

ABSTRACT

EXPLORATION OF PRE-PROCESSING, CONVOLUTIONAL FILTERS, SVM, LSTM, AND KAN FOR ARABIC HANDWRITTEN CHARACTER RECOGNITION

By

Rahmania

21/477740/PA/20698

Arabic is one of the world's largest languages, with approximately 300 million native speakers and 60 million second-language speakers. Arabic-based education is widely used, making the digitalization of Arabic language education increasingly necessary to support the learning process. One approach to this digitalization is handwritten Arabic character recognition using computer technology. Although Convolutional Neural Network (CNN) methods have been widely applied in this field, deeper exploration is still needed.

This study explores the CNN method for handwritten Arabic character recognition, focusing on variations in pre-processing techniques such as closing, convolutional filters, classifiers (SVM and LSTM), and replacing the fully connected layer with the Kolmogorov Arnold Network (KAN). The research is conducted using the Hijja dataset, which consists of handwritten Arabic characters written by children, to evaluate the impact of these modifications on system accuracy. The results indicate that the applied modifications did not significantly improve the model's accuracy. The highest accuracy achieved was 79%, both for the model without modifications and the best-performing modified model.

Keywords: Handwritten Character Recognition, Arabic Language, Convolutional Neural Network (CNN), Closing Pre-processing, SVM, LSTM, Kolmogorov Arnold Network (KAN), Convolutional Filter

INTISARI

EKSPLORASI PRA-PEMROSESAN, FILTER KONVOLUSI, SVM, LSTM, DAN KAN UNTUK REKOGNISI KARAKTER TULISAN TANGAN HURUF ARAB

Oleh

Rahmania

21/477740/PA/20698

Bahasa Arab merupakan salah satu bahasa terbesar di dunia dengan sekitar 300 juta penutur asli dan 60 juta penutur sebagai bahasa kedua. Pendidikan berbasis bahasa Arab digunakan secara luas, sehingga digitalisasi pendidikan berbahasa Arab semakin diperlukan untuk mendukung proses pembelajaran. Salah satu pendekatan dalam digitalisasi ini adalah rekognisi tulisan tangan huruf Arab menggunakan teknologi komputer. Meskipun metode *Convolutional Neural Network* (CNN) telah banyak digunakan dalam rekognisi ini, tetapi eksplorasi lebih mendalam masih perlu dilakukan.

Penelitian ini mengeksplorasi metode CNN untuk rekognisi karakter tulisan tangan huruf Arab, dengan fokus pada variasi pra-pemrosesan seperti *closing*, filter konvolusi, klasifikator (SVM dan LSTM), serta penggantian *fully-connected layer* dengan *Kolmogorov Arnold Network* (KAN). Penelitian dilakukan menggunakan *dataset* Hijja, yang berisi tulisan tangan huruf Arab dari anak-anak, untuk mengevaluasi pengaruh berbagai modifikasi terhadap akurasi sistem. Hasil penelitian menunjukkan bahwa modifikasi yang diterapkan tidak memberikan peningkatan signifikan terhadap akurasi model. Akurasi tertinggi yang diperoleh adalah 79%, baik pada model tanpa modifikasi maupun pada model dengan modifikasi terbaik.

Kata Kunci: Rekognisi Karakter Tulisan Tangan, Bahasa Arab, Convolutional Neural Network (CNN), Pra-pemrosesan *Closing*, SVM, LSTM, Kolmogorov Arnold Network (KAN), Filter Konvolusi