

## ABSTRACT

### **Comparison of K-Nearest Neighbor, Naïve Bayes and Support Vector Machine for Character Recognition in Mobile Application**

by

Christina Caraswati Liantara

12/333043/PA/14810

Due to its variations of techniques and applications, optical character recognition is still the most popular research in digital image processing. As the increasing number of mobile devices and it becomes the future operating system, this application is built on mobile based. With this application, user can easily recognized the characters by capturing image using mobile phone camera. Open Computer Vision library is used to build the application as this library is the most popular one.

Besides, machine learning algorithms are used for the recognition process in the application. There are k-nearest neighbor, naïve Bayes classifier and support vector machine algorithms which are used in this research. The performances of these three algorithms are calculated including recognition accuracy, memory usage and time consumed.

From testing process, k-nearest neighbor show the best performance. It has the highest character recognition accuracy at 94% and the least average memory usage and time consumed in the training and prediction process. The result shows that it consumed average memory of 650300.7 KB in training and 650409.1 KB in the prediction and consumed average time of 128.5 microseconds in training and 520.6 microseconds in prediction.

***Keywords: mobile application, character recognition, OpenCV library, machine learning***

## INTISARI

### Perbandingan K-Nearest Neighbor, Naive Bayes dan Support Vector Machine pada Pengenalan Karakter dalam Aplikasi Piranti Bergerak

oleh

Christina Caraswati Liantara

12/333043/PA/14810

Dikarenakan adanya variasi teknik dan aplikasi, pengenalan karakter optis masih menjadi topik penelitian yang populer dalam bidang pengolahan citra digital. Dengan bertambahnya jumlah piranti bergerak yang menjadi sistem operasi masa depan, aplikasi ini dibuat berbasis piranti bergerak tersebut. Dengan adanya aplikasi ini, pengguna dapat dengan mudah mengenali karakter hanya dengan mengambil gambar melalui kamera telepon seluler. *Open Computer Vision Library* digunakan dalam membangun aplikasi ini karena *library* ini merupakan yang terpopuler.

Selain itu, algoritma *machine learning* juga digunakan dalam aplikasi ini pada proses pengenalan karakter. Beberapa algoritma *machine learning* yang digunakan dalam penelitian ini antara lain, *k-nearest neighbor*, *naïve Bayes classifier* dan *support vector machine*. Performa dari ketiga algoritma ini yang terdiri dari akurasi pengenalan karakter, memori yang digunakan dan waktu yang digunakan akan dibandingkan.

Dalam proses pengujian, algoritma *k-nearest neighbor* menunjukkan performa terbaik. Algoritma ini memiliki akurasi pengenalan karakter paling tinggi sebesar 94% dan menggunakan memori dan waktu paling sedikit dalam proses pelatihan dan proses prediksi. Hasil pengujian menunjukkan memori rata-rata yang digunakan sebesar 650300.7 KB pada proses pelatihan dan 650409.1 KB pada proses prediksi dan waktu rata – rata yang digunakan sebesar 128.5 mikrodetik pada proses pelatihan dan 520.6 mikrodetik pada proses prediksi.

**Kata kunci :** *aplikasi piranti bergerak, pengenalan karakter, OpenCV library, machine learning*