

Reconnaissance de caractères manuscrits disjoints à partir d'un tableau dans un document scanné

MOTS-CLES :

Analyse de documents, reconnaissance de caractères, Temps réel

DESCRIPTION DU PROJET

Dans le cadre d'une application de dématérialisation de documents. Nous proposons de concevoir une application qui permet d'extraire et de reconnaître des caractères inscrits dans un tableau dont le modèle est prédéfini.

- 1) Prétraitement : analyse de la disposition d'un document : détection de l'emplacement des cellules/ détection et suppression des lignes du tableau.
- 2) Utilisation d'un algorithme pour la reconnaissance robuste de caractères qui assure un taux de reconnaissance acceptable.
- 3) Utilisation du GPU pour l'accélération du calcul intensif pour assurer un traitement rapide des données.
- 4) Génération d'un document PDF éditable.

OUTILS MATERIELS / LOGICIELS SUPPORTS

Environnement de travail : Linux/Windows.

CUDA/OpenCL/C/C++, Python