Nom et Prénom : EL KHAYATI MOHSINE
Date de soutenance : 25/06/2022
Directeur de Thèse : EL KETTANI MOHAMED EL YOUSFI

Sujet de Thèse :

Directing the Attention of Convolutional Neural Networks and the Improvement of Arabic Handwriting Recognition

Abstract :

Due to the strong growth in data availability and the great development of machine power, Deep learning algorithms have emerged in many research fields, outperforming conventional classification algorithms. Convolutional Neural Network (CNN) is one of the most successful deep learning algorithms, especially in computer vision. CNNs have shown a great ability to deal with many difficult issues in the literature. Despite the great success of CNNs, they still suffer from many limitations such as complexity, overfitting, lack of data in some areas, etc. Recent research has focused on the challenges of CNNs, and many solutions have been proposed in the literature. In this thesis, we propose three new approaches to direct the attention of CNN aimed at improving performance and combating limitations. The three approaches were integrated within three models namely EFNet, CGNet, and UPNet. The proposed models have been applied to three essential phases of Arabic Handwriting Recognition (AHR): line segmentation, grapheme segmentation, and character recognition.

Compared to other application areas, CNNs have not been widely investigated for AHR.AHR remains a lively subject of research. It is beneficial in many applications including automatic processing of historical documents, bank checks, postal addresses, forms, etc. Generally, AHR is still in its early stages compared to other scripts and this is mainly due to the difficult nature of Arabic handwriting, the absence of huge databases, and the lack of works in this area. In this thesis, our models have been adapted and combined with other techniques in order to effectively address the key phases of AHR. Obtained results have shown the robustness of our models, their capacity to resolve difficult issues, and to overcome some limitations in CNN.

Keywords: Convolutional neural network, Attention-based neural network, Arabic handwriting recognition, Line segmentation, Grapheme segmentation, Character recognition

Résumé:

En raison de la forte croissance de la disponibilité des données et du grand développement de la puissance des machines, les algorithmes d'apprentissage profond ont émergé dans de nombreux domaines de recherche, surpassant les algorithmes de classification classiques. Le réseau de neurones convolutifs (CNN) est l'un des algorithmes d'apprentissage profond les plus performants, en particulier en vision par ordinateur. Les CNN ont montré une grande capacité à traiter de nombreux problèmes difficiles dans la littérature. Malgré le grand succès des CNN, ils souffrent encore de nombreuses limitations telles que la complexité, le surajustement, le manque de données dans certains domaines, etc. Les recherches récentes se sont concentrées sur les enjeux des CNN, et de nombreuses solutions ont été proposées dans la littérature. Dans cette thèse, nous proposons trois nouvelles approches pour diriger l'attention des CNN visant à améliorer les performances et à surmonter leurs limitations. Les trois approches ont été intégrées dans trois modèles, à savoir EFNet, CGNet et UPNet. Les modèles proposés ont été appliqués à trois phases essentielles de la Reconnaissance de l'écriture Manuscrite Arabe (RMA) : la segmentation des lignes, la segmentation des graphèmes et la reconnaissance des caractères.

Par rapport à d'autres domaines d'application, les CNN n'ont pas été largement utilisés pour la RMA. La RMA est un sujet de recherche vivant et bénéfique dans de nombreuses applications, notamment le traitement automatique de documents historiques, de chèques bancaires, d'adresses postales, de formulaires, etc. Généralement, la RMA en est encore à ses débuts par rapport à d'autres scripts et cela est principalement dû à la nature difficile de l'écriture arabe, à

l'absence d'énormes bases de données et au manque de travaux dans ce domaine. Dans cette thèse, nos modèles ont été adaptés et combinés avec d'autres techniques afin d'aborder efficacement les phases clés du système de RMA. Les résultats obtenus ont montré la robustesse de nos modèles, leur capacité à résoudre des problèmes difficiles en RMA et à surmonter certaines limitations en CNN.

Mot Clés : Réseau neuronal convolutif, Réseau neuronal basé sur l'attention, Reconnaissance du manuscrit arabe, Segmentation de lignes, Segmentation de graphèmes, Reconnaissance de caractères