

INTISARI

PENGENALAN KARAKTER TULISAN TANGAN MANUSIA DENGAN NEURAL NETWORK DAN KLASIFIKASI DATA

Oleh

IRFAN MAULANA KAMIL

14/365980/PA/16184

Penelitian ini dilakukan untuk mencari tahu pengaruh dari klasifikasi data dalam proses pengenalan data tulisan tangan manusia. Sebelum data gambar tulisan tangan dilakukan proses pengenalan, data tersebut diklasifikasikan menjadi 4 kelas. Data yang telah terklasifikasi kemudian dilakukan proses pengenalan dengan menggunakan metode pengenalan yang terpisah untuk masing-masing kelas. Asumsi awalnya adalah dengan mengurangi jumlah varian data dalam satu proses pengenalan, akan meningkatkan akurasi pengenalan per kelas, sehingga meningkatkan juga akurasi umumnya.

Percobaan ini dilakukan dengan metode pengenalan *backpropagation neural network* dan metode klasifikasi baru yang diusulkan. Metode klasifikasi ini menggunakan perhitungan pixel aktif dalam sebuah gambar tulisan tangan untuk kemudian diklasifikasikan. Model ini dicoba untuk mengenali dataset EMNIST yang berisi data gambar tulisan tangan (0-9, A-Z, dan a-z). Untuk melihat pengaruh dari klasifikasi data, maka model pengenalan tanpa klasifikasi juga dibuat dalam penelitian ini. Kedua model pengenalan yang dibuat dirancang dengan arsitektur dan pengaturan yang sama agar menciptakan perbandingan yang setara.

Hasil akhir penelitian ini membuktikan bahwa klasifikasi data dapat mempengaruhi akurasi pengenalan data. Model pengenalan dengan klasifikasi data menghasilkan akurasi umum tertinggi mencapai **75%** sedangkan model pengenalan tanpa klasifikasi menghasilkan akurasi tertinggi **64%**. Berdasarkan hasil tersebut, model pengenalan tulisan tangan dengan klasifikasi data lebih unggul kurang lebih **10%** dalam segi akurasi. Dalam segi waktu pelatihan, model baru ini tentunya lebih membutuhkan waktu lama dikarenakan lebih banyaknya *neural network* yang digunakan.

Kata Kunci : Pembelajaran mesin, pengenalan tulisan tangan, klasifikasi data.

ABSTRACT

HUMAN HANDWRITTEN CHARACTERS RECOGNITION WITH NEURAL NETWORK AND DATA CLASSIFICATION

By

IRFAN MAULANA KAMIL

14/365980/PA/16184

This research was conducted to find out the influence of data classification in the process of human handwriting data recognition. Before the process of recognition, the data is classified into 4 classes. The classified data then processed by using a separate recognition method for each group. The initial assumption is that by reducing the number of data variants in an recognition process, it will increase the accuracy of recognition per class, thus increasing general accuracy as well.

This experiment was conducted with backpropagation neural network method and proposed new classification method. This classification method uses active pixel calculations in a handwritten characters image to be classified. This model is attempted to recognize EMNIST datasets containing handwritten image data (0-9, A-Z, and a-z). To see the effect of data classification, an unclassified data recognition model was also made in this experiment. Both of the recognition models are designed with the same architecture to create equal comparisons.

The final result of this experiment proves that the data classification can affect the accuracy of the data recognition. The recognition model with data classification yielded the highest general accuracy of 75% while the unclassified recognition model yielded the highest accuracy of 64%. Based on these results, handwriting recognition model with data classification is more or less 10% better in terms of accuracy. In terms of training time, this new model is certainly more time-consuming due to the more neural network used.

Keywords : Machine learning, handwriting recognition, data classification.