

# Gray Level Co-ocurrence Matrix untuk Pengekstrasian Ciri Topeng Cirebon

Felix Indra Kurniadi<sup>1</sup>, Vinnia Kemala Putri<sup>2</sup>

<sup>1</sup> Fakultas Teknik Informatika, Jurusan Sistem Informasi, Universitas Multimedia Nusantara, Tangerang, Indonesia

felixindra@tau.ac.id

<sup>2</sup> Fakultas Ilmu Komputer, Universitas Indonesia, Depok, Indonesia  
vinnia.kemala51@ui.ac.id

Diterima 5 Maret 2020

Disetujui 17 Juni 2020

**Abstract**—Cirebon mask is one of the intangible cultural heritage in Indonesia. It is one of the prominent cultural assets from Cirebon and becoming one of the identity Cirebon culture. However, the current condition people tend to forget the cultural asset and lack of help from the government makes the Cirebon mask become the third-rate assets. Our concern lays on the extinction of this Mask. We want to implement digitation and automatic identification using image processing techniques. In this paper, we applied the Gray Level Co-occurrence Matrix for extracting the features. K-Nearest Neighbour as the classifier. The best result of this research is done by first order statistical features with the accuracy 40,67%.

**Index Terms**—Cirebon Mask, GLCM, Indonesian Culture, K-Nearest Neighbour

## I. PENDAHULUAN

Topeng Cirebon merupakan salah satu kekayaan budaya Indonesia. Topeng ini merupakan identitas budaya Jawa Barat terutama Cirebon. Topeng Cirebon sendiripun merupakan salah satu bentuk akulturisasi antara adat Jawa Barat dengan Agama. Saat ini topeng Cirebon sering digunakan untuk upacara keadatan maupun pertunjukan hiburan. Sayangnya kebudayaan topeng Cirebon tidak menjadi perhatian dari Pemerintahan maupun kaum generasi muda bahkan berdasarkan data di UNESCO yang menjadi Warisan budaya takbenda dari Indonesia adalah pencak silat, pinisi, tiga jenis tarian bali, Noken, Tari Saman, Angklung, Batik, Kris dan Teater Wayang [1]. Hal ini ditakutkan akan membuat kesenian topeng Cirebon makin menghilang dan dilupakan oleh orang-orang.

Berdasarkan problem ini kami melakukan proses digitalisasi untuk menyimpan kebudayaan topeng Cirebon dalam bentuk gambar. Sayangnya proses mengklasifikasikan topeng-topeng tersebut sangat susah dan memakan waktu yang lama dikarenakan beberapa hal seperti keunikan setiap jenis topeng Cirebon baik secara ornament maupun ekspresinya. Hal inilah yang membuat penulis ingin melakukan proses klasifikasi secara otomatis topeng-topeng Cirebon.

Pengklasifikasian topeng Cirebon pernah dilakukan oleh Kurniadi. Kurniadi mengusulkan penyelesaian klasifikasi topeng Cirebon menggunakan metode robust statistic menggunakan z-score untuk mencari data yang dianggap sebagai outlier. Metode yang diusulkan oleh Kurniadi memberikan hasil penambahan akurasi akan tetapi kelemahan terbesar dari penelitian ini adalah penggunaan metode ekstraksi fitur yang sederhana dalam menyelesaikan data yang kompleks [2].

Saat ini penelitian terhadap topeng Cirebon masih sedikit, oleh karena hal ini maka untuk mendapatkan sebuah keterbaruan yang layak digunakan dalam proses klasifikasi dari topeng Cirebon maka perlu dilakukan proses pencarian state of the art dari berbagai penelitian yang mencoba menyelesaikan permasalahan klasifikasi dan identifikasi terhadap budaya Indonesia lainnya.

Penelitian pernah dilakukan untuk menyelesaikan permasalahan klasifikasi terhadap warisan budaya Indonesia. Harjoko et.al. mencoba mengatasi permasalahan terhadap klasifikasi topeng Bali menggunakan *Convolutional Neural Network* (CNN). Hasil yang diberikan dengan menggunakan metode CNN sudah baik walaupun dalam beberapa kasus tidak mampu membedakan topeng yang kompleks [3]. Aditya et.al. mengusulkan penggunaan *Gray Level Co-occurrence Matrix* (GLCM) dan *statistical color* RGB menggunakan *neural network* sebagai *classifier*-nya hasil yang didapatkan pada penelitian ini adalah kombinasi dari kedua fitur ini menghasilkan nilai akurasi yang baik [4]. Sulistianingsih et.al. mengusulkan metode GLCM untuk ekstraksi fitur dengan *Correlation Based Feature Selection* (CFS) untuk seleksi fitur dan klasifikasi menggunakan metode *K-Nearest Neighbour* (KNN) dan *Backpropagation Neural Network* sebagai klasifier. Hasil terbaik didapatkan menggunakan *Backpropagation Neural Network* [5]. Irawan et.al. menggunakan GLCM untuk klasifikasi Batik Lasem dengan klasifier menggunakan KNN dimana hasil yang didapatkan didapatkan bahwa penggunaan data train yang banyak akan memberikan hasil yang lebih baik dimana akurasi terbaik adalah 73.33% [6].

Berdasarkan penelitian sebelumnya dapat disimpulkan bahwa penggunaan GLCM adalah salah satu metode yang banyak digunakan untuk melakukan proses ekstraksi ciri. Sehingga pada penelitian ini akan menggunakan ekstraksi ciri menggunakan metode GLCM dan *classifier* menggunakan *K-Nearest Neighbour*. Secara garis besar yang dilakukan mirip dengan yang dilakukan oleh Kurniadi akan tetapi perbedaan antara penelitian ini dengan yang dilakukan pada penelitian sebelumnya adalah tidak adanya penggunaan outlier detection, hal ini dilakukan untuk membuktikan bahwa metode GLCM lebih baik dibandingkan metode statistical feature sederhana

Kontribusi penulis dalam penelitian ini adalah:

1. Proses digitalisasi dan identifikasi otomatis berbasis pendekatan *artificial intelligence*.
2. Pengimplementasian metode GLCM pada dataset topeng Cirebon. Dimana GLCM bekerja sebagai ekstraksi fitur.
3. Pengimplementasian metode *classifier* seperti *K-Nearest Neighbour*.
4. Penggunaan augmentasi data untuk mengukur keakuratan model.

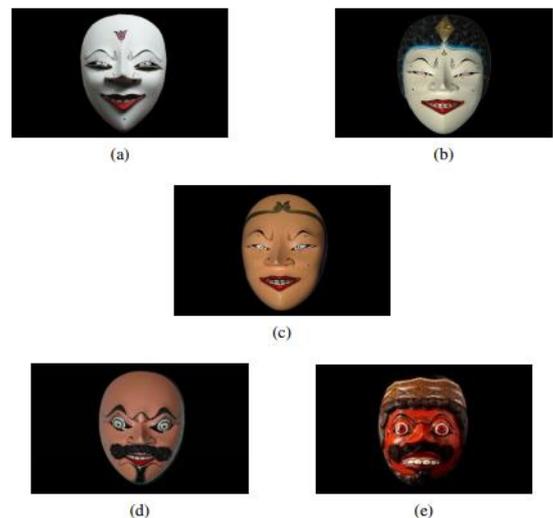
## II. LANDASAN TEORI

Pada bagian ini akan menjelaskan beberapa teori dasar yang mendukung penelitian ini. Secara berurutan akan dijelaskan mengenai Topeng Cirebon, *Gray Level Co-occurrence Matrix*, dan *K-Nearest Neighbour*.

### A. Topeng Cirebon

Topeng Cirebon diciptakan oleh Sunan Kalijaha yang bekerja sama dengan Sunan Gunung Jati, Topeng Cirebon awalnya merupakan salah satu sarana untuk media dakwah agama Islam di pesisir utara pada abad ke-15. Topeng Cirebon sendiri merupakan adaptasi dari kisah Panji yang berasal dari kerajaan Jenggala Jawa Timur [7].

Topeng Cirebon sendiri memiliki lima jenis topeng yang paling terkenal dan sering digunakan dalam pertunjukan yaitu Panji, Samba, Rummyang, Kelana dan Tumenggung. Gambar 1 memberikan gambaran kelima jenis topeng Cirebon.



Gambar 1. Jenis-jenis topeng Cirebon: a) Panji; b) Samba; c) Rummyang; d) Tumenggung; e) Kelana

Setiap topeng Cirebon memiliki karakteristik penokohan yang unik dan terkadang kompleks seperti sifat manusia dimana Panji yang bertindak sebagai gambaran bayi yang baru lahir dan kewibawaan dari seorang raja dimana Panji memiliki ketenangan dan kewibawaan. Samba merepresentasikan sikap suka merendahkan serta kawan dan penuh kaingin tuhan. Rummyang merepresentasikan seseorang yang penuh dengan rasa ingin tahu terhadap lingkungan sekitarnya dan refleksi diri dalam keragu-raguan. Tumenggung merepresentasikan kedewasaan dan kemauan keras serta keberanian dan ambisius. Kelana merepresentasikan Hasrat duniawi yang penuh nafsu dan kekerasan [8].

### B. *Gray Level Co-occurrence Matrix*

*Gray Level Co-occurrence Matrix* (GLCM) adalah sebuah fitur ekstraksi yang mengambil konsep dari seon-order statistical features yang merupakan pecahan dari Harallick features. GLCM diketahui baik dalam mengekstraksi ciri dari tekstur sebuah gambar. Secara general, GLCM membagi gambar dalam 4 angular direction yaitu  $\theta = 00, 45, 90, \text{ and } 135$ . GLCM menggunakan beberapa konsep yaitu *Contrast*, *Homogeneity*, *Energy* dan *Correlation*. Sehingga dari hasil ini setiap gambar akan menghasilkan 16 fitur vektor. *Contrast* merupakan pengukuran perbedaan antara *gray level* dalam sebuah area dalam gambar. *Contrast* sangat bergantung pada proses pengambilan gambar yang diambil [6].

$$\text{contrast} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} (i-j)^2 \cdot P_{ij} \quad (1)$$

*Homogeneity* merujuk pada *Inverse Difference Moment* yang merupakan kemiripan terhadap intensitas yang didapatkan dalam sebuah gambar. Secara garis besar gambar yang homogenus akan

memiliki nilai *gray level* yang *uniform* dan memiliki nilai *inverse value* yang tinggi [6].

$$\text{homogeneity} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} \frac{1}{1 + (i - j)^2} P_{ij} \quad (2)$$

*Correlation* adalah sebuah teknik untuk menghitung secara akurat perubahan pada gambar. Biasanya *correlation* digunakan untuk mengukur ketergantungan linear terhadap intensitas [6].

$$\text{correlation} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} \frac{(i - \mu_i)(j - \mu_j)P_{ij}}{\sigma_i \sigma_j} \quad (3)$$

Energi menghitung keseragaman antar *pixel* yang memiliki pasangan berulang dalam citra [6].

$$\text{energy} = \sum_{i=1}^M \sum_{j=1}^M (P_{i,j})^2 \quad (4)$$

Dimana  $P(i,j)$  adalah matriks GLCM,  $\mu$  adalah mean dan  $\sigma$  adalah standar deviasi dari matriks GLCM dan  $i$  serta  $j$  merepresentasikan baris dan kolom dari matriks GLCM.

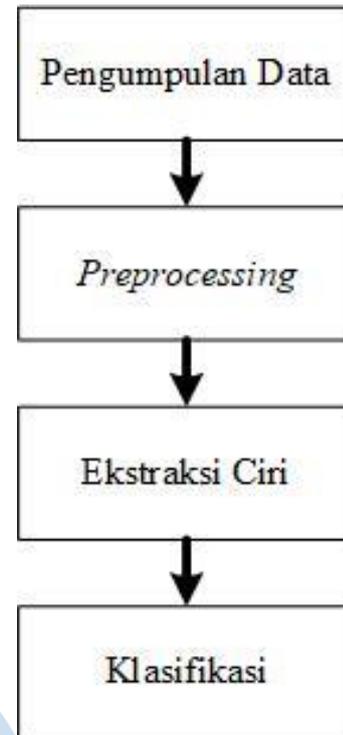
### C. K-Nearest Neighbour

*K-Nearest Neighbour* (KNN) merupakan salah satu algoritma sederhana yang sering digunakan dikarenakan banyak data yang secara linear dapat dipisahkan. Hal inilah yang membuat KNN merupakan salah satu algoritma yang tangguh [9]. Secara garis besar KNN melihat kedekatan antar  $k$ -data. Dimana data test yang dimasukkan akan melihat persebaran tetangga terdekat dan mengambil kelas berdasarkan kelas terbanyak dari tetangga terdekat [9].

Metode pengukuran jarak yang digunakan adalah Euclidean *distance* dan berdasarkan penelitian yang pernah dilakukan oleh Kurniadi mengenai perbandingan *distance metric* [10] bahwa penggunaan Euclidean dan Minkowski *distance* memberikan hasil terbaik dibandingkan dengan metode *distance metric* lainnya. Persamaan dibawah merupakan persamaan untuk Euclidean *distance*.

$$d(a, b) = \sqrt{\sum_{j=1}^n (a_j - b_j)^2} \quad (5)$$

Dimana  $a, b$  adalah dua data dan  $d$  merepresentasikan jarak dan  $j$  adalah banyaknya data.



Gambar 2. Metodologi penelitian klasifikasi topeng Cirebon

## III. METODOLOGI

Pada Bagian ini akan dibahas mengenai proses pengumpulan data dan bagaimana data tersebut diolah. Pra-proses terhadap data yang digunakan, proses ekstraksi ciri dan proses klasifikasi. Gambar 2 menggambarkan secara runtut proses yang dilakukan dalam penelitian yang dilakukan saat ini.

### A. Pengumpulan Data

Data dikumpulkan menggunakan search engine dan pengambilan data primer ke lapangan. Penggunaan data search engine yang tersedia untuk publik dilakukan sebagai benchmark apakah data yang bersifat publik cukup relevan dengan data yang didapatkan dari lapangan. Hasil dari pengumpulan data menggunakan search engine dan pengambilan data primer menghasilkan total 150 gambar yang dibagi menjadi 30 untuk setiap kelas. Untuk membuat representasi yang baik maka pembagian data untuk menjadi data latih dan data uji menggunakan perbandingan 80:20. Dimana data latih menjadi sebesar 24 dan data uji sebesar 6.

Data yang didapatkan secara manual ini dikembangkan dengan menggunakan proses augmentasi. Proses augmentasi dilakukan untuk membuat dataset menjadi lebih banyak dengan melakukan proses pemberian derau, rotasi dan peningkatan intensitas. Hasil dari augmentasi citra ini mendapatkan total 4500 data dengan 900 citra untuk setiap kelas. Hasil ini kemudian dibagi menjadi data

latih dan data uji. Tabel 1 sampai Tabel 2 menggambarkan pembagian data dan jumlah data.

Tabel 1. Total dataset dari *original* dataset dengan augmentasi dataset

Original	150
Augmentasi	4500

Tabel 2. Dataset per-kelas dari original dataset dengan augmentasi dataset dan pembagian antara data latih dan data uji

	train	test	Total
Original	24	6	30
Augmentasi	792	198	900

### B. Pra Proses

Pengontrolan dataset dilakukan dengan menghapus *background* pada citra dan merubah gambar menjadi keabuan. Selain kedua proses tersebut *window size* dari gambar juga diubah menjadi  $100 \times 100$  *pixel*. Proses ini dilakukan dengan bantuan perangkat lunak untuk memanipulasi gambar.

### C. Fitur Ekstraksi

Pada proses ekstraksi fitur dilakukan pengekstrasian ciri citra dengan menggunakan metode GLCM yang sudah dijelaskan pada Bagian II. Proses ini dilakukan dengan menggunakan semua  $\theta$  sehingga menghasilkan 16 fitur vector yang merupakan gabungan antara 4  $\theta$  dan 4 GLCM algoritma yaitu *homogeneity*, *contrast*, *correlation* dan *energy*.

Hasil dari proses ekstraksi fitur ini kemudian akan dibuatkan model dengan menggunakan metode yang diusulkan yaitu *K-Nearest Neighbour*.

### D. Klasifikasi

Pada proses klasifikasi dilakukan pemasangan model secara garis besar proses klasifikasi berbeda antara proses latih dan uji, dimana pada proses latih fitur yang didapatkan pada proses ekstraksi fitur dimasukkan kedalam metode KNN dan metode KNN akan melakukan persebaran data yang nantinya akan menjadi model utama yang akan digunakan dalam proses pengujian.

Pada proses pengujian fitur yang dimasukkan ke dalam model dan model akan menghasilkan nilai kelas dan jika nilai kelas yang dihasilkan sesuai dengan nilai kelas yang diinginkan maka model dapat mengenali data tersebut dengan baik.

Proses evaluasi pada penelitian ini akan menggunakan akurasi, presisi dan *recall*.

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn} * 100\% \quad (6)$$

$$precision = \frac{tp}{tp + fp} * 100\% \quad (7)$$

$$recall = \frac{tp}{tp + fn} * 100\% \quad (8)$$

Dimana *tp*, *tn*, *fp* dan *fn* merupakan *true positive*, *false positive*, *true negative* dan *false negative*. Konsep ini diambil dari konsep *confusion matrix*.

## IV. EKSPERIMEN DAN HASIL

Pada bagian ini akan dijelaskan mengenai skenario eksperimen yang dilakukan serta hasil yang didapatkan dari penelitian ini. Selain itu akan dijelaskan mengenai hasil yang didapatkan.

### A. Skenario Eksperimen dan Konfigurasi Eksperimen

Pada penelitian ini akan digunakan KNN dengan membandingkan 3 jenis *K-Neighbour* yaitu bernilai 3, 5 dan 7. Selain itu data yang akan digunakan adalah data *original* dan augmentasi.

Eksperimen ini dilakukan menggunakan CPU dengan prosesor Intel i7 dan RAM sebesar 4GB. Proses pembagian train dan test dilakukan secara manual.

Tabel 3. Hasil eksperimen pada data original

Metode	Neighbour	akurasi	presisi	recall
First order statistical[2]	k=3	40	0.47	0.4
	k=5	40	0.42	0.44
	k=7	36.67	0.43	0.37
GLCM	k=3	16.67	0.11	0.16
	k=5	26.67	0.2	0.27
	k=7	26.67	0.23	0.27

Tabel 4. Hasil eksperimen pada data augmentasi

Metode	Neighbour	akurasi (%)	presisi	recall
First order statistical[2]	k=3	40.69	0.4	0.41
	k=5	40.69	0.4	0.41
	k=7	40.49	0.39	0.41
GLCM	k=3	33.14	0.34	0.33
	k=5	33.17	0.34	0.33
	k=7	31.07	0.31	0.31

Berdasarkan hasil di atas bahwa metode GLCM yang diusulkan masih tidak mampu mengalahkan metode first order statistic yang sudah dilakukan oleh Kurniadi [2] sebelumnya. Secara keseluruhan akurasi yang didapatkan baik dari data augmentasi dan bukan augmentasi mengecewakan dikarenakan hasil tidak dapat menghasilkan nilai lebih dari 50%. Hal ini merupakan problem yang dikemukakan oleh [2] pada penelitian sebelumnya dimana dataset public yang tersedia tidak merepresentasikan data yang dilapangan sehingga tidak cocok digunakan.

Proses augmentasi datapun tidak merubah hasil untuk *first order statistical feature* akan tetapi pada metode GLCM memberikan peningkatan hasil. Sayangnya penggunaan data augmentasi tidak terlalu berperan banyak dalam bahkan peningkatan signifikan hanya terjadi pada  $k=3$  dengan metode GLCM.

#### V. SIMPULAN

Penelitian pengklasifikasian topeng Cirebon menggunakan metode GLCM memberikan beberapa hasil baru salah satunya adalah data publik yang digunakan tidak dapat digunakan menjadi acuan yang tepat sehingga harus dilakukan observasi lapangan yang lebih detail. Metode GLCM yang diusulkan tidak memberikan hasil yang memuaskan bahkan metode yang diusulkan tidak dapat menandingi metode yang pernah diusulkan sebelumnya, yang dimana hasil akurasi terbaik dari metode GLCM adalah 33.14% pada data augmentasi. Selain itu penggunaan data augmentasi pada penelitian ini tidak terlalu berpengaruh.

Penelitian ini bisa menjadikan dasar untuk penelitian lain yang ingin berfokus pada klasifikasi topeng Cirebon.

#### UCAPAN TERIMA KASIH

Penulis ingin mengucapkan terima kasih kepada Fendy Hendriyanto sebagai pengumpul data dan penulis juga berterima kasih kepada Kementerian Riset dan Teknologi Republik Indonesia.

#### DAFTAR PUSTAKA

- [1] United Nations Educational, Scientific and Cultural Organization, "Browse the Lists of Intangible Cultural Heritage and the Register of good safeguarding practices," Intangible Cultural Heritage. [Online]. Available: [https://ich.unesco.org/en/lists?text=&country\[\]=00104&multi-national=3&display=inscriptionID#tabs](https://ich.unesco.org/en/lists?text=&country[]=00104&multi-national=3&display=inscriptionID#tabs). [Accessed: 21-Feb-2020].
- [2] F. I. Kurniadi and F. Hendriyanto, "Cirebon Mask Classification using Robust K-Nearest Neighbour," in 2019 International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2019, pp. 143–146, doi: 10.1109/ICOIACT46704.2019.8938499.
- [3] B. P. Ida Bagus Gede and A. Harjoko, "Pengenalan Topeng Bali Menggunakan Algoritma Convolutional Neural Network," Master's Thesis, Universitas Gadjah Mada, 2017.
- [4] C. S. K. Aditya, M. Hani'ah, R. R. Bintana, and N. Suciati, "Batik classification using neural network with gray level co-occurrence matrix and statistical color feature extraction," in 2015 International Conference on Information Communication Technology and Systems (ICTS), 2015, pp. 163–168, doi: 10.1109/ICTS.2015.7379892.
- [5] N. Sulistianingsih, I. Soesanti, and R. Hartanto, "Classification of Batik Image using Grey Level Co-occurrence Matrix Feature Extraction and Correlation Based Feature Selection," in 2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 2018, pp. 492–497, doi: 10.1109/ISRITI.2018.8864237.
- [6] C. Irawan, E. N. Ardyastiti, D. R. I. M. Setiadi, E. H. Rachmawanto, and C. A. Sari, "A Survey: Effect of the Number of GLCM Features on Classification Accuracy of Lasem Batik Images using K-Nearest Neighbor," in 2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 2018, pp. 33–38, doi: 10.1109/ISRITI.2018.8864443.
- [7] Lasmiyati, "Sejarah Pertumbuhan dan Pengembangan Tari Topeng Cirebon Abad XV-XX," Patanjala, vol. 3, no. 3, pp. 472–487, Sep. 2011.
- [8] Pusat Studi Sunda, Kujang, bedog, dan topeng dan kajian lainnya mengenai budaya Sunda, Seri Sunda. Pusat Studi Sunda, 2008.
- [9] F. I. Kurniadi and V. K. Putri, "A Comparison of Human Crafted Features and Machine Crafted Features on White Blood Cells Classification," J. Phys.: Conf. Ser., vol. 1201, p. 012045, May 2019, doi: 10.1088/1742-6596/1201/1/012045.
- [10] F. I. Kurniadi and V. K. Putri, "Perbandingan Distance Metric pada Nearest Neighbour untuk Klasifikasi Sel Darah Putih," Ultimatics, vol. XI, no. 1, p. 4.