

University of Mosul
College of Engineering



A Holistic Handwriting Recognition System Based on Convolutional Neural Networks

Amina Mohammed Abdu Alhameed Hussein

Master Thesis of Science in Computer Engineering
/ Computer Engineering

Supervised by

Asst. Prof. Dr. Mayada Faris Ghanim

Abstract

Handwriting recognition is a significant application in the field of pattern recognition, and it has a wide range of usages where the goal is the recognition of digits, alphabets, words, and other characters. The clear benefit of deep learning architectures like convolutional neural networks (CNN) is the automatic feature extraction from the data rather than hand-engineered features. In this work, deep learning based on a CNN was employed to perform recognition of computer science vocabulary. The model works as an offline electronic dictionary by recognizing the word that is written and presenting a simple explanation of it. The model is constructed and validated using our previously produced handwritten Computer Science Vocabulary Dataset (HCSVD). The dataset contains the handwritten vocabulary chosen from a dictionary of computer science and engineering and it includes 27682 images for 500 different classes, 120 different students in the College of Engineering participated in the data gathering. They were asked to fill out a certain form printed on A4-sized paper with the terms from a predetermined vocabulary set. Each word has been written 55-60 times. After building the dataset the preprocessing was done. The preprocessing stage has many phases that contribute to enhancing the images and making them clearer which include converting the image to grayscale, filtering with stretched contrast, and resizing to the standard size. The proposed CNN model has been evaluated and trained using three different optimizers: Adam, SGD with momentum, and RMSprop. The results showed that SGD with momentum gave the best results in both accuracy and time consumed. The highest accuracy of 0.98 is obtained with a learning rate of 0.001, batch size of 32, and data splitting of 80: 20. The proposed model also showed good performance

in terms of accuracy and time compared to the Alexnet model which is a pioneering CNN that is mostly employed for tasks involving image recognition and classification.



جامعة الموصل
كلية الهندسة

نظام كلي للتعرف على خط اليد على أساس الشبكات العصبية اللافوفية

آمنة محمد عبد الحميد حسين

رسالة ماجستير علوم في هندسة الحاسوب / هندسة الحاسوب

بإشراف

أ.م.د. ميادة فارس غانم

الخلاصة

يعد التعرف على الكتابة اليدوية تطبيقًا مهمًا في مجال التعرف على الأنماط، وله نطاق واسع من الاستخدامات حيث يكون الهدف هو التعرف على الأرقام والحروف الهجائية والكلمات والأحرف الأخرى. تتمثل الفائدة الواضحة لبنيات التعلم العميق مثل الشبكات العصبية التلافيفية (CNN) في استخراج الميزات تلقائيًا من البيانات بدلاً من الميزات المصممة يدويًا. في هذا العمل، تم استخدام التعلم العميق المعتمد على شبكة CNN للتعرف على مفردات علوم الكمبيوتر. يعمل النموذج بمثابة قاموس إلكتروني غير متصل بالإنترنت من خلال التعرف على الكلمة المكتوبة وتقديم شرح بسيط لها. تم إنشاء النموذج والتحقق من صحته باستخدام مجموعة بيانات مفردات علوم الكمبيوتر المكتوبة بخط اليد (HCSVD) والتي تم إنتاجها مسبقًا. تحتوي مجموعة البيانات على المفردات المكتوبة بخط اليد المختارة من قاموس علوم وهندسة الحاسوب وتتضمن 27682 صورة لـ 500 فصل مختلف، شارك 120 طالبًا مختلفًا في كلية الهندسة في جمع البيانات. وقد طُلب منهم ملء نموذج معين مطبوع على ورق بحجم A4 يتضمن مصطلحات من مجموعة مفردات محددة مسبقًا. تمت كتابة كل كلمة 55-60 مرة. بعد بناء مجموعة البيانات تم إجراء المعالجة المسبقة. تشتمل مرحلة المعالجة المسبقة على العديد من المراحل التي تساهم في تحسين الصور وجعلها أكثر وضوحًا والتي تشمل تحويل الصورة إلى التدرج الرمادي، التصنيفية بتباين ممتد وتغيير الحجم إلى الحجم القياسي. تم تقييم نموذج CNN المقترح وتدريبه باستخدام ثلاثة أدوات تحسين مختلفة: Adam و SGD مع الزخم و

RMSprop أظهرت النتائج أن SGD مع الزخم أعطى أفضل النتائج من حيث الدقة والوقت المستهلك. تم الحصول على أعلى دقة 0.98 بمعدل تعلم 0.001 وحجم دفعة 32 وتقسيم البيانات 80:20. كما أظهر النموذج المقترح أداءً جيدًا من حيث الدقة والوقت مقارنة بنموذج AlexNet وهو نموذج CNN الرائد الذي يستخدم في الغالب للمهام التي تتضمن التعرف على الصور وتصنيفها.