

Klasifikasi Topeng Pandawa dengan SVM

Andi Sanjaya a, 1, *, Endang Setyati a, 2, Herman Budianto a, 3

a Department of Information Technology, Institute Sains dan Teknologi Terpadu Surabaya,
Ngagel Jaya Tengah 73-77, Surabaya-60284, Indonesia

Email: ¹4ndisanjaya@gmail.com; ² endang@stts.edu*; ³ herman.budianto@gmail.com

Abstract. Classification is an advanced stage of a computer vision science. Because the purpose of a recognition application is to recognize. The way to recognize is by way of classification. Many classification methods exist, but in this study using Support Vector Machine (SVM). SVM was chosen because it can handle data with very large dimensions without reducing data, working with linear or nonlinear data and creating a hyperplane that separates data between classes. In this study using Pandava mask data with five classes. Five classes consist of the Yudhistira, Bima, Arjuna, Nakula and Sadewa classes. The kernel used in this study uses the Radial Basis Function (RBF). The results of trials in this study have an average f-measure of 0.82.

Keywords: Classification, Support Vector Machine, Pandava Mask

Abstrak. Klasifikasi merupakan tahapan tingkat lanjut dari sebuah keilmuan computer vision. Karena tujuan dari sebuah aplikasi rekognisi yaitu mengenali. Cara mengenali yaitu dengan cara klasifikasi. Banyak metode klasifikasi yang ada, namun pada penelitian ini menggunakan Support Vector Machine (SVM). SVM dipilih karena bisa mengatasi data dengan dimensi yang sangat besar tanpa mereduksi data, bekerja dengan data linier atau nonlinier dan membuat sebuah hyperplane yang memisahkan data antar kelas. Pada penelitian ini menggunakan data topeng pandawa dengan lima kelas. Lima kelas terdiri dari kelas yudhistira, bima, arjuna, nakula dan sadewa. Kernel yang digunakan pada penelitian ini menggunakan Radial Basis Function (RBF). Hasil ujicoba pada penelitian mempunyai rata-rata f-measure sebesar 0.82.

Kata Kunci: Klasifikasi, Support Vector Machine, Topeng Pandawa

1. Pendahuluan

Saat ini, ada banyak penelitian yang membahas klasifikasi gambar dengan berbagai objek dan metode. Klasifikasi gambar biasanya digunakan untuk membuat aplikasi pengenalan objek. Klasifikasi gambar digunakan untuk membuat aplikasi pengenalan objek. Misalnya adalah Teks dan karakter pengakuan pada lembaran logam [1], Pembelajaran mendalam untuk pengenalan karakter Jawa tulisan tangan [2], Pengambilan gambar berbasis konten untuk pengenalan buah multi-objek menggunakan k-means dan k-terdekat tetangga [3], Lisensi Pengenalan Plat Otomatis berdasarkan deteksi tepi [4] dan Arab pengenalan tulisan tangan menggunakan optimasi minimal sekuensial [5].

Tujuan dari penelitian ini adalah untuk melakukan klasifikasi gambar topeng pandawa. Metode klasifikasi yang digunakan adalah Support Vector Machine (SVM). Fitur yang digunakan dalam proses klasifikasi merupakan seluruh piksel dari seluruh gambar yang ada. Gambar topeng pandawa diklasifikasikan ke dalam 5 kelas. Kelas tersebut antara lain yudhistira, bima, arjuna, nakula dan sadewa. Penulis berharap penelitian ini bermanfaat sebagai bentuk upaya pelestarian peninggalan sejarah Indonesia dengan memanfaatkan teknologi.

SVM adalah metode klasifikasi yang memanfaatkan jarak margin terjauh dari hyperplane [6] [7]. SVM juga dikenal sebagai metode yang mampu mengolah data dengan dimensi tinggi tanpa mengurangi dimensi tertentu data [8] [9]. SVM sering memberikan model dan solusi yang baik. Hasil model bisa dimanfaatkan lagi dalam proses pengujian. SVM mampu memisahkan distribusi data yang bersifat linear atau non-linear.

Ada beberapa penelitian serupa yang menggunakan metode SVM untuk klasifikasi gambar seperti penelitian yang dilakukan oleh [7], [11] dan [12]. Namun, pada penelitian tersebut menggunakan objek gambar yang berbeda. Dalam penelitian yang dilakukan oleh [7], penulis menggunakan gambar tumor dan hasil deteksi uji coba memperoleh skor presisi 93,3%. Dalam penelitian yang dilakukan oleh [11], penulis menggunakan gambar digit angka tahun peninggalan Kerajaan Majapahit dan hasil dari rata-rata klasifikasi sebesar 77%. Sedangkan pada penelitian [12], penulis menggunakan *data mining* untuk diklasifikasikan sebagai *malware* atau bukan dan menghasilkan nilai rata-rata klasifikasi sebesar 74-83%.

Untuk memudahkan pemahaman tentang proses penelitian ini, maka penulis menggunakan artikel struktur penulisan dengan membagi menjadi tiga bagian, metodologi penelitian, hasil eksperimen, dan yang terakhir bagian

adalah kesimpulannya. Pada bagian metodologi penelitian, penulis menyediakan tiga sub-bagian sebagai berikut: pengumpulan data, metode, dan analisis data. Sub-bagian pengumpulan data menjelaskan bagaimana penulis memperolehnya set data yang digunakan dalam penelitian ini. Sub-bagian metode menjelaskan bagaimana cara kerja SVM. Sementara itu, sub-bagian analisis data membahas bagaimana dataset yang digunakan dapat diimplementasikan oleh metode yang ditentukan. Pada bagian hasil percobaan, ini memaparkan hasil uji coba dengan menerapkan metode SVM.

2. Metode

A. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah gambar topeng pandawa dengan total data sebanyak 500 gambar. Gambar topeng pandawa terdiri dari 5 kelas, masing-masing kelas mempunyai 100 gambar. Untuk contoh gambar bisa dilihat pada gambar 1.



Gambar 1. Contoh gambar pandawa

C. Klasifikasi

Klasifikasi adalah teknik utama dalam memisahkan data dan banyak digunakan di berbagai bidang. Dengan definisi sederhana, klasifikasi bisa dikatakan sebagai proses menganalisis sekumpulan data dan menghasilkan sekumpulan aturan pengelompokan yang dapat digunakan untuk mengklasifikasikan data yang baru. Untuk klasifikasi kami menggunakan pada SVM [12]. SVM memiliki regularisasi yang kuat, yang mengacu pada generalisasi model data baru [12] [13]. SVM biasa digunakan dalam hal *machine learning*. Diproses melalui dua langkah: pelatihan dan pengujian [6].

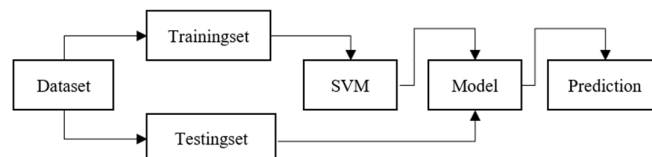
SVM didasarkan pada konsep pengambilan keputusan [7]. Pengambilan keputusan didasarkan pada pemisahan anggota fitur dari kelas yang berbeda. SVM dipilih, karena mampu mengatasi dimensi fitur yang sangat besar dan tanpa mengurangi fitur yang ada. SVM melakukan klasifikasi dengan membuat *hyperplane n-dimensional* [13]. Tujuan SVM adalah untuk menemukan margin *hyperplane* terbesar [10].

SVM adalah sebuah algoritma untuk memaksimalkan fungsi matematika tertentu sehubungan dengan koleksi data yang diberikan. SVM dapat menangani data yang linear dan data non-linear. SVM sederhana biasanya linier dalam membagi fitur ke dalam kedua kelas.

SVM dikelompokkan menjadi dua jenis klasifikasi linier dan non-linear. Klasifikasi SVM linier bermanfaat bagi klasifier nonlinear untuk memetakan pola input ke dalam ruang fitur dimensi yang lebih tinggi. Di sini, kami menggunakan SVM non-linear, yang dibagi menjadi beberapa kelas. Karena itu, SVM membutuhkan kernel. Kernel Gaussian atau yang biasa dikenal dengan Radial Basis Function (RBF).

D. Skema Ujicoba

Dalam penelitian ini, menggunakan *dataset* sebanyak 500 gambar topeng pandawa. *Dataset* tersebut dibagi menjadi dua yaitu sebagai *trainingset* dan *testingset* seperti pada gambar 2.



Gambar. 2. Alur kerja

Pembagian antara *trainingset* dan *testingset* yaitu 50:50. Jadi pemrosesan data yang digunakan adalah 250 data untuk *trainingset* dan 250 untuk *testingset*. Data pada *trainingset* diproses dengan menggunakan SVM. Jadi fitur yang digunakan SVM untuk melakukan pemrosesan *dataset* hanya menggunakan semua nilai perpixel semua gambar yang ada. Setelah *trainingset* diproses melalui SVM maka muncul sebuah model. Model yang dihasilkan inilah yang digunakan untuk melakukan proses pada *testingset*. Sehingga pada tahap prediksi bagus atau tidaknya hasil dikarenakan pada model yang dihasilkan. Apabila menghasilkan model yang baik, maka hasil akan baik.

3. Hasil Percobaan

Pada bagian ini dibahas tentang hasil ujicoba dari skema ujicoba yang telah direncanakan pada bagian sebelumnya. Percobaan berfokus pada kemampuan SVM dalam melakukan klasifikasi gambar menjadi lima kelas yaitu yudhistira, bima, arjuna, nakula dan sadewa. Untuk hasil klasifikasi bisa dilihat pada tabel 1.

Tabel 1. Hasil Klasifikasi

| Kelas | Prediksi | | | | | Total data |
|--------------------------|------------|-----------|-----------|-----------|-----------|------------|
| | yudhistira | bima | arjuna | nakula | sadewa | |
| yudhistira | 44 | 0 | 3 | 1 | 2 | 50 |
| bima | 0 | 49 | 0 | 0 | 1 | 50 |
| arjuna | 4 | 3 | 35 | 1 | 7 | 50 |
| nakula | 0 | 0 | 0 | 50 | 0 | 50 |
| sadewa | 3 | 8 | 11 | 0 | 28 | 50 |
| Total terprediksi | 51 | 60 | 49 | 52 | 38 | |

Dari hasil percobaan pada tabel 1 bisa diketahui bahwa total benar kelas yudhistira sejumlah 44, total benar kelas bima sejumlah 49, total benar kelas arjuna sejumlah 35, total benar kelas nakula sejumlah 50 dan total benar kelas sadewa sejumlah 28. Dari tabel 1 dicari nilai presisi, *recall* dan *f-measure*. Untuk rumus presisi bisa dilihat pada (1), untuk rumus *recall* bisa dilihat pada (2) sedangkan untuk rumus *f-measure* bisa dilihat pada (3).

$$presisi = \frac{total\ benar}{total\ terprediksi} \quad (1)$$

$$recall = \frac{total\ benar}{total\ data} \quad (2)$$

$$f - measure = \frac{2 \times presisi \times recall}{presisi + recall} \quad (3)$$

Presisi merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. *Recall* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Sedangkan *f-measure* merupakan perbandingan rata-rata presisi dan *recall* yang dibobotkan. Untuk mengetahui nilai presisi, *recall* dan *f-measure* dari tabel 1 bisa dilihat pada tabel 2.

Tabel 2. Hasil Presisi, Recall dan F-measure

| Kelas | Presisi | Recall | F-Measure |
|------------------|---------|--------|-----------|
| yudhistira | 0.86 | 0.88 | 0.87 |
| bima | 0.82 | 0.98 | 0.89 |
| arjuna | 0.71 | 0.7 | 0.71 |
| nakula | 0.96 | 1 | 0.98 |
| sadewa | 0.74 | 0.56 | 0.64 |
| Rata-rata | 0.82 | 0.824 | 0.82 |

Dari hasil tabel 2 kelas yudhistira mempunyai nilai presisi 0.86, nilai *recall* 0.88 dan *f-measure* 0.87. Pada kelas bima mempunyai nilai presisi 0.82, nilai *recall* 0.98 dan *f-measure* 0.89. Pada kelas arjuna mempunyai nilai presisi 0.71, nilai *recall* 0.7 dan *f-measure* 0.71. Pada kelas nakula mempunyai nilai presisi 0.96, nilai *recall* 1 dan *f-measure* 0.98. Sedangkan pada kelas sadewa mempunyai nilai presisi 0.74, nilai *recall* 0.56 dan *f-measure* 0.64. Dari keseluruhan hasil semua kelas rata-rata presisi sebesar 0.82, *recall* sebesar 0.824 dan *f-measure* sebesar 0.82.

4. Kesimpulan

Klasifikasi yang dilakukan oleh SVM bertujuan untuk membuat *hyperplane*. *Hyperplane* menjadi sebuah model yang digunakan untuk melakukan klasifikasi. Nakula dan sadewa merupakan tokoh kembar yang gambar datanya hampir mirip. Namun pada proses prediksi kelas dengan menggunakan data *testingset*. Kelas nakula bisa memperoleh hasil lebih baik dibanding dengan kelas sadewa. Dari penelitian ini harapan penulis bisa melakukan penelitian dengan data yang sama menggunakan fitur lain misalnya menggunakan fitur yang ada di wajah seperti pemrosesan rekognisi wajah.

Referensi

- [1] J. Kronenberger, D. Malysiak, and U. Handman, "Text and character recognition on metal-sheets," *2017 IEEE Int. Conf. Inf. Autom. ICIA 2017*, no. July, pp. 392–397, 2017, doi: 10.1109/ICInFA.2017.8078940.
- [2] Rismiyati, Khadijah, and A. Nurhadiyatna, "Deep learning for handwritten Javanese character recognition," *Proc. - 2017 1st Int. Conf. Informatics Comput. Sci. ICICoS 2017*, vol. 2018-January, pp. 59–63, 2018, doi: 10.1109/ICICOS.2017.8276338.
- [3] Erwin, M. Fachrurrozi, A. Fiqih, B. R. Saputra, R. Algani, and A. Primanita, "Content based image retrieval for multi-objects fruits recognition using k-means and k-nearest neighbor," *Proc. 2017 Int. Conf. Data Softw. Eng. ICoDSE 2017*, vol. 2018-January, pp. 1–6, 2018, doi: 10.1109/ICODSE.2017.8285855.
- [4] P. S. Ha and M. Shakeri, "License Plate Automatic Recognition based on edge detection," *2016 Artif. Intell. Robot. IRANOPEN 2016*, pp. 170–174, 2016, doi: 10.1109/RIOS.2016.7529509.
- [5] H. Hassen and S. Al-Maadeed, "Arabic handwriting recognition using sequential minimal optimization," pp. 79–84, 2017, doi: 10.1109/asar.2017.8067764.
- [6] K. Machhale, H. B. Nandpuru, V. Kapur, and L. Kosta, "MRI brain cancer classification using hybrid classifier (SVM-KNN)," *2015 Int. Conf. Ind. Instrum. Control. ICIC 2015*, no. Icic, pp. 60–65, 2015, doi: 10.1109/IIC.2015.7150592.
- [7] M. Mohamed Fathima, D. Manimegalai, and S. Thaiyalnayaki, "Automatic detection of tumor subtype in mammograms based on GLCM and DWT features using SVM," *2013 Int. Conf. Inf. Commun. Embed. Syst. ICICES 2013*, pp. 809–813, 2013, doi: 10.1109/ICICES.2013.6508213.

- [8] A. Patel and T. V. Kalyani, "Support Vector Machine with Inverse Fringe as Feature for MNIST Dataset," *Proc. - 6th Int. Adv. Comput. Conf. IACC 2016*, pp. 123–126, 2016, doi: 10.1109/IACC.2016.32.
- [9] V. Wasule and P. Sonar, "Classification of brain MRI using SVM and KNN classifier," *Proc. 2017 3rd IEEE Int. Conf. Sensing, Signal Process. Secur. ICSSS 2017*, pp. 218–223, 2017, doi: 10.1109/SSPS.2017.8071594.
- [10] T. Septianto, E. Setyati, and J. Santoso, "Digit Classification of Majapahit Relic Inscription using GLCM-SVM," vol. 1, no. 2, pp. 46–54, 2018.
- [11] B. Sanjaa and E. Chuluun, "Malware detection using linear SVM," *Ifost, Ulaanbaatar*, 2013, pp. 136–138, doi: 10.1109/IFOST.2013.6616872.
- [12] G. Kesavaraj and S. Sukumaran, "A Study On Classification Techniques in Data," 2013, doi: 10.1109/ICCCNT.2013.6726842.
- [13] A. Patle and D. S. Chouhan, "SVM kernel functions for classification," *2013 Int. Conf. Adv. Technol. Eng. ICATE 2013*, 2013, doi: 10.1109/ICAdTE.2013.6524743.