

INTISARI
CNN-SVM UNTUK PENGENALAN POLA KARAKTER TULISAN
TANGAN HURUF JEPANG HIRAGANA DAN KATAKANA

Oleh

Nicolaus Euclides Wahyu Nugroho

18/433789/PPA/05604

Telah banyak metode berbasis *deep learning* yang telah diteliti selama bertahun-tahun. *Deep learning* banyak diteliti untuk memecahkan masalah pengenalan pola terutama terutama karakter tulisan tangan atau yang sering disebut *character recognition*. Dari sekian banyak metode tidak banyak yang melakukan penelitian pengenalan karakter tulisan tangan *hiragana* dan *katakana*. Beberapa penelitian telah dilakukan untuk pengenalan karakter tulisan tangan *hiragana* dan *katakana*. Variabel yang menentukan kesulitan dalam pengenalan karakter *hiragana* dan *katakana* adalah banyaknya karakter yang hampir mirip satu dengan yang lainnya. Terdapat pula kemiripan antara beberapa huruf yang merupakan transformasi dari huruf dasar yang ada. Penelitian ini dilakukan karena masih terdapat masih terdapat beberapa kesalahan pembacaan karakter terutama pada huruf *katakana* dengan *tenten* (*dakuten*).

Penelitian ini dilakukan untuk memecahkan masalah tersebut sehingga bisa didapatkan hasil yang lebih akurat dalam pengenalan pola tulisan tangan huruf *hiragana* dan *katakana*. Pada penelitian ini dilakukan berbagai kombinasi dari beberapa metode yang telah dipelajari sehingga dapat menghasilkan algoritma pengenalan karakter *hiragana* dan *katakana* yang cukup akurat. Metode yang digunakan adalah metode *Convolutional Neural Network* (CNN) sebagai pengekstrak ciri dan pengolah fitur dan digabungkan dengan *Support Vector Machine* (SVM) sebagai alat untuk pengenalan dari data yang digunakan. Pada penelitian ini juga dilakukan perbandingan antara pengenalan yang dilakukan oleh CNN sendiri dengan CNN yang digabungkan dengan SVM.

Hasilnya didapatkan akurasi dengan parameter sebagai berikut ukuran gambar 69×69, nilai *patience* 3, *callbacks monitor val_loss*, fungsi optimasi Nadam, nilai *learning rate* 0.001, nilai *epochs* 30, dan kernel SVM rbf. jika menggunakan sistem yang hanya menggunakan jaringan CNN akurasi yang didapat sebesar 87,82%. Hasil yang didapat jika menggunakan gabungan CNN dan SVM sebesar 88,21%.

Kata Kunci: *Character-Recognition, SVM, CNN, Deep-Learning, Japanese Character Recognition*

ABSTRACT
CNN-SVM FOR HANDWRITTEN JAPANESE HIRAGANA AND
KATAKANA CHARACTER PATTERN RECOGNITION

by

Nicolaus Euclides Wahyu Nugroho

18/433789/PPA/05604

There have been many deep learning based methods that have been researched over the years. Deep learning is widely researched to solve pattern recognition problems, especially handwritten characters or what is often called character recognition. Of the many methods, not many have conducted research on handwritten hiragana and katakana character recognition. Several studies have been conducted on the handwritten hiragana and katakana character recognition. The variable that determines the difficulty in recognizing hiragana and katakana characters is the number of characters that are almost similar to one another. There are also similarities between several letters which are transformations of the existing basic letters. This research was conducted because there are still some misreading of characters, especially in katakana letters with tenten (dakuten).

This research was conducted to solve this problem so that more accurate results can be obtained in the recognition of hiragana and katakana handwriting patterns. In this research, various combinations of several methods have been studied so as to produce a fairly accurate hiragana and katakana character recognition algorithm. The method used is the Convolutional Neural Network (CNN) method as a feature extractor and feature processor combined with a Support Vector Machine (SVM) as a tool for recognition of the data used. In this study also conducted a comparison between the introduction carried out by CNN itself and CNN combined with the SVM.

The result obtained accuracy with the following parameters: 69×69 image size, 3 patience value, val_loss monitor callbacks, Nadam optimization function, 0.001 learning rate value, 30 epochs value, and SVM rbf kernel. If using a system that only uses the CNN network, the accuracy is 87.82%. The results obtained when using a combination of CNN and SVM were 88.21%.

Keywords: *Character-Recognition, SVM, CNN, Deep-Learning, Japanese Character Recognition*