

ABSTRAK

Jumlah karakter dalam bahasa Jepang Hiragana yang melimpah, kemiripan bentuk karakter, dan kurangnya familiaritas masyarakat terhadap Hiragana dalam kehidupan sehari-hari membuatnya sulit untuk dipelajari. Masyarakat cenderung lebih terbiasa dengan tulisan romaji (abjad) daripada karakter tertentu, yang menyebabkan kesulitan dalam memahami Hiragana dengan variasi ukuran dan bentuk yang beragam.

Penelitian ini bertujuan untuk mengembangkan sistem pengenalan tulisan tangan bahasa Jepang Hiragana yang efektif dan sistematis menggunakan metode SVM (*Support Vector Machine*) dan SIFT (*Scale Invariant Feature Transform*). Metodologi penelitian mencakup identifikasi masalah, studi literatur, pengumpulan data, *data preprocessing*, perancangan sistem, implementasi, dan evaluasi.

Data yang diperoleh mengalami proses augmentasi dan *image preprocessing* untuk menciptakan variasi dan jumlah data yang lebih besar. Selanjutnya, dilakukan ekstraksi fitur pada data menggunakan metode SIFT sebelum melatih model menggunakan SVM. Hasil penelitian menunjukkan bahwa model SVM-SIFT mencapai akurasi sebesar 0.989130 pada 92 data pengujian dengan Skor CV terbaik 0.8509327303561335, yang lebih unggul dibandingkan model SVM tanpa SIFT dengan akurasi 0.978261, dengan Skor CV terbaik untuk model SVM tanpa SIFT sebesar 0.7746709410609622. Pengujian membuktikan bahwa penggunaan SVM-SIFT efektif untuk klasifikasi tulisan tangan yang bervariasi dalam bentuk dan ukuran.

Kata kunci: Bahasa Jepang, Pengenalan Hiragana, Klasifikasi tulisan, SVM-SIFT

ABSTRACT

The abundance of characters in Japanese Hiragana, the similarity in character shapes, and the lack of familiarity among the general public with Hiragana in daily life make it difficult to learn. People tend to be more accustomed to romaji (alphabet) writing than specific characters, which leads to difficulties in understanding Hiragana with various sizes and shapes.

This research aims to develop an effective and systematic Japanese Hiragana handwritten recognition system using SVM (Support Vector Machine) and SIFT (Scale Invariant Feature Transform) methods. The research methodology includes problem identification, literature review, data collection, data preprocessing, system design, implementation, and evaluation.

The obtained data undergoes augmentation and image preprocessing processes to create a larger variety and quantity of data. Subsequently, feature extraction is performed on the data using the SIFT method before training the model using SVM. The research results indicate that the SVM-SIFT model achieves an accuracy of 0.989130 on 92 test data with the best CV score of 0.8509327303561335, which is superior to the SVM model without SIFT with an accuracy of 0.978261, and the best CV score for the SVM model without SIFT is 0.7746709410609622. Testing proves that the use of SVM-SIFT is effective for classifying handwritten characters with variations in shape and size.

Keywords: *Japanese Language, Hiragana Recognition, Handwriting Classification, SVM-SIFT*