



AVIS DE PRESENTATION DE THESE EN SOUTENANCE POUR L'OBTENTION DU DIPLOME NATIONAL DE DOCTEUR

Madame Wafa KHLIF

Présentera ses travaux intitulés :

« Détection de texte de scène multilingue basée sur des réseaux de neurones convolutifs »

Spécialité : Informatique et Applications

Le 2 juin 2022 à 15H00

Lieu :

**La Rochelle Université
Pôle Communication, Multimédia et Réseaux
Amphithéâtre Michel Crépeau
44 Av. Albert Einstein
17000 LA ROCHELLE**

Composition du jury :

**M. ABID Mohamed
M. ALIMI Adel
MME BEN AMARA Najoua
M. BURIE Jean-Christophe
Mme EGLIN Véronique
M. GHAMRI -DOUDANE Yacine
Mme NAYEF Nibal
M. RAMEL Jean-Yves**

**Professeur, École Nationale d'Ingénieurs de Sfax
Professeur, École Nationale d'Ingénieurs de Sfax
Professeure, École Nationale d'Ingénieurs de Sousse
Professeur, La Rochelle Université
Professeure, INSA de Lyon
Professeur, la Rochelle Université
Ingénieure de recherche, Myscript
Professeur, Université de Tours**

Résumé :

Cette thèse propose des approches de détection de texte par des techniques d'apprentissage profond pour explorer et récupérer des contenus faiblement structurés dans des images de scène naturelles.

Ces travaux proposent, dans un premier temps, une méthode de détection de texte dans des images de scène naturelle basée sur une analyse multi-niveau des composantes connexes (CC) et l'apprentissage des caractéristiques du texte par un réseau de neurones convolutionnel (CNN), suivie d'un regroupement des zones de texte détectées par une méthode à base de graphes. Les caractéristiques des composantes texte brut/non-texte obtenues à différents niveaux de granularité sont apprises via un CNN.

Une deuxième méthode est présentée dans cette thèse inspirée du système YOLO. Le système réalise la détection du texte et l'identification du script simultanément. Nous considérons la tâche de détection de texte multiscrit comme un problème de détection d'objets, où l'objet est le script du texte. La détection de texte et l'identification des scripts sont réalisées avec une approche holistique en utilisant un réseau neuronal convolutionnel unique.

Les évaluations expérimentales de ces approches sont réalisées sur le jeu de données MLT (Multi-Lingual Text dataset), nous avons contribué à la création de ce nouveau jeu de données. Il est composé d'images de scènes naturelles et synthétiques contenant du texte, tels que des panneaux de circulation et publicitaires, des noms de magasins, d'images extraites des réseaux sociaux. Ce type d'images représente l'un des types d'images les plus fréquemment rencontrés sur Internet, à savoir les images avec du texte incorporé dans les réseaux sociaux.