وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم البرمجيات

تطوير برمجيات التعرف على النصوص المكتوبة بخط اليد والقابلة لإعادة الاستخدام بالاعتماد على TensorFlow

رسالة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة ماجستير علوم في
البرمجيات

من قبل
احمد عبدالرحمن ادريس القدو

بإشراف
أ.د. دجان بشير طه

١٤٤٤ هـ ٢٠٢٢ م

المستخلص

تُعد عملية التعرف على خط اليد واحدة من أهم التحديات التي يواجها الباحثون للحصول على نتائج عالية الدقة والسبب وراء هذه التحديات هو الحاجة لها في عدة مجالات تستخدم لمعالجة المستندات الرقمية، مثل التحقق من الملفات وإدخال البيانات والمكتبات الرقمية وقراءة العناوين البريدية واستخراج البيانات من الشيكات والأوراق الضريبية وغيرها الكثير. إنَّ وظائف عمليات التعرف على خط اليد التقليدية هو العمل على تمييز كل حرف من حروف الكلمة بشكل منفصل.

تم اقتراح في هذه الرسالة خوارزمية تقوم باستخدام طريقة هجينة تدمج مزايا كل من الشبكات العصبية التلافيفية (CNN) والشبكات العصبية المتكررة (RNN). إذ انها تقوم بقراءة ادخالات متغيرة الاطوال وبصورة تسلسلية وهذا ما يجعلها مميزة كونها تتعامل مع ابعاد صورة مختلفة مثل الكلمة ولا تتطلب تحديد او إشارة لهذه الحروف التي في داخل الصورة في مرحلة التدريب ليتمكن النموذج من تعلمها، انما تكون لها القدرة على معالجة هذه الصورة بشكل مباشر. كل هذه الميزات الموجودة في CRNN تجعلها من أفضل الطرائق لقراءة وتمييز البيانات المتسلسلة بدون الحاجه الى عمليات التجزئة للحروف المتسلسلة انما ستتولى الخوارزمية عمليات استخلاص الميزات وتمييز الحروف المتسلسلة. ومن المعروف ان من اهم الأمور التي يتوجب توفرها لبناء نموذج قوي وعالي الدقة هو توفير بيانات كافية لتدريبه ليتسنى له تمييز مختلف الأنماط والاشكال والتنبؤ بها بصورة دقيقة. ولهذا تم استخدام IAM dataset والمتكونة من 657 كاتب جميعهم ساهموا في بناء هذه البيانات والتي تُعد من أضخم قواعد البيانات المتعلقة بخط اليد. هناك ايضاً معالجات مسبقة طبقت اثناء هذه العمليات ومن أهم المعالجات المسبقة التي تم التركيز عليها هي ان الصورة المدخلة (الكلمة) التي سيتم تمييزها يجب ان تطابق الصور التي استخدمت لتدريب النموذج من حيث عمليات المعالجة هذه. وبالاستعانة بهذه الطريقة الكفوءة والبيانات الضخمة تم بناء نظام برمجي يقوم بقراءة الصورة المتكونة من مجموعة من الاسطر (Lines). حيث يقوم النظام البرمجي باكتشاف كل سطر على حدى وتحديده وتجزئته وإدخاله للنموذج المدرب كي يتم تمييز كل الكلمات في هذه الاسطر وبصورة متسلسلة.

تم تدريب ثمان نماذج مختلفة وعلى عدة مراحل وبأكثر من مُحسن ليتم الاخُذ بنظر الاعتبار انسب النماذج التي ستكون جاهزة وقادرة على التنبؤ بمختلف الأنماط. أظهرت النتائج أن معدل الخطأ في الحرف (Character Error Rate) وصلت الى 4.53 ومعدل الخطأ في الكلمات (Word Error Rate) وصلت الى 8.4، وتم مقارنة هذه النسبة مع العديد من المشاريع السابقة فحقق النموذج المقترح دقة اعلى بلغت 95.47٪ في عمليات التمييز والتعرف على خط اليد.

Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Software

Developing Reusable Handwritten Text Recognition Software Based on TensorFlow

**A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of Master
in Software**

By
Ahmed Abdulrahman Idris Alkaddo

Supervised by
Prof.Dr. Dujan Basheer Taha

2022 A.D. **1444 A.H.**

Abstract

The process of handwriting recognition is one of the most important challenges that researchers face to obtain high-accuracy results. The purpose behind these challenges is that it is needed in several areas used for processing digital documents, such as checking files, data entry, digital libraries, extracting data from cheques and tax papers, etc. The traditional handwriting recognition process works to recognize letters in every single word, but the recognition of the whole word process is much better compared with the recognition of letters separately. In addition, connecting words is one of the most difficult situations a handwriting recognition system may encounter. In this paper, we propose an algorithm that uses a hybrid (CNN & RNN). It will sequentially read all letterings and this is what makes it distinctive because it deals with different image dimensions (such as the word) and does not require identification of these letters that are inside the image at the training stage, because it can process this photo directly. All of these features in CRNN make it one of the best ways to read and distinguish sequential data because we will not have to hash the sequential characters, but the algorithm will do the feature extraction and distinguish the sequential characters. It is known that one of the significant portions needed to build a robust and high-accuracy model is to provide sufficient data to train this model so that it can distinguish and predict different patterns accurately. We used the IAM dataset, which is one of the largest databases related to handwriting. There is also preprocessing applied during these processes, but one of the most important preprocessing that was focused on is that where applied on the input image (the word) because it should match the images that were used to train our model. Using this efficient method and big data, a program was built to read the image consisting of a set of lines. Where the program detects each line separately, identifies it, breaks it into words, and enters these words into the trained model to distinguish all the words in these lines in a sequential manner. Different models have been trained on several stages with different data and with more than one optimizer to take into account the most appropriate models that will be ready and able to predict various patterns. The results showed that the error rate of the trained model for the metric Character Error Rate (CER) is 4.53 and Word Error Rate (WER) is 8.4. The proposed model achieved a higher accuracy of 95.47 in the processes of discrimination and handwriting recognition.