

KLASIFIKASI MOTIF BATIK JAWA MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBORS (KNN)

Muhamad Deni Akbar^{1)*}, Martanto²⁾, Yudhistira Arie Wijaya³⁾

^{1,2,3)} Teknik Informatika, STMIK IKMI Cirebon, Kota Cirebon, Jawa Barat
email: denyakbar05@gmail.com¹⁾, martantomusijo@gmail.com²⁾,
yudhistira010471@gmail.com³⁾

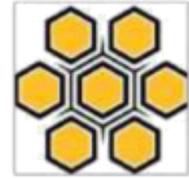
Abstrak

Batik merupakan salah satu warisan Indonesia yang indah dan terkenal di seluruh dunia, batik sebagai warisan tradisional Nusantara hadir dengan beragam motif. Setiap wilayah mempunyai motif yang berbeda dan melambangkan filosofi yang beragam. Banyaknya motif batik Indonesia yang tersebar dari Aceh hingga Papua, sehingga tidak semua orang dapat membedakan motif batik. Penelitian ini bertujuan untuk membedakan motif batik yang berasal dari pulau Jawa dan motif batik non Jawa. Motif batik yang diambil peneliti sebagai sampel dari wilayah Jawa adalah motif batik Mega Mendung, motif batik Lasem, motif batik Sekar Jagad, motif batik Kawung, dan untuk wilayah non Jawa peneliti mengambil sampel motif batik Cendrawasih, motif batik Dayak, motif batik Pala, dan motif batik Bali. Metode penelitian menggunakan algoritma K-Nearest Neighbors (KNN) yang memiliki tahapan yaitu pengumpulan citra motif batik, penentuan motif batik Jawa dan non Jawa, pre-processing, ekstraksi fitur, klasifikasi citra, dan evaluasi motif. Ekstraksi fitur warna dilakukan dengan menggunakan metode gray level co-occurrence matrix (GLCM). Berdasarkan hasil pengujian menunjukkan bahwa dengan penerapan metode GLCM dan KNN menggunakan size gambar 200x200 dengan ratio $k=5$ presentasi split citra 80% test dan 20% training mampu menghasilkan akurasi sebesar 65%.

Kata Kunci : Batik, Klasifikasi, K-Nearest Neighbors (KNN), Size.

Abstract

Batik is one of Indonesia's beautiful and well-known heritages throughout the world, batik as a traditional heritage of the archipelago comes with a variety of motifs. Each region has different motifs and different philosophies. The number of Indonesian batik motifs spread from Aceh to Papua, so not everyone can distinguish batik motifs. This study aims to distinguish between Javanese batik motifs and non-Javanese batik motifs. The batik motifs taken by researchers as samples from the Java region were the Mega Mendung batik motif, Lasem batik motif, Sekar Jagad batik motif, Kawung batik motifs and motifs, and for non-Javanese researchers took samples of Cendrawasih batik motifs, Dayak batik motifs, and batik motifs. nutmeg, and Balinese batik motifs. The research method uses the K-Nearest Neighbors (KNN) algorithm which has stages of collecting image on batik motifs, knowing Javanese and non-Javanese batik motifs, pre-processing, feature extraction, image classification, and evaluation of motifs. Color feature extraction is carried out using gray level co-occurrence matrix (GLCM) methods. Based on the test results show that with the application of GLCM and KNN with an image size ratio of 200x200 with a ratio of $k=5$ the percentage of split image 80% test and 20% training is able to produce an accuracy of 65%..



Keywords: Batik, Classification, K-Nearest Neighbors (KNN), Size.

PENDAHULUAN

Sebagai salah satu peninggalan Indonesia, batik berkembang dari pakaian resmi hingga baju santai. Perihal ini didukung oleh realitas karena seni batik digunakan dalam banyak baju, sepatu, helm, dompet, ataupun hal-hal lain. Kain batik dibuat dengan menyiapkan kain mori, membuat desain batik, melukis di kain dengan memanfaatkan parafin yang sudah dicairkan, mengisi seluruh bagian putih pada kain dengan memberi warna pada pola dengan parafin, setelah itu beri warna kain pada bagian yang tidak diberi parafin, lukis kembali dengan canting, hilangkan parafin dan kembali membatik, hilangkan lapisan parafin agar motif [1].

Batik sudah tumbuh dalam bermacam macam motif semenjak seni batik diadopsi oleh banyak budaya yang terdapat di pulau pulau di Indonesia. Untuk melestarikan batik, banyak seniman membuat pembaruan dalam motif yang sudah ada agar lebih di minati. Batik telah diturunkan oleh nenek moyang Indonesia, tetapi pengetahuan tidak diturunkan kepada keturunannya dalam pembelajaran resmi.

Dari yang terlansir di situs nasional kemenparekraf.go.id setidaknya, saat ini ada sekitar 5.849 motif batik di Indonesia yang tersebar dari Aceh hingga Papua. Banyaknya motif batik tersebut hendak mengakibatkan sulitnya dalam membedakan motif dari tiap-tiap batik. Sangat banyak warga Indonesia modern tidak bisa mengidentifikasi tipe ataupun batik yang mereka pakai [2], maka diperlukan klasifikasi untuk menangani hal ini menggunakan algoritma KNN.

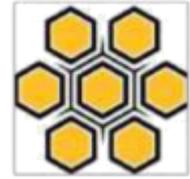
Klasifikasi ialah suatu proses analisa pada citra yang menghasilkan model guna

menggambarkan kelas yang terdapat pada citra tersebut [3].

K-Nearest Neighbors (KNN) adalah metode yang menggunakan algoritma supervised yang mana hasil dari query instance baru diklasifikasi bersumber pada kebanyakan dari jenis KNN. Algoritma KNN memakai supervised learning guna mendapatkan pola baru dalam informasi dengan menghubungkan pola informasi yang telah terdapat dengan informasi yang baru [4]. KNN merupakan cara guna menerapkan klasifikasi terhadap objek bersumber pada informasi yang amat mirip (Nearest Neighbors) dengan jumlah K yang sudah ditentukan serta mengklasifikasikannya ke dalam kelas baru [5].

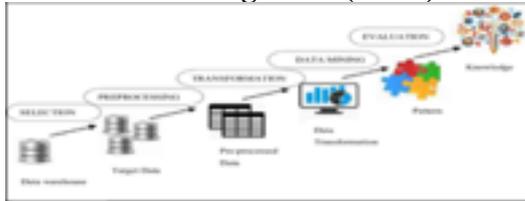
Klasifikasi citra merupakan pengelompokan piksel-piksel tertentu kelas dengan rujukan yang sudah ditentukan. Proses pengelompokannya di tetapkan berdasarkan kedekatan, kesamaan ataupun kesamaan nilai piksel pada citra. KNN merupakan tata cara mengelompokkan objek bersumber pada informasi yang mempunyai jarak terdekat dengan objek [6], dengan bantuan fitur Color to Grayscale mengekstraksi fitur dengan menghitung tingkat keabuan citra yang dekat / citra tetangga [7] yang hanya membaca motifnya saja tidak dari warna gambar tersebut sehingga sistem bisa dengan dengan mudah mempelajari dan melatih citra.

Berdasarkan hal tersebut penelitian ini bertujuan untuk mengetahui berapa tingkat akurasi pada klasifikasi gambar motif batik jawa menggunakan aplikasi rapid miner dan mengetahui faktor apa saja yang berpengaruh dalam hasil akurasi.



LANDASAN TEORI

A. K-Nearest Neighbors (KNN)



Gambar 1. Alur KDD

Algoritma K-Nearest Neighbors (KNN) merupakan salah satu algoritma yang dapat diterapkan sebagai solusi untuk mengklasifikasi sebuah citra. KNN merupakan suatu tata cara guna menerapkan klasifikasi terhadap objek bersumber pada citra pembelajaran yang jaraknya sangat dekat dengan objek tersebut. Tingkatan akurasi sesuatu kategori motif batik dipengaruhi oleh seberapa besar nilai K.

Menurut penelitian Ida Ayu, A. Angreni, Sakti Adji Adisasmita, M. Isran Ramli, dan Sumarni Hamid dengan judul “Pengaruh Nilai K Pada Metode K-Nearest Neighbor (KNN) Terhadap Tingkat Akurasi Identifikasi Kerusakan Jalan” Pengolahan citra pada penelitian ini meliputi 2 sesi, ialah proses ekstraksi dengan tahapan: wiener filtering serta thresholding, sebaliknya proses klasifikasi dengan tata cara KNN.

Uji coba nilai k yang berbeda-beda, ialah 1, 8, serta 15, menciptakan tingkatan akurasi yang berbeda yakni 84%, 8%, 18% masing-masing tipe kerusakan. Nilai k menerangkan berapa banyak jumlah citra yang terdekat dengan suatu objek, Jumlah citra yang terdekat yang berbeda dapat mempengaruhi hasil klasifikasi terhadap suatu citra [8].

B. Size

Size merupakan ukuran seberapa besar pixel sebuah gambar, besar kecilnya sebuah

gambar dapat mempengaruhi hasil dari klasifikasi citra. Semakin besar dan jelas sebuah gambar akan menghasilkan akurasi lebih besar.

Menurut penelitian Jiamei Sun & Alexander Binder dengan judul “Comparison of Deep Learning Architectures for H&E Histopathology Images” Dampak dari ukuran size pada evaluasi training sample sangat berpengaruh. Penelitian ini Menggunakan dataset BreKHis yang diri dari pemindaian bernoda H&E dari jaringan kanker hasil menunjukkan bahwa semakin besar gambar dari sebuah data training maka akan membuat performasi dalam akurasi semakin baik pula [9].

METODE PENELITIAN

Metode penelitian yang digunakan pada penelitian Knowledge Discovery of Databases (KDD). Dimana ada beberapa tahapan seperti pada gambar 1.

Data penelitian diperoleh dengan mengunduh data di situs <https://www.kaggle.com/alfanme/indonesian-batik-motifs-corak-app>. Data batik berjumlah 1049 File dengan jenis 15 data motif batik nasional.

1. Data Selection

Pemilihan dari dataset harus dilakukan sebelum tahap penggalian informasi dalam KDD, Pemilihan data adalah proses menganalisa data yang relevan dari database.

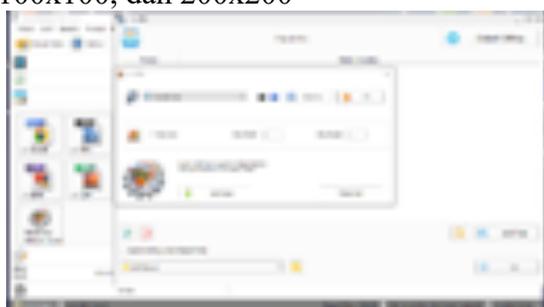
Untuk mempermudah proses menganalisa data menggunakan algoritma KNN maka perlu dilakukan proses pembersihan data yang bertujuan untuk menghilangkan data yang tidak relevan, data duplikat, pemilihan motif batik yang digunakan peneliti, pemilihan kualitas size gambar terbaik dan pemilihan format



gambar terbaik. Pada penelitian ini data yang diambil ialah data yang terbagi menjadi 2 kelas dengan masing – masing data 50 gambar yaitu data motif batik jawa dengan data batik kawung, batik lasem, batik mega mendung, dan batik sekar jagad. Motif batik non jawa dengan data batik bali, batik cendrawasih, batik dayak, dan batik palu.

2. Preprocessing

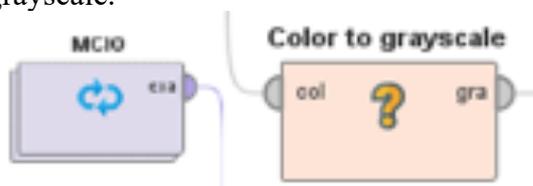
Proses cleansing sebelum proses data mining dapat dilakukan, pada tahap ini peneliti memilih gambar dengan merubah ukuran gambar dari size 224x224 menjadi beberapa size yakni 20x20, 80x80, 100x100, dan 200x200



Gambar 2. Format Factory

3. Transformation

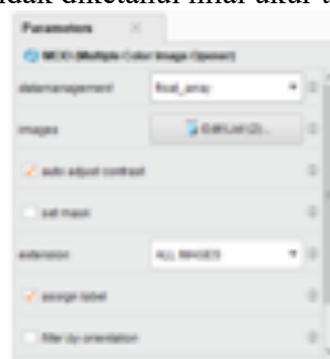
Pada tahapan ini dilakukan proses perubahan data, sebelum data dapat diolah dengan menggunakan algoritma K-Nearest Neighbors. Berdasarkan gambar di atas menjelaskan bahwa teknik transformation menggunakan operator MCIO dan Color to grayscale.



Gambar 3. MCIO & Color to Srayscale

Pada Operator MCIO parameter yang di pilih pada data management yaitu float

array berfungsi untuk mengkonversi seberapa banyak angka yang angka di konversi menjadi bilangan desimal yang angkanya tidak boleh lebih dari 7 di belakang koma (,) pemilihan float array dalam operator ini adalah agar mengetahui nilai detail dalam angka karena jika yang digunakan int array nilai yang akan keluar dari contoh hasil ukur 0,33 adalah 0 yang berarti tidak diketahui nilai ukur tersebut.



Gambar 4. Float Array

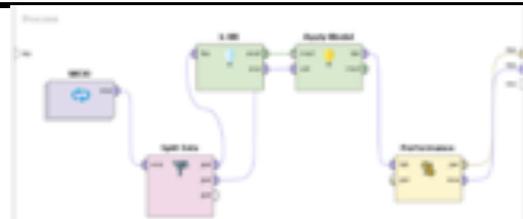
Fungsi dari operator Color to Grayscale bertujuan mengubah gambar berwarna menjadi skala abu abu agar fokus dalam sistem mempelajari motifnya saja bukan dari warna gambar lalu, mengubahnya menjadi data integer.



Gambar 5. Example Set MCIO

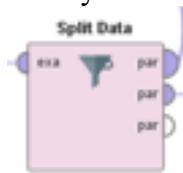
4. Data Mining

Pada tahapan ini penelitian dilakukan dengan mengelompokkan data batik mejadi 2 class yaitu batik jawa dan batik non jawa, masing masing kelas berisikan 200 data dengan total jumlah yaitu 400 data dilakukan dari operator MCIO hingga operator Performance.



Gambar 6. Alur Klasifikasi Batik KNN

Data yang digunakan berjumlah 400 data, size gambar 200X200 dengan format gambar JPG, data mining dilakukan dengan menggunakan operator MCIO yang berfungsi mengolah gambar sebelum dilakukan klasifikasi. Dalam klasifikasi dilakukan data test dan data train menggunakan operator split data, data yang di test maupun di train tidak boleh kurang dari 0,2 jumlah rasio test dan train harus berjumlah 1. Rasio pembelajaran sampel memutuskan berapa besar sistem harus belajar dan melatihnya.



Gambar 7. Split Data

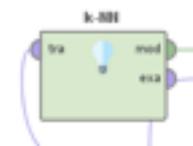
Pada operator split data sampling type yang dipilih stratified sampling membangun himpunan bagian acak dan memastikan bahwa distribusi kelas dalam himpunan bagian sama dengan di seluruh kumpulan data.



Gambar 8. Stratified Sampling

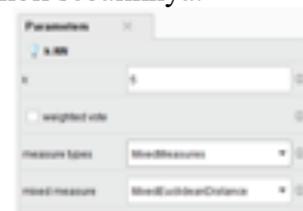
Algoritma KNN digunakan berdasarkan perbandingan jumlah k yang ditentukan, k

sendiri dalam knn merupakan seberapa banyak jumlah data terdekat yang akan dijadikan contoh. Jumlah minimal k yang harus digunakan minimal 1 dan tidak boleh 0 karena membutuhkan seberapa banyak yang akan nilai k baca.



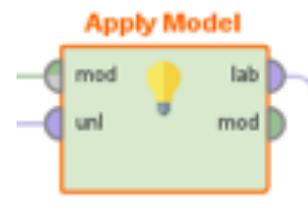
Gambar 9. KKN

Pemilihan measure types yang digunakan mixed measures digunakan untuk menghitung jarak dalam hal atribut nominal dan numerik. Pemilihan tipe mixed measure yang digunakan adalah mixed euclidean distance, untuk nilai nominal, jarak 0 diambil jika kedua nilai sama dan jarak satu di ambil sebaliknya.



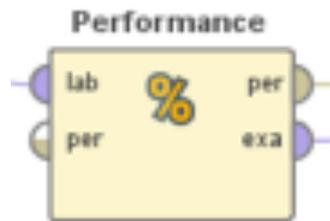
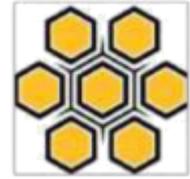
Gambar 10. Mixed Measures

Apply Model merupakan operator untuk menerapkan model yang diperoleh dari hasil data train dan data test.



Gambar 11. Apply Model

Performance digunakan untuk mengevaluasi hasil kinerja dalam klasifikasi data, Operator ini secara otomatis menentukan hasil apa yang kita ingin tampilkan.



Gambar 12. Performace

5. Interpretasi / Evaluation

Pengujian akurasi dataset dengan ratio test 8 train 2, test 2 train 8, dan test 5 train 5, dengan nilai $k=5$ pada gambar yang beragam size menunjukkan semakin besar size sebuah gambar akan memberikan akurasi yang besar pula.

Tabel 1. Akurasi Size

| No | Pixels | K | Ratio | | Accuracy |
|----|---------|---|-------|-------|----------|
| | | | Test | Train | |
| 1 | 20X20 | 5 | 8 | 2 | 52,70% |
| 2 | 20X20 | 5 | 2 | 8 | 56,95% |
| 3 | 20X20 | 5 | 5 | 5 | 57,07% |
| 4 | 80X80 | 5 | 8 | 2 | 63,75% |
| 5 | 80X80 | 5 | 2 | 8 | 50,31% |
| 6 | 80X80 | 5 | 5 | 5 | 60,00% |
| 7 | 100X100 | 5 | 8 | 2 | 63,75% |
| 8 | 100X100 | 5 | 2 | 8 | 56,25% |
| 9 | 100X100 | 5 | 5 | 5 | 59,00% |
| 10 | 200X200 | 5 | 8 | 2 | 65,00% |
| 11 | 200X200 | 5 | 2 | 8 | 54,06% |
| 12 | 200X200 | 5 | 5 | 5 | 61,50% |

Perbandingan nilai resio test dan train mempengaruhi hasil dari akurasi pada sebuah klasifikasi gambar.

Tabel 2. Akurasi Ratio

| No | Pixels | K | Ratio | | Accuracy |
|----|---------|---|-------|-------|----------|
| | | | Test | Train | |
| 1 | 200X200 | 5 | 2 | 8 | 54,06% |
| 3 | 200X200 | 5 | 4 | 6 | 58,75% |
| 4 | 200X200 | 5 | 8 | 2 | 65,00% |
| 6 | 200X200 | 5 | 6 | 4 | 63,75% |
| 7 | 200X200 | 5 | 5 | 5 | 56,00% |

Perbandingan Nilai K pada klasifikasi gambar mempengaruhi hasil akurasi.

Tabel 3. Akurasi K

| No | Pixels | K | Ratio | | Accuracy |
|----|---------|---|-------|-------|----------|
| | | | Test | Train | |
| 1 | 200X200 | 1 | 8 | 2 | 50,00% |
| 2 | 200X200 | 2 | 8 | 2 | 52,50% |
| 3 | 200X200 | 3 | 8 | 2 | 58,75% |
| 4 | 200X200 | 4 | 8 | 2 | 57,50% |
| 5 | 200X200 | 5 | 8 | 2 | 65,00% |
| 6 | 200X200 | 6 | 8 | 2 | 61,25% |
| 7 | 200X200 | 7 | 8 | 2 | 62,50% |



| | | | | | |
|----|---------|----|---|---|--------|
| 8 | 200X200 | 8 | 8 | 2 | 56,25% |
| 9 | 200X200 | 9 | 8 | 2 | 57,50% |
| 10 | 200X200 | 10 | 8 | 2 | 53,75% |

HASIL DAN PEMBAHASAN

Proses klasifikasi dilakukan dengan meneliti 3 faktor yaitu, perbedaan data pada size gambar, proses training testing, dan nilai k. Proses training dan testing atau pelatihan bertujuan untuk melatih sistem, proses yang dilakukan peneliti pada percobaan rasio sebanyak 12 kali dengan berbagai size dan rasio percobaan yang berbeda.

Proses terakhir adalah percobaan dengan membedakan nilai k pada size gambar 200x200 dengan rasio 8,2. Nilai K yang di uji sebanyak 1 sampai 10 nilai k, hasil akurasi terbaik terdapat dengan nilai k 5.

| | Real Jawa | Real Non Jawa | Total |
|---------------|-----------|---------------|-----------|
| Pred Jawa | 8 (TP) | 2 (FP) | 10 (80%) |
| Pred Non Jawa | 2 (FN) | 8 (TN) | 10 (80%) |
| Total | 10 (80%) | 10 (80%) | 20 (100%) |

Gambar 13. Nilai K

Keterangan:

True Positive (TP): pred jawa – true jawa, jumlah data yang diprediksi jawa dan kenyataannya jawa

True Negative (TN): pred non jawa – true non jawa, jumlah data yang diprediksi non jawa dan kenyataannya non jawa

False Positive (FP): pred jawa – true non jawa, jumlah data yang diprediksi jawa tapi kenyataannya adalah non jawa

False Negative (FN): pred non jawa – true jawa, jumlah data yang diprediksi non jawa tapi kenyataannya adalah jawa

KESIMPULAN

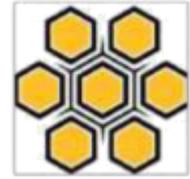
Dari hasil akurasi terhadap 10 nilai k tersebut, dapat dilihat kecenderungan nilai akurasi akan bertambah maupun berkurang dengan besar kecilnya nilai k. Penurunan akurasi terlihat drastis untuk nilai k=1 dengan akurasi sebesar 50% dan hasil akurasi terbesar terlihat dengan nilai k=5 sebesar 65%. Perbedaan tersebut bisa disebabkan oleh penyebaran data yang berbeda untuk masing masing jenis data batik

UCAPAN TERIMA KASIH

Puji syukur kami panjatkan kepada Allah Subhanahu Wa Ta'ala atas segala rahmat yang telah diberikan-Nya, sehingga penulis dapat menyelesaikan jurnal yang berjudul “Klasifikasi Motif Batik Jawa Menggunakan Algoritma K-Nearest Neighbors (KNN).

Dalam penyusunan jurnal teknik informatika ini kami sangat banyak mendapatkan bantuan macam-macam dari pihak, Oleh karena itu dengan penuh hormat kami ingin mengucapkan terima kasih kepada:

1. Kedua orang tua tercinta yang telah banyak memberikan bantuan baik secara moral maupun materi, serta dorongan semangat dan berkat do'a yang tak ternilai.
2. Bapak Dr. Dadang Sudrajat, S.Si., M.Kom., selaku Ketua STMIK IKMI Cirebon.
3. Ibu Ade Irma Purnamasari, M.Kom, selaku Sekretaris Ketua STMIK IKMI Cirebon
4. Bapak Dian Ade Kurnia, M.Kom, selaku Wakil Ketua I Bidang Akademik, Riset dan Inovasi



5. Ibu Dra. Nining R, M.Si., selaku Wakil Ketua II Bidang Keuangan.
 6. Bapak H. Eka Jayawangsa, BBA., selaku Wakil Ketua III Bidang Kerjasama dan Kemahasiswaan
 7. Ibu Gifthera Dwilestari, S.I.Kom., M.Kom, Sebagai Ketua Program Studi Teknik Informatika.
 8. Bapak Martanto M.Kom., Sebagai Pembimbing Utama.
 9. Bapak Yudhistira Arie Wijaya M.Kom., Sebagai Pembimbing Kedua.
 10. Teman-teman dan semua pihak yang sudah membantu, terimakasih banyak.
- DAFTAR PUSTAKA**
- [1] I. M. Ihdal, "Klasifikasi Kain Khas Batik Dan Kain Khas Sasirangan Dengan Menggunakan Metode Convolutional Neural Network," *J. Teknol. Inf. Univ. Lambung Mangkurat*, vol. 6, no. 1, pp. 25–30, Apr. 2021, doi: 10.20527/jtiulm.v6i1.62.
 - [2] I. M. A. Agastya and A. Setyanto, "Classification of Indonesian batik using deep learning techniques and data augmentation," *Proc. - 2018 3rd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2018*, pp. 27–31, 2018, doi: 10.1109/ICITISEE.2018.8720990.
 - [3] Y. A. Wijaya et al., "Analisa Klasifikasi menggunakan Algoritma Decision Tree pada Data Log Firewall Jurnal Sistem Informasi dan Manajemen," *J. STMIK GICI*, vol. 9, no. 3, 2021, [Online]. Available: <https://ejournal.stmikgici.ac.id/index.php/jursima/article/view/303>.
 - [4] Y. Kusumawati et al., "A Classification Of Batik Kudus Based On Patterns Using K-NN And GLCM," no. February, 2021.
 - [5] R. Ripai, "Pengenalan Motif Batik Pandeglang Menggunakan Deteksi Tepi Canny dan Metode K-NN Berbasis Android ini mengambil beberapa ringkasan deteksi motif batik berbasis platform android . Dari hasil," vol. XVI, 2021.
 - [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," *2018 Int. Semin. Res. Inf. Technol. Intell. Syst. ISRITI 2018*, pp. 33–38, 2018, doi: 10.1109/ISRITI.2018.8864443.
 - [7] Z. A. A. Feri Agustina, "Identifikasi Citra Daging Ayam Kampung dan Broiler Menggunakan Metode GLCM dan Klasifikasi-NN," *J. Infokam*, vol. XVI, no. 1, pp. 25–36, 2020.
 - [8] I. A. Angreni, S. A. Adisasmita, M. I. Ramli, and S. Hamid, "Pengaruh Nilai K Pada Metode K-Nearest Neighbor (Knn) Terhadap Tingkat Akurasi Identifikasi Kerusakan Jalan," *Rekayasa Sipil*, vol. 7, no. 2, p. 63, 2018, doi: 10.22441/jrs.2018.v07.i2.01.
 - [9] J. Sun and A. Binder, "Comparison of deep learning architectures for H&E histopathology images," *2017 IEEE Conf. Big Data Anal. ICBDA 2017*, vol. 2018-Janua, pp. 43–48, 2018, doi: 10.1109/ICBDAA.2017.8284105.