ABSTRAK

Masyarakat Sunda merupakan suku terbesar kedua (15,5 persen) setelah suku Jawa. Namun, tidak banyak penutur yang memahami aksara Sunda. Beragam lambang membuat aksara Sunda sulit dihafal. Perda Nomor Tahun 2003 tentang pemeliharaan, Aksara, Bahasa, dan Sastra Daerah yaitu aksara Sunda. Inovasi komputerisasi dapat dilakukan melalui sistem pengenalan aksara Sunda untuk menujang kelestarian budaya tersebut.

Penelitian ini bertujuan untuk mengembangkan sistem pengenalan tulisan tangan aksara Sunda yang efektif dan sistematis menggunakan metode SVM (*Support Vector Machine*), dengan bantuan ekstraksi fitur SIFT (*Scale Invariant Feature Transform*), dan mengetahui nilai akurasinya. Metodologi penelitian mencakup *bussines understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation*, dan *deployment*. Data yang diperoleh mengalami proses augmentasi data, kemudian dilakukan *preprocessing* (*resize, grayscale, centered image,* dan *histogram equalization*). Selanjutnya, dilakukan ekstraksi fitur pada data menggunakan metode SIFT, sebelum akhirnya melatih model menggunakan SVM.

Hasil penelitian menunjukkan bahwa hasil model SVM-SIFT mencapai akurasi sebesar 85% data uji dengan rasio *splitting* 95:5, $k = 300$, $C = 10$, kernel RBF dan $gamma = 0.231$. Model tersebut lebih unggul dibandingkan model SVM tanpa SIFT dengan peningkatan *precision* berbagai kelas dan akurasi 81% pada data uji.

Kata kunci: Aksara Sunda, Klasifikasi Tulisan, *Support Vector Machine* (SVM), *Scale Invariant Feature Transform* (SIFT)

ABSTRACT

The Sundanese people are the second largest ethnic group (15.5 percent) after the Javanese. However, not many speakers understand Sundanese script. Various symbols make Sundanese script difficult to memorize. Regional Regulation Number 2003 concerning the preservation of Regional Scripts, Languages, and Literature, namely Sundanese script. Computerized innovation can be done through a Sundanese script recognition system to support the preservation of this culture.

This research aims to develop an effective and systematic Sundanese handwriting recognition system using the SVM (Support Vector Machine) method, with the help of SIFT (Scale Invariant Feature Transform) feature extraction, and determine its accuracy value. The research methodology includes business understanding, data understanding, data preparation, modelling, evaluation, and deployment. The obtained data undergoes a data augmentation process, followed by preprocessing (resize, grayscale, centered image, and histogram equalization). Furthermore, feature extraction is performed on the data using the SIFT method, before finally training the model using SVM

The results showed that the SVM-SIFT model achieved an accuracy of 85% test data with a 95:5 splitting ratio, $k = 300$, $C = 10$, *kernel RBF and* $gamma = 0.231$. *This model outperforms the SVM model without SIFT with increase in precision for various classes and 81% accuracy on test data.*

Keywords: Sundanese Script, Text Classification, Support Vector Machine (SVM), Scale Invariant Feature Transform (SIFT)