

INTISARI

PENGENALAN KATA BAHASA INDONESIA PADA CITRA DENGAN FONT TULISAN TANGAN BERBASIS TRANSFORMER

Oleh

Dianita Alfi R

18/427490/PA/18450

Perkembangan teknologi yang semakin maju mampu membuat proses dokumen digitalisasi semakin cepat. Terutama perihal digitalisasi tulisan tangan yang semakin dibutuhkan. Pengenalan teks dari citra sudah banyak dilakukan. Banyak arsitektur baru yang mampu untuk mengenali teks terutama citra tulisan tangan. Salah satu arsitektur yang dapat melakukannya yaitu transformer. pengenalan teks dalam bahasa Indonesia dengan citra tulisan tangan yang sudah dilakukan sebelumnya masih banyak menggunakan dataset citra dengan tulisan huruf tegak. Sedangkan variasi tulisan tangan juga ada yang ditulis bersambung sehingga kurang bervariasi. penelitian ini berfokus mengimplementasikan dan mengembangkan sistem pada Transformer dengan pengujian dataset yang lebih bervariasi sehingga dapat memperkaya pengetahuan dalam pembelajaran pemodelan.

Dataset yang digunakan berupa citra dengan objek tulisan kata-kata dalam bahasa Indonesia. Dataset ini dimasukkan ke tahapan *pre-processing*. Untuk anotasi citra, akan diubah menjadi token yang berisi label kelas dan titik koordinat *bounding box*. Setelah tahap ini, dataset akan dilatih dengan menggunakan arsitektur transformer. Komponen utama dari arsitektur Transformer adalah *encoder-decoder*. Setelah model dilatih, tahapan selanjutnya yaitu pengujian. Pengujian ini dilakukan dengan mencari nilai *mean Average Precision* (mAP).

Sistem yang dikembangkan dapat secara efektif mengenali dan mengklasifikasikan benda-benda dari data gambar tulisan tangan yang termasuk objek kata bahasa Indonesia. *Hyperparameter* yang paling optimal didapatkan *batch* dan jumlah *epoch* masing-masing 32 dan 40. Hasil evaluasi model yang didapatkan pada data Train menggunakan mAP dan *confusion matrix* sebesar 0,98 dan nilai rata-rata *accuracy* 0,997, *precision* 0,984, *recall* 0,984, dan *f1-score* 0,984. Sedangkan hasil evaluasi performa model pada data Test mendapatkan nilai mAP sebesar 0,95 dan nilai rata-rata *accuracy* 0,990, *precision* 0,957, *recall* 0,950, dan *f1-score* 0,950.

Kata Kunci: *Transformer, Image Recognition, Tulisan Tangan*

ABSTRACT

RECOGNITION OF INDONESIAN WORDS IN THE IMAGE OF HANDWRITING FONT BASED ON TRANSFORMER

By

Dianita Alfi R

18/427490/PA/18450

The development of increasingly advanced technology is able to make the document digitization process faster. Especially regarding the digitization of handwriting which is increasingly needed. Text recognition from images has been done a lot. Many new architectures are able to recognize text, especially handwritten images. One architecture that can do this is a transformer. text recognition in Indonesian with handwritten imagery that has been done previously still uses a lot of image datasets with upright lettering. Meanwhile, handwriting variations are also written in succession so that they are less varied. This research focuses on implementing and developing systems on Transformers with more varied dataset testing so as to enrich knowledge in learning modeling.

The dataset used is in the form of images with written objects in Indonesian. This dataset is entered into the pre-processing stage. For image annotations, it will be converted into a token containing the class label and bounding box coordinates. After this stage, the dataset will be compared using the transformer architecture. The main components of the Transformer architecture are the encoder-decoder. After the completion of the model, the next stage is testing. This test is carried out by looking for the mean Average Precision (mAP) value.

The developed system can effectively recognize and classify objects from handwritten image data which include Indonesian word objects. The most optimal hyperparameters were obtained from the batch and the number of epochs, respectively 32 and 40. The model evaluation results obtained on the Train data using the mAP and confusion matrix were 0.98 and the average accuracy was 0.997, precision was 0.984, recall was 0.984, and f1-score was 0.984. While the results of the model performance evaluation on the Test data obtained a mAP value of 0.95 and an average accuracy value of 0.990, precision 0.957, recall 0.950, and f1-score 0.950.

Keywords: *Transformer, Image Recognition, Handwriting*