité de la caméra... A notre connaissance, il n'existe pas de base de donnée publique d'images de CNI. Dans ce contexte, est-il possible d'utiliser l'apprentissage automatique sans données annotées voire sans données du tout ?

En l'absence de données, nous proposons de créer une simulation qui permet de projeter des documents structurés synthétiques dans un environnement 3D. De cette façon, nous pouvons reproduire et contrôler les différentes difficultés qui seraient rencontrées avec l'image réelle d'un document. Cette solution nous permet de générer des exemples synthétiques d'apprentissage pour entraîner nos modèles d'intelligence artificielle de réseaux de neurones. Les étapes clés de la chaîne de traitements de création de données sont illustrées sur la figure 1.

Nous construisons d'abord une CNI vierge à partir d'un échantillon de Wikipédia, sur laquelle nous rajoutons du texte et une image de profil générée à l'aide du modèle StyleGAN. Cette image est alors projetée dans un environnement 3D fabriqué sur le logiciel Blender. Nous faisons ensuite varier la texture de la table et les paramètres de la simulation (tels que la position de la carte, la distance de la caméra, la distance focale...) afin d'obtenir une variété de rendus synthétiques annotés, pouvant être utilisés pour entraîner des modèles d'intelligence artificielle. Les jeux de données que nous avons ainsi générés sont désormais accessibles au public.¹ Néanmoins, en décidant de simuler entièrement les données, nous nous exposons au



FIG. 1 – *Étapes clés de la chaîne de traitement pour la création de données*

problème de la représentativité des exemples générés. Afin de minimiser ce risque, nous proposons l'utilisation d'une procédure d'apprentissage actif guidée par la lisibilité pour régler

1. <https://github.com/ResearchPaper0/Learning-without-real-data>