

KLASIFIKASI JENIS BATIK MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBORS (KNN)

Azarin Benita Suciani¹, Gunawansyah², Kiki Maulana Adhinugraha³

^{1,2,3} Teknik Informatika, Universitas Sangga Buana

¹ azarinbenitas@gmail.com

ABSTRACT

Batik is a world heritage that has been recognized by UNESCO, has high philosophical and artistic value and economic value. However, with so many types of batik in Indonesia, many people still have difficulty knowing and distinguishing between these types of batik, which can threaten its sustainability. Lack of public knowledge about the types of batik can lead to reduced appreciation and loss of cultural identity. Therefore, appropriate treatment is needed to overcome this problem. This research aims to build a classification system for types of batik using the K-Nearest Neighbors (KNN) algorithm to help the public recognize and differentiate types of batik, especially Batik Lasem, Batik Sekar Jagad, Batik Tambal, and Batik Truntum. It is hoped that this system can increase public awareness and knowledge about the richness of batik culture and help in preserving Indonesia's cultural heritage. The system testing results show differences in accuracy in each experiment with a dataset of 200 images, consisting of 50 images per type of batik. The highest accuracy was obtained from the third experiment, namely 85%, while the second experiment resulted in an accuracy of 72.5%, and the first experiment obtained an accuracy of 65%

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

ABSTRAK

Batik merupakan salah satu warisan dunia yang telah diakui oleh UNESCO, memiliki nilai filosofis dan seni yang tinggi serta bernilai ekonomis. Namun, dengan banyaknya jenis batik di Indonesia, banyak masyarakat masih kesulitan untuk mengetahui dan membedakan jenis-jenis batik tersebut, yang dapat mengancam kelestariannya. Kurangnya pengetahuan masyarakat tentang jenis-jenis batik dapat menyebabkan berkurangnya apresiasi dan kehilangan identitas budaya. Oleh karena itu, diperlukan penanganan yang tepat untuk mengatasi permasalahan ini. Penelitian ini bertujuan untuk membangun sistem klasifikasi jenis batik menggunakan algoritma K-Nearest Neighbors (KNN) untuk membantu masyarakat mengenali dan membedakan jenis-jenis batik, khususnya Batik Lasem, Batik Sekar Jagad, Batik Tambal, dan Batik Truntum. Diharapkan sistem ini dapat meningkatkan kesadaran dan pengetahuan masyarakat tentang kekayaan budaya batik serta membantu dalam pelestarian warisan budaya Indonesia. Hasil pengujian sistem menunjukkan perbedaan akurasi di setiap percobaan dengan dataset berjumlah 200 gambar, terdiri dari 50 gambar per jenis batik. Akurasi tertinggi diperoleh dari percobaan ketiga, yaitu sebesar 85%, sedangkan percobaan kedua menghasilkan akurasi sebesar 72.5%, dan percobaan pertama memperoleh akurasi sebesar 65%.

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

PENDAHULUAN

Batik merupakan salah satu warisan budaya Indonesia yang telah diakui oleh UNESCO sebagai warisan budaya takbenda sejak tahun 2009 [1]. Batik memiliki nilai filosofis, seni, dan ekonomi yang tinggi, serta mencerminkan identitas budaya Indonesia. Di Indonesia, batik sudah sangat dikenal di masyarakat. Namun, dengan banyaknya jenis batik yang

ada, masih banyak masyarakat yang kesulitan dalam mengenali dan membedakan berbagai jenis batik tersebut. Bahkan perhatian untuk melestarikan batik di Indonesia pada umumnya masih sebatas perlakuan normal memakai dan menggunakan batik saja, padahal sudah seharusnya kita sebagai masyarakat Indonesia untuk melestarikan batik tidak hanya dengan memakai batik saja,

akan tetapi dengan mengetahui jenis - jenis batik yang sangat beragam di Indonesia.

Kurangnya pengetahuan masyarakat akan solusi yang tepat dalam menghadapi masalah kesulitan mengetahui dan membedakan jenis - jenis batik yang sangat beragam di Indonesia juga menjadi faktor kehilangan warisan budaya. Tanpa pengetahuan yang cukup, masyarakat mungkin tidak dapat mengidentifikasi dan menghargai keunikan serta makna dari setiap jenis batik. Hal ini dapat menyebabkan kehilangan pemahaman terhadap nilai warisan budaya yang tertanam dalam setiap motif dan teknik pembuatan batik.

Jika permasalahan diatas tidak ditangani dengan tepat, maka dapat menyebabkan kurangnya kesadaran masyarakat akan kekayaan budaya batik tersebut, dan menjadikan batik kurang diapresiasi dan dihargai sehingga potensi kehilangan identitas budaya akan terjadi. Berdasarkan permasalahan diatas, diperlukan cara penanggulangan yang efektif, sebagai salah satu cara dalam mempertahankan warisan budaya batik di Indonesia.

Penelitian terdahulu telah berhasil membangun sistem klasifikasi motif batik khas Banyuwangi menggunakan algoritma KNN. Hasil penelitian tersebut menunjukkan bahwa sistem mampu menghasilkan tingkat akurasi mencapai 85% [2]. Sementara itu, dalam penelitian Dalam penelitian dengan judul Klasifikasi Motif Batik Karawang Berbasis Citra Digital dengan *Principal*

Component Analysis dan *K-Nearest Neighbors* dihasilkan bahwa, Algoritma *K-Nearest Neighbors (KNN)* dapat diimplementasikan dengan baik dalam melakukan klasifikasi motif batik Karawang, dengan menghasilkan akurasi sebesar 95% yang diperoleh dari nilai $K=3$ [3]. Selanjutnya, penelitian yang berbeda dengan judul *Komparasi Metode Klasifikasi Batik Menggunakan Neural Network Dan K-Nearest Neighbors Berbasis Ekstraksi Fitur Tekstur*, mampu menghasilkan nilai akurasi tertinggi yaitu sebesar 96,00% pada parameter $K=1$ [4]. Berdasarkan penelitian diatas, dapat memberikan dasar yang kuat untuk menerapkan metode klasifikasi batik menggunakan algoritma KNN. Namun, belum banyak penelitian yang mengkaji klasifikasi batik dengan fokus pada jenis - jenis batik tertentu di lokasi yang berbeda, seperti Rumah Batik Komar. Oleh karena itu, penelitian ini bertujuan untuk menerapkan algoritma KNN pada dataset yang spesifik dan berbeda dari penelitian sebelumnya.

Dalam penelitian ini, diharapkan dapat membantu masyarakat dalam mempermudah untuk mengenali dan membedakan jenis - jenis batik yang ada di Rumah Batik Komar. Dari sudut pandang ilmiah, penelitian ini memberikan kontribusi dengan menerapkan algoritma KNN untuk membangun sistem klasifikasi yang dapat digunakan sebagai referensi untuk pengembangan aplikasi serupa di masa depan, khususnya dalam bidang klasifikasi citra.

Kebaruan dari penelitian ini terletak pada jumlah dataset yang digunakan, serta fokus pada jenis Batik Lasem, Batik Sekar Jagad, Batik Tambal, dan Batik Truntum yang terdapat di Rumah Batik Komar. Penelitian ini menghasilkan tingkat akurasi yang berbeda serta konteks yang lebih spesifik, sehingga membedakannya dari penelitian sebelumnya. Selain itu dalam penelitian ini juga menyediakan fitur yang menjelaskan jenis batik yang diklasifikasikan, yang dapat menambah wawasan pengguna mengenai batik tersebut. Hal ini menjadi pembeda utama dengan penelitian terdahulu.

Tujuan utama penelitian ini adalah untuk membangun sistem klasifikasi jenis batik di Rumah Batik Komar serta untuk menentukan metode terbaik yang digunakan dalam membangun sistem klasifikasi jenis batik tersebut.

METODE

Penelitian ini dimulai dengan pengumpulan data berupa gambar batik yang diperoleh dari Rumah Batik Komar. Dataset yang digunakan terdiri dari 200 gambar, yang dibagi menjadi empat jenis batik, yaitu Batik Lasem, Batik Sekar Jagad, Batik Tambal, dan Batik Truntum. Masing-masing jenis batik diwakili oleh 50 gambar. Penggunaan dataset yang beragam ini bertujuan untuk menguji kemampuan sistem dalam mengklasifikasikan batik berdasarkan motif yang berbeda, sehingga dapat memberikan hasil yang akurat dan relevan.

Metode yang diterapkan dalam penelitian ini adalah menggunakan algoritma *K-Nearest Neighbors* (KNN). Algoritma KNN merupakan salah satu metode dalam *machine learning* yang termasuk ke dalam *supervised learning*, digunakan untuk klasifikasi berdasarkan kedekatan jarak suatu data dengan data yang lain [5]. Algoritma ini bekerja dengan membandingkan jarak antara data yang akan diklasifikasikan dengan data yang sudah diketahui kelasnya [6]. Proses ini dilakukan dengan mengukur jarak antara data uji dan data latih menggunakan metode jarak *Euclidean*. Algoritma K-NN memiliki tujuan untuk mengklasifikasi objek baru berdasarkan atribut dan *train-ing samples*, berdasarkan mayoritas dari kategori pada K-NN [7]. Penentuan nilai parameter K signifikan berpengaruh terhadap kinerja algoritma KNN. Pengukuran jarak adalah aspek penting dalam algoritma KNN. Jarak ini menggambarkan tingkat kemiripan antara data latih dan data uji. Semakin besar jarak, semakin kecil kemiripan antara keduanya, sedangkan semakin kecil jarak, semakin tinggi kemiripan. Biasanya, untuk menentukan jarak antara data, digunakan rumus jarak *Euclidean Distance* [8]. Berikut adalah rumus *Euclidean Distance*.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots (1)$$

Keterangan:

d(x,y) = Jarak kedekatan antara data x ke data y

xi = Data testing (data uji) ke i

y_i = Data training (data latih) ke i

N = Jumlah atribut 1 sampai n

Penentuan nilai parameter K signifikan berpengaruh terhadap kinerja algoritma KNN. Nilai K yang kecil dapat menyebabkan model terlalu peka terhadap noise, sementara nilai K yang besar dapat menyebabkan model kehilangan kemampuan untuk menangkap pola-pola lokal dalam data [9]. Dan pada prinsipnya KNN memiliki dua parameter yang penting yaitu jumlah tetangga (K) dan cara mengukur jarak antar titik data [10].

HASIL DAN PEMBAHASAN

Pengujian sistem *K-Nearest Neighbors (KNN)* dilakukan dengan melalui tiga skenario pengujian untuk mengevaluasi pengaruh pembagian data latih dan data uji terhadap akurasi model. Hasil penelitian menunjukkan bahwa semakin besar proporsi data latih, semakin tinggi pula akurasi model. Pada skenario pertama, 70% data digunakan untuk pelatihan dan 30% untuk pengujian serta nilai $K=15$, menghasilkan akurasi model mencapai 65%. Skenario kedua, dengan 80% data latih dan 20% data uji serta nilai $K=22$, menghasilkan akurasi sebesar 72,5%. Pada skenario ketiga, dengan 90% data latih dan 10% data uji serta nilai $K=26$, model mencapai akurasi tertinggi, yaitu 85%. Hal ini menunjukkan bahwa penambahan data latih dan pemilihan nilai K yang optimal dapat meningkatkan performa model KNN.

Tabel 1: Hasil Skenario Pengujian

No	Data Latih	Data Uji	Nilai K Optimal	Akurasi
1	70%	30%	15	65%
2	80%	20%	22	72.5%
3	90%	10%	26	85%

Secara detail, performa klasifikasi setiap jenis batik bervariasi. Secara keseluruhan, hasil ini mengindikasikan bahwa akurasi model KNN cenderung membaik seiring dengan meningkatnya proporsi data latih dan pemilihan nilai K yang lebih optimal. Untuk memperjelas hasil skenario penelitian tersebut, maka pada tabel 2 sampai tabel 4 akan dijelaskan hasil training dan testing pada skenario penelitian yaitu sebagai berikut.

Tabel 2: Training dan Testing pada Skenario Penelitian 1

Jenis Batik	Precision	Recall	F1-Score	Accuracy
Batik Lasem	0.88	0.58	0.70	
Batik Sekar Jagad	0.78	0.50	0.61	
Batik Tambal	0.50	0.87	0.63	65%
Batik Truntum	0.71	0.63	0.67	

Berdasarkan tabel diatas, Batik Lasem menunjukkan performa yang cukup baik namun tidak optimal. Selanjutnya Batik Sekar Jagad menunjukkan bahwa model kurang efektif dalam mendeteksi semua gambar Batik Sekar Jagad. Sementara itu, Batik Tambal menunjukkan kemampuan deteksi yang baik namun akurasi yang kurang. Sedangkan Batik Truntum menunjukkan performa yang seimbang.

Secara keseluruhan, meskipun akurasi model 65%, ada variasi signifikan dalam kemampuan klasifikasi antar jenis batik, menunjukkan perlunya perbaikan untuk keseimbangan *Precision* dan *Recall*.

Tabel 3: Training dan Testing pada Skenario Penelitian 2

Jenis Batik	Precision	Recall	F1-Score	Accuracy
Batik Lasem	0.43	0.50	0.46	72.5%
Batik Sekar Jagad	0.80	0.62	0.70	
Batik Tambal	0.73	0.80	0.76	
Batik Truntum	0.83	0.91	0.87	

Pada tabel diatas, Batik Lasem menunjukkan performa yang kurang memuaskan. Sementara itu, Batik Sekar Jagad menunjukkan akurasi yang baik tetapi tidak optimal dalam mendeteksi semua gambar. Selanjutnya Batik Tambal menunjukkan keseimbangan yang baik antara akurasi dan deteksi. Serta Batik Truntum menandakan performa terbaik dalam identifikasi dan deteksi.

Secara keseluruhan, hasil ini menunjukkan bahwa model KNN bekerja lebih baik untuk beberapa jenis batik dibandingkan yang lain, dengan variasi signifikan dalam efektivitas klasifikasi.

Tabel 4: Training dan Testing pada Skenario Penelitian 3

Jenis Batik	Precision	Recall	F1-Score	Accuracy
Batik Lasem	1.00	0.25	0.40	

Batik Sekar Jagad	0.71	1.00	0.83	8.5%
Batik Tambal	1.00	1.00	1.00	
Batik Truntum	0.86	1.00	0.92	

Berdasarkan tabel diatas, bahwa setiap gambar yang diidentifikasi sebagai Batik Lasem benar-benar adalah Batik Lasem, tetapi hanya menunjukkan 25% dari gambar Batik Lasem yang berhasil terdeteksi. Serta performa klasifikasi untuk Batik Lasem kurang memuaskan. Selanjutnya Batik Sekar Jagad menunjukkan akurasi yang baik dalam identifikasi dan deteksi jenis batik ini. Kemudian Batik Tambal mencapai hasil sempurna yang menandakan klasifikasi yang sangat akurat dan efektif. Sementara itu, Batik Truntum menunjukkan performa yang sangat baik dalam identifikasi dan deteksi, meskipun akurasi keseluruhan model sangat rendah.

Hasil klasifikasi untuk ketiga skenario menunjukkan bagaimana model *K-Nearest Neighbors (KNN)* mengidentifikasi berbagai jenis batik dengan detail yang berbeda. Setelah dilakukan uji coba terhadap ketiga skenario, maka skenario yang terpilih yaitu skenario penelitian ketiga.

Pada skenario penelitian ketiga dengan pembagian data latih 90% dan data uji 10%, menghasilkan nilai $K=26$, model *K-Nearest Neighbors (KNN)* mencapai akurasi keseluruhan yang tinggi dari skenario penelitian yang lainnya yaitu sebesar 85%. Batik Lasem menunjukkan *Precision* sempurna (1.00) tetapi *Recall*-nya hanya 0.25,

artinya meskipun setiap gambar yang teridentifikasi sebagai Batik Lasem benar-benar Batik Lasem, model hanya mendeteksi 25% dari gambar Batik Lasem yang ada.

Sedangkan Batik Sekar Jagad memiliki akurasi yang baik dengan *F1-Score* 0.83, menandakan deteksi yang efektif. Sementara itu, Batik Tambal menunjukkan performa optimal dengan *Precision* dan *Recall* sempurna, mencapai *F1-Score* 1.00, menunjukkan klasifikasi yang sangat akurat. Sedangkan Batik Truntum juga memiliki *F1-Score* tinggi 0.92, menunjukkan performa deteksi yang sangat baik.

Hasil klasifikasi dari ketiga pengujian menunjukkan bahwa peningkatan ukuran data latih secara signifikan meningkatkan akurasi, *Precision*, *Recall*, dan *F1-score* model *K-Nearest Neighbors* (KNN), terutama dalam mengenali kelas-kelas batik yang lebih umum. Namun, ukuran data uji yang terlalu kecil menurunkan efektivitas model dalam mendeteksi kelas yang jarang muncul, dengan *precision* tinggi tetapi *recall* rendah. Oleh karena itu, keseimbangan antara data latih dan uji sangat penting agar model tidak hanya akurat pada kelas dominan, tetapi juga mampu mendeteksi kelas yang kurang umum.

SIMPULAN

Penelitian ini berhasil menghasilkan sebuah sistem yang mampu mengklasifikasikan berbagai jenis batik yang terdapat di Rumah Batik Komar. Sistem ini menggunakan metode *K-Nearest Neighbors* (KNN) sebagai

pendekatan utama dalam proses klasifikasi. Berdasarkan hasil pengujian, metode KNN terbukti efektif dan menghasilkan tingkat akurasi tertinggi sebesar 85%. Hal ini menunjukkan bahwa sistem yang dibangun mampu mengklasifikasikan jenis batik dengan tingkat akurasi yang baik, sehingga dapat mendukung upaya pelestarian budaya batik melalui teknologi.

Dalam penelitian ini, terdapat beberapa kesimpulan yang dapat diambil. Penelitian berhasil menghasilkan sebuah sistem yang mampu mengklasifikasikan berbagai jenis batik yang terdapat di Rumah Batik Komar. Sistem ini menggunakan metode *K-Nearest Neighbors* (KNN) sebagai pendekatan utama dalam proses klasifikasi. Berdasarkan hasil pengujian, metode KNN terbukti efektif dan menghasilkan tingkat akurasi tertinggi sebesar 85%. Hal ini menunjukkan bahwa sistem yang dibangun mampu mengklasifikasikan jenis batik dengan tingkat akurasi yang baik, sehingga dapat mendukung upaya pelestarian budaya batik melalui teknologi.

DAFTAR PUSTAKA

- [1] Supriono Primus, *Ensiklopedia: the heritage of Batik : identitas pemersatu kebanggaan bangsa*. 2016. [Online]. Available: https://books.google.co.id/books?id=P09AswEACAAJ&dq=ensiklopedia+batik&hl=id&newbks=1&newbks_redir=0&sa=X&ved=2ahUKEwjD9cLnuGDAXVzxTgGH9ECWUQ6AF6BAgEEAE
- [2] T. Harlina and E. Handayani, "Klasifikasi motif batik banyuwangi menggunakan metode k-nearest neighbor (k-nn) berbasis android," vol.

- 07, pp. 82–96, 2022.
- [3] A. P. B. Salsabila, C. Rozikin, and R. I. Adam, “Klasifikasi Motif Batik Karawang Berbasis Citra Digital dengan Principal Component Analysis dan K-Nearest Neighbor,” *J. Sist. dan Teknol. Inf.*, vol. 11, no. 1, p. 20, 2023, doi: 10.26418/justin.v11i1.46936.
- [4] B. Zaman, A. Rifai, and M. B. Hanif, “Komparasi Metode Klasifikasi Batik Menggunakan Neural Network Dan K-Nearest Neighbor Berbasis Ekstraksi Fitur Tekstur,” *J. Inf. Syst. Informatics*, vol. 3, no. 4, pp. 582–595, 2021, doi: 10.51519/journalisi.v3i4.213.
- [5] M. Lestari, “Penerapan Algoritma Klasifikasi Nearest Neighbor (K-NN) untuk Mendeteksi Penyakit Jantung,” *Fakt. Exacta*, vol. 7, no. September 2010, pp. 366–371, 2014.
- [6] P. Sharma, *Simplified Machine Learning: The essential bulding blocks for Machine Learning expertise*. 2024.
- [7] F. Liantoni, “Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode K-Nearest Neighbor,” *J. Ultim.*, vol. 7, no. 2, pp. 98–104, 2016, doi: 10.31937/ti.v7i2.356.
- [8] A. A. Permana *et al.*, *MACHINE LEARNING*. 2023. [Online]. Available: https://books.google.co.id/books?id=1_muEAAAQBAJ&printsec=frontcover&hl=id&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false
- [9] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*. 2013. [Online]. Available: https://www.stat.berkeley.edu/users/rabbbee/s154/ISLR_First_Printing.pdf
- [10] A. C. Muller and S. Guido, *Introduction to Machine Learning with Phyton*. 2016. [Online]. Available: https://books.google.co.id/books?id=1-4IDQAAQBAJ&printsec=frontcover&redir_esc=y#v=onepage&q&f=false