

ANALISIS TEKSTUR PADA CITRA MOTIF BATIK UNTUK KLASIFIKASI MENGGUNAKAN K-NN

Kristian Adi Nugraha¹ Widi Hapsari² Nugroho Agus Haryono³
adinugraha@ti.ukdw.ac.id widi@ti.ukdw.ac.id nugroho@ti.ukdw.ac.id

Abstract

Indonesian's Batik is one of culture heritage that recognized around the world. Batik has many variations of pattern based on their region. In this research, Batik would be used as subject for texture feature extraction. The value of this feature extraction would be used for classification using K-Nearest Neighbor (K-NN) method. Texture Feature Extraction components that used in this research were Entropy, Correlation, Homogeneity, and Energy. This research will investigate which component would give dominant effect for Batik's pattern recognition. Batik pattern used in this research is pattern from Yogyakarta region. There are four patterns namely Ceplok, Parang, Semen, and Nitik. The result showed that there was no component from Texture Feature Extraction that gave dominant effect (average = 53,96%). Component with the highest value of accuracy is Correlation with a percentage of 55,83%. Whereas for K-NN classification, the best accuracy is 60% for K = 5.

Kata Kunci: Klasifikasi K-NN, Ekstraksi Fitur Tekstur, Batik

1. PENDAHULUAN

Klasifikasi motif batik bisa dilakukan secara manual maupun secara komputasi digital menggunakan metode pengolahan citra digital. Proses klasifikasi citra digital dapat dilakukan berdasarkan fitur dari masing-masing citra. Klasifikasi motif batik yang baik dapat dimanfaatkan untuk pengenalan dan bisa juga untuk mempercepat proses pencarian citra batik. Proses klasifikasi citra dapat dilakukan dengan menggunakan fitur-fitur citra seperti warna, bentuk, dan tekstur citra. Setiap citra dalam masing-masing kelas diidentifikasi fitur-fiturnya yang membedakan dari kelas citra yang lain {(Moertini & Sitohang, 2005), (Mouine, Yahiaoui, & Verroust-Blondet, 2013), (Kamavisdar, Saluja, & Agrawal, 2013)}. Pada penelitian ini akan digunakan fitur tekstur untuk klasifikasi citra Batik. Komponen fitur tekstur yang digunakan adalah *Entropy*, *Correlation*, *Homogeneity*, dan *Energy*. Tujuan utama penelitian ini adalah untuk mengetahui fitur mana yang memberikan pengaruh paling besar dalam proses klasifikasi citra Batik. Setelah memperoleh fitur tekstur dari masing-masing citra Batik, maka proses selanjutnya adalah dengan melakukan klasifikasi dengan metode K-NN. K-NN termasuk kategori *supervised learning*, dimana terdapat *dataset* yang harus diberikan terlebih dahulu untuk dilatih sebelum dilakukan proses klasifikasi itu sendiri.

Materi penelitian berupa kumpulan data motif batik beserta dengan nama-nama kelasnya diambil dari buku motif batik karya (Kusrianto, 2013), buku karya (Ramadhan, 2013), dan buku motif batik Yogya yang disusun oleh Dinas Perindustrian Perdagangan dan Koperasi Provinsi Daerah Istimewa Yogyakarta (Dinas Perindustrian Perdagangan dan Koperasi DIY, 2007). Sedangkan data uji diambil dari sumber yang sama yang belum dimasukkan ke dalam kumpulan data. Jumlah *dataset* yang digunakan adalah 20 untuk masing-masing jenis motif,

¹ Program Studi Teknik Informatika Fakultas Teknologi Informasi UKDW Yogyakarta

² Program Studi Teknik Informatika Fakultas Teknologi Informasi UKDW Yogyakarta

³ Program Studi Teknik Informatika Fakultas Teknologi Informasi UKDW Yogyakarta

sehingga total keseluruhan terdapat 80 buah data. Sedangkan jumlah data yang digunakan untuk pengujian adalah sebanyak 10 data untuk masing-masing jenis motif, sehingga total berjumlah 40 data.

2. EKSTRAKSI FITUR TEKSTUR CITRA DIGITAL

Tekstur merupakan elemen penting dari persepsi visual yang dapat digunakan untuk memisahkan daerah menarik dari sebuah image. Tekstur dapat memuat informasi-informasi penting tentang struktur dan relasinya terhadap sekitarnya. Pengertian tekstur adalah keteraturan pola-pola tertentu yang terbentuk dari susunan piksel-piksel dalam citra digital. Salah satu bagian penting dalam analisis tekstur adalah menggunakan matriks pasangan intensitas (*Gray Level Co-occurrence Matrix/GLCM*) yang merupakan matriks keterkaitan dua dimensi. Matriks pasangan intensitas adalah suatu matriks yang menggambarkan frekuensi munculnya pasangan dua piksel dengan intensitas tertentu dalam jarak dan arah tertentu dalam citra. GLCM dapat dihitung menggunakan beberapa arah offset spasial yaitu 0 derajat, 45 derajat, 90 derajat, dan 135 derajat. Dari matriks GLCM yang diperoleh, ciri tekstur dari citra yang digunakan dalam penelitian ini bisa diukur dengan menggunakan fitur-fitur berikut ini (Gonzalez & Woods, 2008).

1.1. Entropy

Entropy, digunakan untuk mengukur keteracakan dari distribusi intensitas dengan menggunakan rumus persamaan [1].

$$entropy = - \sum_{i=1}^K \sum_{j=2}^K p(i_1, i_2) \log p(i_1, i_2) \quad [1]$$

1.2. Correlation

Correlation, digunakan untuk mengukur korelasi antara sebuah piksel terhadap piksel tetangga dari keseluruhan citra dengan menggunakan rumus persamaan [2] sampai dengan persamaan [5].

$$correlation = \sum_{i=1}^K \sum_{j=1}^K \frac{(i - m_r) - (j - m_c) P_{ij}}{\sigma_r \sigma_c} \quad [2]$$

Keterangan:

$$p(i_1, i_2) = \text{nilai probabilitas matriks GLCM} \quad [3]$$

$$m_r = \sum_{i=1}^K i \sum_{j=1}^K P_{ij} \quad [4]$$

$$m_c = \sum_{j=1}^K j \sum_{i=1}^K P_{ij} \quad [5]$$

1.3. Homogeneity

Homogeneity, digunakan untuk mengukur kehomogenan variasi intensitas dalam citra dengan menggunakan rumus persamaan [6].

$$homogeneity = \sum_{i_1} \sum_{i_2} \frac{p(i_1, i_2)}{1 + |i_1 - i_2|} \quad [6]$$

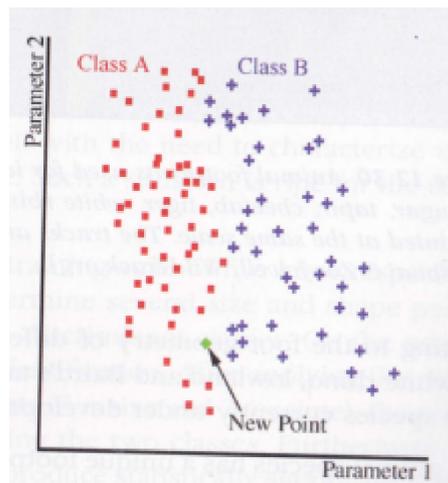
1.4. Energy

Energy, digunakan untuk mengukur konsentrasi pasangan intensitas pada matriks co-occurrence dengan menggunakan rumus persamaan [7].

$$energy = \sum_{i_1} \sum_{i_2} p^2(i_1, i_2) \quad [7]$$

3. KLASIFIKASI K-NN

Terdapat beberapa metode yang dapat digunakan untuk menentukan klasifikasi dari objek-objek. Salah satu metode yang paling banyak digunakan adalah K-Nearest Neighbor (K-NN). Metode klasifikasi K-NN termasuk dalam model *supervised learning* yang mana database populasi yang dipakai sudah diidentifikasi terlebih dahulu ke dalam kelas tertentu. Metode ini bekerja dengan melakukan pencarian sebanyak k-objek pada database yang memiliki ukuran terdekat dengan objek baru yang diuji. Ukuran kedekatan yang dipakai bergantung dari fitur yang dipakai. Gambar 1 memberikan contoh klasifikasi K-NN titik baru (bentuk *diamond* warna hijau) pada dua kelas yang ada menggunakan dua parameter dengan k = 5. Tiga dari lima titik terdekat terhadap titik baru berada dalam kelas B (bentuk *plus* warna biru), sehingga titik baru bisa dimasukkan dalam kelas B. (Russ, 2011)

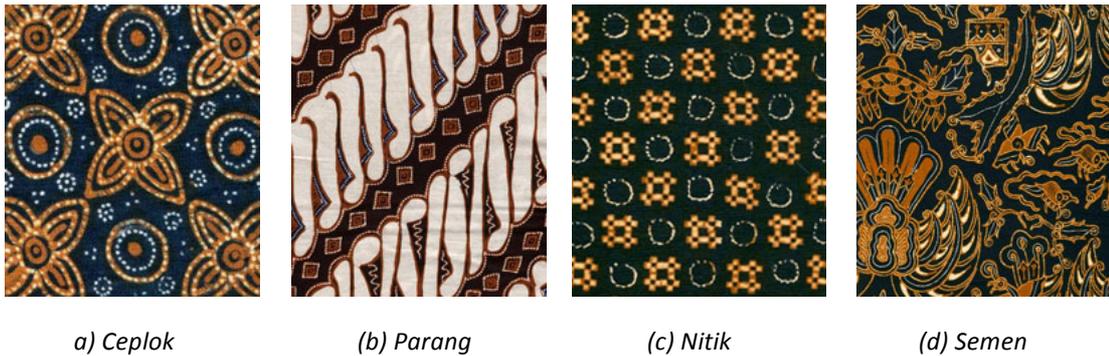


Gambar 1. Klasifikasi K-NN: tiga dari tetangga terdekat titik baru berada dalam kelas B, sehingga titik tersebut masuk dalam kelas B. (Russ, 2011)

Dalam penelitian ini, proses klasifikasi diawali dengan identifikasi database citra motif batik ke dalam empat kelas motif batik, yaitu Ceplok, Parang, Nitik, dan Semen seperti pada Gambar 2.

Dari masing-masing citra dalam database disimpan vector-vector fiturnya. Setelah semua citra dalam database diidentifikasi ke dalam masing-masing kelas, maka proses berikutnya adalah melakukan pengujian terhadap citra batik baru untuk diklasifikasi. Dalam penelitian ini diambil nilai-nilai k, yaitu: 5, 10, dan 15. Dari setiap nilai k akan dihitung keakuratan proses klasifikasi dengan menggunakan rumus akurasi seperti pada persamaan [8].

$$akurasi = \frac{jumlah\ data\ uji\ yang\ benar}{jumlah\ total\ data\ uji} \times 100\% \quad [8]$$



Gambar 2. Empat Kelas Motif Batik yang dipakai (Dinas Perindustrian Perdagangan dan Koperasi DIY, 2007)

Jumlah data uji yang akan dimasukkan sebagai sampel uji sebanyak 40 citra dari sumber yang sama yang belum dimasukkan ke dalam database dengan masing-masing kelas 10 citra. Proses perhitungan jarak antar citra di dalam klasifikasi k-NN untuk fitur tekstur menggunakan nilai absolut dari selisih nilai parameter fitur tekstur masing-masing citra seperti dalam persamaan [9].

$$\delta = |p(c_1) - p(c_2)| \quad [9]$$

4. HASIL PENELITIAN PADA BATIK

Sistem klasifikasi motif Batik yang dibuat oleh penulis dibangun dengan menggunakan MATLAB. Hampir seluruh fungsi yang dibutuhkan untuk pembangunan sistem telah disediakan oleh MATLAB, sehingga penulis cukup memanggil fungsi-fungsi tersebut dari *library* milik MATLAB. Klasifikasi dilakukan dengan menghitung selisih jarak yang diperoleh dari nilai ekstraksi yang telah dinormalisasi, sehingga nilai jarak antar ekstraksi fitur tersebut berkisar antara 0 hingga 1. Kemudian jarak dari keempat komponen ekstraksi fitur dikombinasikan menjadi satu, untuk komponen utama yang akan diujikan diberi bobot sebesar 70%, sedangkan tiga komponen sisanya masing-masing sebesar 10%. Berdasarkan hasil pengujian yang dilakukan oleh penulis, didapatkan hasil seperti yang ditunjukkan pada tabel 1.

Tabel 1
Hasil Pengujian Klasifikasi menggunakan Ekstraksi Fitur

Nama	K = 5				K = 10				K = 15				Total	%
	Ent r	Cor r	Ene r	Hom o	Ent r	Cor r	Ene r	Hom o	Ent r	Cor r	Ene r	Hom o		
Ceplo k	3	6	4	4	4	5	4	5	4	4	7	4	54	45
Parang	6	5	6	4	6	6	6	5	6	4	4	6	64	53.3
Semen	6	10	6	5	5	6	5	5	5	6	5	5	69	57.5
Nitik	8	7	8	8	5	3	8	7	5	5	3	5	72	60
Total	23	28	24	21	20	20	23	22	20	19	19	20	259	53.9 6

Pada tabel 1 diujikan sebanyak 10 data uji untuk masing-masing motif Batik. Pengujian klasifikasi dilakukan sebanyak 3 kali untuk masing-masing komponen ekstraksi fitur, yaitu untuk nilai K = 5, K = 10, dan K = 15. Pada tabel 1 dapat dilihat bahwa motif Batik yang memiliki tingkat akurasi paling tinggi adalah motif Nitik, yaitu sebesar 60%. Sedangkan rata-rata keseluruhan tingkat akurasi adalah sebesar 53,96%. Untuk pengaruh masing-masing

komponen ekstraksi fitur tekstur dapat dilihat pada tabel 2. Pada tabel tersebut dapat dilihat bahwa seluruh komponen ekstraksi fitur (*entropy, correlation, homogeneity, & energy*) memiliki pengaruh yang tidak jauh berbeda, yaitu antara 52,5% hingga 55,83%, dimana nilai tertinggi diperoleh oleh komponen *Correlation*.

Tabel 2
Kesimpulan Hasil Ekstraksi Fitur

Nama	Jumlah	%
Entropy	63	52,5
Correlation	67	55,83
Homogeneity	66	55
Energy	63	52,5

Sedangkan untuk nilai K pada proses klasifikasi yang memiliki tingkat akurasi paling tinggi dapat dilihat pada tabel 3 yaitu adalah untuk K = 5 dengan persentase sebesar 60%.

Tabel 3
Kesimpulan Hasil Ekstraksi Fitur

K	Jumlah	%
5	96	60
10	85	53,13
15	78	48,75

Pada pengujian klasifikasi yang dilakukan oleh penulis, terdapat banyak motif yang dikenali tidak sesuai dengan jenis sebenarnya. Menurut pengamatan penulis, proses klasifikasi yang tidak sesuai dengan motif yang sebenarnya disebabkan karena antar motif Batik memiliki pola tekstur yang hampir sama. Misalnya untuk motif Ceplok seringkali justru dikenali sebagai motif Semen, demikian juga dengan motif Semen yang sering dikenali sebagai Ceplok.

5. KESIMPULAN

Berdasarkan hasil analisa, dari semua nilai K yang telah diuji, nilai K dengan tingkat akurasi paling baik adalah K = 5. Semakin tinggi nilai K justru memberikan performa yang semakin menurun. Secara keseluruhan, motif Nitik adalah motif yang memiliki tingkat akurasi pengenalan yang paling baik dibanding motif lainnya, yaitu dengan persentasi sebesar 60%. Sedangkan untuk pengujian pengaruh komponen ekstraksi fitur, ternyata tidak ada yang memberikan pengaruh paling dominan yaitu dengan nilai antara 52,5% s/d 55,83%, dengan nilai tertinggi diperoleh oleh komponen Homogeneity. Terdapat cukup banyak motif Batik yang kurang tepat dikenali pada saat klasifikasi. Hal tersebut disebabkan karena antar jenis motif Batik bisa jadi memiliki kemiripan tekstur. Karena itu perlu diteliti lebih lanjut agar dapat membedakan motif-motif Batik yang berbeda jenis namun memiliki kemiripan tekstur, salah satunya dengan cara melihat bentuk dari motif Batik tersebut.

Daftar Pustaka

- Dinas Perindustrian Perdagangan dan Koperasi DIY. (2007). Buku Motif Batik Yogya Ceplok (1 ed.). Yogyakarta, Indonesia: Pena Persada Desktop Publishing.
- Dinas Perindustrian Perdagangan dan Koperasi DIY. (2007). Buku Motif Batik Yogya Nitik (1 ed.). Yogyakarta, Indonesia: Pena Persada Desktop Publishing.
- Dinas Perindustrian Perdagangan dan Koperasi DIY. (2007). Buku Motif Batik Yogya Parang dan Lereng (1 ed.). Yogyakarta, Indonesia: Pena Persada Desktop Publishing.

- Dinas Perindustrian Perdagangan dan Koperasi DIY. (2007). Buku Motif Batik Yogya Semen (1 ed.). Yogyakarta, Indonesia: Pena Persada Desktop Publishing.*
- Gonzalez, R. C., & Woods, R. E. (2008). Digital Image Processing. New Jersey: Pearson Education, Inc.*
- Kamavisdar, P., Saluja, S., & Agrawal, S. (2013). Survey on Image Classification Approaches and Techniques. International Journal of Advanced Research in Computer and Communication Engineering Volume 9, August, 2009, 1005-1009.*
- Kusrianto, A. (2013). Batik Filosofi, Motif dan Kegunaan. Yogyakarta: Andi.*
- Moertini, V. S., & Sitohang, B. (2005). Algorithms of Clustering and Classifying Batik Images Based on Color, Contrast and Motif. PROC. ITB Eng. Science Vol. 37 B, No. 2, 2005, 141-160.*
- Mouine, S., Yahiaoui, I., & Verroust-Blondet, A. (2013). A Shape-based Approach for Leaf Classification using Multiscale Traingular Representation. ACM International Conference on Multimedia Retrieval (pp. 127 - 134). Dallas, Texas: ACM.*
- Ramadhan, I. (2013). Cerita Batik. Tangerang: Literati.*
- Russ, J. C. (2011). The Image Processing Handbook. India: CRC Press Taylor & Francis Group.*