

## **BAB II**

### **TINJAUAN PUSTAKA**

Dalam bab ini menjelaskan tentang landasan teori yang bertujuan untuk menyusun konsep berkaitan dengan penelitian, yang terdiri dari penjelasan studi-studi sebelumnya dan dasar-dasar teori yang digunakan.

#### **2.1. Penelitian Sebelumnya**

Beberapa penelitian sebelumnya yang sejenis tentang ekstraksi citra dan klasifikasi telah banyak dilakukan dengan berbagai metode. Pada Penelitian yang dilakukan oleh (Jatmoko & Sinaga, 2019) yang melakukan penelitian untuk ekstraksi fitur *Gray Level Co-Occurrence Matrix* (GLCM) pada *K-Nearest Neighbours* (KNN) dalam mengklasifikasi motif batik. Klasifikasi dilakukan sebanyak 2 kali, perbedaannya terletak pada penggunaan jumlah fitur pada saat ekstraksi fitur tekstur. Yang pertama menggunakan 4 fitur yaitu *energy*, *contrast*, *correlation*, dan *homogeneity*. Klasifikasi selanjutnya menggunakan 5 fitur yang memberikan informasi nilai-nilai *energy*, *contrast*, *correlation*, *homogeneity*, dan *entropy*. Pada uji coba pertama menghasilkan akurasi tertinggi yaitu 100%, maka sistem layak dilanjutkan untuk menguji data testing. Uji coba kedua diperoleh akurasi 73,33% dan Uji coba ketiga memperoleh hasil akurasi 66,67%, Hasil akurasi terbaik yaitu 100% pada percobaan pertama dimana menggunakan 15 data testing dan 135 data training.

Penelitian yang lainya yaitu analisis tekstur pada citra motif batik untuk klasifikasi menggunakan k-nn (Nugraha dkk, 2014) Pada penelitiannya menggunakan fitur tekstur untuk mengklasifikasi citra Batik. Komponen fitur yang digunakan adalah *Entropy*, *Correlation*, *Homogeneity*, dan *Energy*. Penelitiannya bertujuan untuk mengetahui fitur mana yang memberikan pengaruh paling besar dalam proses klasifikasi citra Batik. Hasil yang diperoleh yaitu nilai akurasi tertinggi 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*.

Penelitian selanjutnya Klasifikasi Motif Citra Batik Yogyakarta Menggunakan Metode *Adaptive Neuro Fuzzy Inference System* (Hardiyanto dkk, 2019), penelitian bertujuan untuk mengklasifikasikan motif citra batik Yogyakarta (menggunakan motif batik Parang Kusumo dan motif batik Kawung). Metode yang digunakan adalah ekstraksi fitur GLCM (*Grey Level Co-Occurrence Matrix*) dan klasifikasi menggunakan metode ANFIS (*Adaptive Neuro Fuzzy Inference System*). Dengan menggunakan 100 data citra uji, Trapezoidal Shaped-Membership Function memperoleh akurasi terbaik yakni 80%, sedangkan *Gaussian Shaped-Membership Function* memperoleh akurasi terendah yakni 77%.

## 2.2. Citra

Citra adalah suatu representasi (gambaran), kemiripan, atau imitasi dari suatu objek. Citra terbagi 2 yaitu ada citra yang bersifat analog dan ada citra yang bersifat digital. Citra analog adalah citra yang bersifat kontinu seperti gambar pada monitor televisi, foto sinar X, hasil CT Scan dll. Sedangkan pada citra digital adalah citra yang dapat diolah oleh komputer( T,Sutoyo, 2009: 9).

Berdasarkan cara penyimpanan atau pembentukannya, citra digital dapat dibagi menjadi dua jenis. Jenis pertama adalah citra digital yang dibentuk oleh kumpulan pixel dalam array dua dimensi. Citra jenis ini disebut citra bitmap atau citra raster. Jenis citra yang kedua adalah citra yang dibentuk oleh fungsi-fungsi geometri dan matematika. Jenis citra ini yang disebut grafik vektor. Citra digital dihasilkan dari citra analog melalui digitalisasi. Digitalisasi citra analog terdiri dari sampling dan quantization. Sampling adalah pembagian citra ke dalam elemen-elemen diskrit (pixel), sedangkan quantization adalah pemberian nilai intensitas warna pada setiap pixel dengan nilai yang berupa bilangan bulat (G.W. Awcock,1996).

### 2.3.1 Jenis–Jenis Citra

Menurut seorang Frank Jefkins dalam Soemirat dan Elvinaro Ardianto (2007:117), memilah citra dalam beberapa jenis, antara lain:

- a. *The mirror image* (cerminan citra), yaitu bagaimana taksiran (citra) manajemen terhadap public eksternal dalam melihat perusahaannya.
- b. *The current image* (citra masih hangat), yaitu citra yang terdapat pada publik eksternal, yang berdasarkan pengalaman atau menyangkut miskinnya informasi dan pemahanman publik. eksternal. Citra ini bisa saja berbenturan dengan mirror image.
- c. *The wish image* (citra yang diinginkan), yaitu manajemen mengharapakan perolehan prestasi tertentu. Citra ini diaplikasikan untuk sesuatu yang baru sebelum dipublikasikan secara eksternal untuk memperoleh informasi secara lengkap.
- d. *The multiple image* (citra yang berlapis), ialah sekelompok individu, kantor cabang atau perwakilan perusahaan lainnya dapat membentuk citra yang belum sesuai dengan kesamaan citra seluruh organisasi atau perusahaan.

### 2.3.2 Elemen Citra Digital

Menurut seorang (T Sutoyo, 2009:24) Citra Digital memiliki elemen-elemen sebagai berikut:

- a. Kecerahan (*Brightness*)  
Mengambarkan kekuatan cahaya yang dikeluarkan melalui piksel dari citra sehingga dapat ditangkap oleh sistem penglihatan.
- b. Kontras (*Contrast*)
- c. Merupakan komposisi terang dan gelap dalam sebuah citra, Citra yang baik memiliki komposisi gelap dan terang secara merata.
- d. Kontur (*Countour*)  
Merupakan keadaan yang ditimbulkan oleh alterasi intensitas pada piksel-piksel yang berdekatan. Dengan adanya transisi intensitas inilah yang membuat mata mampu mendeteksi tepi-tepi objek didalam citra.

e. Warna (*Colour*)

Warna sebagai persepsi yang bisa ditangkap sistem visual terhadap panjang gelombang cahaya yang dipantulkan melalui objek.

### 2.3.3 Pengolahan Citra

Pengolahan citra digital adalah sebuah disiplin ilmu yang mempelajari hal-hal yang berkaitan dengan perbaikan kualitas gambar (peningkatan kontras, transformasi warna, restorasi citra), transformasi gambar (rotasi, translasi, skala, transformasi geometrik), melakukan pemilihan citra ciri (*feature images*) yang optimal untuk tujuan analisis, melakukan proses penarikan informasi atau deskripsi objek atau pengenalan objek yang terkandung pada citra, melakukan kompresi atau reduksi data untuk tujuan penyimpanan data, transmisi data, dan waktu proses data. Input dari pengolahan citra adalah citra, sedangkan outputnya adalah citra hasil pengolahan (T, Sutoyo, 2009: 5).

### 2.3.4 Teknik Pengolahan Citra

Menurut (Basuki, 2005:11) ada beberapa teknik dalam pengolahan citra yaitu sebagai berikut:

a. *Image Enhancement*

Suatu proses perbaikan citra dengan cara meningkatkan kualitas citra, baik kontras maupun kecerahan.

b. *Image restoration*

Suatu proses memperbaiki model citra, sehingga menjadi bentuk citra yang sesuai.

c. *Color Image Processing*

Suatu proses yang dikaitkan dengan citra warna, baik itu berupa *image enhancement*, *image restoration*, ataupun yang lainnya.

d. *Wavelet and multiresolution processing*

Suatu proses citra yang menyatakan dalam berapa resolusi.

*e. Image Compression*

Suatu proses yang digunakan untuk mengubah ukuran data dalam suatu citra.

*f. Morphological Processing*

Proses untuk mendapatkan sebuah informasi yang menyatakan deskripsi suatu bentuk dari sebuah citra.

*g. Segmentation*

Proses untuk membedakan atau memisahkan objek-objek di dalam suatu citra, seperti memisahkan objek dengan background.

*h. Object Recognition*

Proses yang dilakukan untuk mengenali suatu objek apa saja yang ada didalam suatu citra.

### **2.3. Citra RGB**

Pada color image ini masing-masing piksel memiliki warna tertentu, warna tersebut adalah merah (Red), hijau (Green) dan biru (Blue). Jika masing-masing warna memiliki range 0-255, maka totalnya adalah  $255^3 = 16.581.375$  (16 K) variasi warna berbeda pada gambar, dimana variasi warna ini cukup untuk gambar apapun. Karena jumlah bit yang diperlukan untuk setiap pixel, gambar tersebut juga disebut gambar-bit warna. Color image ini terdiri dari tiga matriks yang mewakili nilai-nilai merah, hijau dan biru untuk setiap pikselnya (Kusumanto & Tompunu, 2011).

### **2.4. Citra Grayscale**

Merupakan citra yang hanya memiliki warna tingkat keabuan. Penggunaan fitur warna grayscale dikarenakan hanya membutuhkan sedikit informasi pada setiap piksel dibandingkan dengan citra berwarna. Warna abu-abu pada citra grayscale adalah warna R (red), G (green), dan B (blue) yang memiliki intensitas yang sama. Sehingga citra grayscale hanya membutuhkan nilai intensitas tunggal daripada citra berwarna yang membutuhkan tiga intensitas untuk setiap pikselnya. Intensitas dari citra grayscale akan disimpan dalam

8bit integer yang nantinya memberikan 256 kemungkinan yang mana dimulai dari level 0 sampai dengan 255 (0 untuk hitam dan 255 untuk putih dan nilai diantaranya adalah derajat keabuan). Tingkat keabuan atau Grayscale level (Anggraini dkk, 2017).

## 2.5. Klasifikasi

Klasifikasi merupakan sebuah metode yang digunakan untuk mengelompokkan data. Klasifikasi bisa juga diartikan sebagai pengelompokan data atau objek baru kedalam suatu kelas atau kategori berdasarkan variabel-variabel tertentu. Klasifikasi memiliki teknik data mining yang melihat dari kelompok data yang sudah didefinisikan sebelumnya. Atribut ini digunakan sebagai variabel dalam menentukan kelas suatu objek yang baru. Klasifikasi ini memiliki tujuan yaitu menentukan kelas dari suatu objek yang kelasnya belum diketahui dengan akurat.

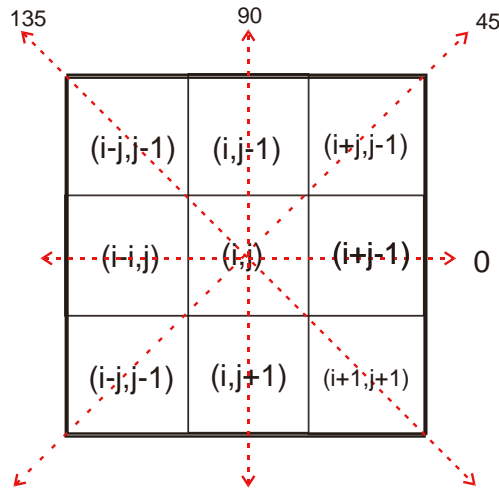
Dalam proses klasifikasi terdiri dari dua fase, yaitu fase learning dan fase testing. Fase learning merupakan sebagian data yang kelas datanya telah diketahui sebelumnya dan dijadikan untuk model yang akan dibangun. Sedangkan fase testing memiliki arti fase model yang sudah terbentuk diuji dengan sebagian data lainnya untuk mengetahui akurasi dari model tersebut. Jika model akurasinya mencukupi, model ini bisa digunakan untuk memprediksi kelas data yang belum diketahui.

## 2.6. Gray Level Co-Occurrence Matrix (GLCM)

*Gray Level Co-Occurrence Matrix* (GLCM) merupakan suatu citra sebagai distribusi dari saling-kemunculan (cooccurrence) nilai piksel pada area tertentu dan dengan ukuran tertentu (Arnia & Munadi, 2018). Metode ini biasanya digunakan dalam pengenalan tekstur, segmentasi citra, analisis warna pada citra, klasifikasi citra, dan pengenalan objek).

Komponen utama dalam GLCM adalah arah dan jarak antara dua piksel. Arah ketetangaan yang mungkin antara dua buah piksel adalah  $0^\circ$ ,  $45^\circ$ ,

90°, 135°, 180°, 225°, 270°, dan 315° seperti pada Gambar 2.1 dibawah ini



**Gambar 2.1 Arah Ketetangaan**

Beberapa ciri atau fitur statistik yang akan digunakan dalam penelitian ini adalah sebagai berikut :

a. *Energy*

Digunakan untuk mengukur homogenitas sebuah citra, Dimana  $P_{i,j}$  menyatakan nilai pada baris  $i$  dan kolom  $j$  pada matriks kookurensi.

$$Energy = \sum_{i=0}^{G-1} | \sum_{j=0}^{G-1} P_{i,j}^2 \dots\dots\dots (2.1)$$

b. *Correlation*

Digunakan untuk menghitung keterkaitan piksel yang memiliki level keabuan  $i$  dengan piksel yang memiliki level keabuan  $j$ .

$$Correlation = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{(i - \mu_i)(j - \mu_j)}{\sigma_i \sigma_j} P_{i,j} \dots\dots\dots (2.2)$$

c. *Homogenity*

Digunakan untuk mengukur homogenitas citra dengan level keabuan sejenis.

$$\text{Homogeneity} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1 + (i - j)^2} P_{i,j} \quad \dots\dots\dots (2.3)$$

*d. Contrast*

Digunakan untuk mengukur variasi pasangan tingkat keabuan dalam sebuah citra.

$$\text{Contrast} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - j)^2 P_{i,j} \quad \dots\dots\dots (2.4)$$

## 2.7. Metode K- Nearest Neighbors

K- Nearest Neighbors adalah satu dari sekian banyaknya metode dimana fungsinya yaitu mengerjakan suatu pengelompokan terhadap suatu objek yang masukannya disesuaikan dari jarak yang mendekati objek itu sendiri. Dari data yang dihasilkan maka akan di proyeksikan ke dalam ruang dimensi paling banyak, dan disitu setiap dimensi mempunyai presentasi berbagai macam konfigurasi dan warna. Dalam ruang nantinya akan dibagi lebih dari satu dengan pengelompokan dan pembenahan. KNN yang merupakan metode dimana ia mempunyai sifat supervised, yang mana hasil data dari query instance akan dikelompokan yang berlandaskan dari banyaknya kategori metode tersebut.

Tujuan dari algoritma ini adalah mengklasifikasikan obyek baru berdasarkan jarak suatu obyek yang akan diklasifikasikan terhadap data contoh. Classifier hanya menggunakan fungsi jarak dari data baru ke data training. Prinsip kerja K-Nearest Neighbors (KNN) adalah mencari jarak terdekat antara data yang akan dievaluasi dengan K tetangga (neighbor) terdekatnya dalam data pelatihan.



Data pelatihan diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi data pelatihan. Sebuah titik pada ruang ini ditandai kelas  $c$ , jika kelas  $c$  merupakan klasifikasi yang paling banyak ditemui pada  $k$  tetangga terdekat titik tersebut. Dekat atau jauhnya tetangga bisa dihitung berdasarkan jarak Euclidean dengan rumus persamaan 2.5 sebagai berikut :

$$d_1 = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \dots\dots\dots (2.5)$$

Keterangan:

$x_1$  = Sampel Data

$x_2$  = Data Uji / Testing

$i$  = Variabel Data

$d$  = Jarak

$p$  = Dimensi Data

Pada fase pembelajaran, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi dari data pembelajaran. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk data test (yang klasifikasinya tidak diketahui). Jarak dari vektor yang baru ini terhadap seluruh vektor data pembelajaran dihitung, dan sejumlah  $k$  yang paling dekat diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak dari titik-titik tersebut.

Nilai  $k$  yang terbaik untuk algoritma ini tergantung pada data. Umumnya, nilai  $k$  yang tinggi akan mengurangi efek noise pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi lebih kabur. Nilai  $k$  yang bagus dapat dipilih dengan optimasi parameter, misalnya dengan menggunakan cross-validation. Kasus khusus di mana klasifikasi diprediksikan berdasarkan data pembelajaran yang paling dekat (dengan kata lain,  $k = 1$ ) disebut algoritma nearest neighbor.

Ketepatan algoritma KNN ini sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi. Ketika jumlah data mendekati tak hingga, algoritma ini menjamin error rate yang tidak lebih dari dua kali Bayes error rate.

## **2.8. Diagram Alir**

Merupakan sebuah diagram dengan simbol yang menggambarkan suatu proses pada suatu sistem secara urut dan runtut, hubungan antar satu tahap ke tahap yang lain, dan intruksi yang tergantung di didalam masing-masing simbol. Diagram alir terbagi menjadi lebih dari satu bagian yaitu sebagai berikut :

### **2.8.1. Alir Sistem**

Merupakan suatu bagian yang menjelaskan tentang proses pekerjaan disebuah sistem. Pada bagian inilah yang menggambarkan arus secara rinci dan menyeluruh

### **2.8.2. Dagram Alir Program**

Merupakan suatu bagian alur yang menggambarkan tahapan-tahapan dalam sebuah program. Hal ini sangat membantu bagi seorang programmer dalam memahami suatu program.

### **2.8.3 Diagram alir Proses**

Diagram alir proses ini banyak digunakan di sektor industri dan juga analisis sistem. Yang memiliki fungsi untuk melihat prosedur dalam suatu proses produksi. Diagram ini juga sering digunakan untuk melihat langkah awal sampai langkah terakhir.

## **2.9. Matlab**

Matlab (*Matrix Laboratory*) adalah suatu program untuk analisis dan komputasi numerik dan merupakan suatu Bahasa pemrograman matematika lanjutan yang dibentuk dengan dasar pemikiran menggunakan sifat dan bentuk matriks. Pada awalnya, program ini merupakan *interface* untuk

koleksi rutin-rutin numeric dari proyek LINPACK dan EISPACK, dan dikembangkan menggunakan bahasa FORTRAN namun sekarang merupakan produk komersial dari perusahaan Mathworks, Inc. yang dalam perkembangan selanjutnya dikembangkan menggunakan bahasa C++ dan assembler (utamanya untuk fungsi-fungsi dasar Matlab).

Matlab telah berkembang menjadi sebuah environment pemrograman yang canggih yang berisi fungsi-fungsi *built-in* untuk melakukan tugas pengolahan sinyal, aljabar linier, dan kalkulasi matematis lainnya. Matlab juga berisi *toolbox* yang berisi fungsi-fungsi tambahan untuk aplikasi khusus. Matlab bersifat *extensible*, dalam arti bahwa seorang pengguna dapat menulis fungsi baru untuk ditambahkan pada *library* ketika fungsi-fungsi built-in yang tersedia tidak dapat melakukan tugas tertentu. Kemampuan pemrograman yang dibutuhkan tidak terlalu sulit bila anda telah memiliki pengalaman dalam pemrograman bahasa lain seperti C++, PASCAL, atau FORTRAN. Matlab merupakan merk software yang dikembangkan oleh Mathworks, Inc yang merupakan software yang paling efisien untuk perhitungan numerik berbasis matriks. Dengan demikian jika di dalam perhitungan kita dapat menformulasikan masalah kedalam format matriks maka Matlab merupakan software terbaik untuk penyelesaiannya. Matlab yang merupakan bahasa pemrograman tingkat tinggi berbasis pada matriks sering digunakan untuk teknik komputasi numerik, untuk menyelesaikan masalah-masalah yang melibatkan operasi matematika elemen, matrik, optimasi, aproksimasi dan lain-lain. Sehingga Matlab banyak digunakan pada: (1) Matematika dan Komputansi, (2) Pengembangan dan Algoritma, (3) Pemrograman modeling, simulasi, dan pembuatan *prototype*, (4) Analisa Data, eksplorasi dan visualisasi, (5) Analisis numerik dan statistic, dan (6) Pengembangan aplikasi teknik.

## 2.10. Batik

Batik merupakan hal yang tidak asing bagi masyarakat Indonesia saat ini. Batik merupakan salah satu warisan nusantara yang unik. Keunikannya

ditunjukkan dengan berbagai macam motif yang memiliki makna tersendiri. Menurut Asti & Ambar B (2011: 1). Dalam Kamus Besar Bahasa Indonesia, batik memiliki arti kain bergambar yang pembuatannya secara khusus dengan menuliskan atau menerakan malam pada kain itu, kemudian pengolahannya diproses dengan cara tertentu.

Perkembangan batik saat ini sangat pesat hingga pada akhirnya menghasilkan motif atau corak yang baru pada setiap daerah yang biasanya menggambarkan potensi atau keunggulan dari daerah tersebut. Seperti halnya pada Motif batik Bojonegoro atau biasa disebut dengan batik Jonegoroan. Motif-motif Batik Jonegoroan yaitu motif Gastro Rinonce (motif kilang minyak dan gas bumi), Jagung Miji Emas (motif jagung), Mliwis Mukti (motif burung legendaris jelmaan Angling Dharma, mliwis putih), Parang Dahono Munggal (motif wisata api abadi, kahyangan api), Parang Jembul Sekar Rinandar (motif hewan sapi), Pari Sumilak (motif padi), Rancak Thengul (motif wayang thengul, khas Bojonegoro), Sata Gondo Wangi (motif tembakau), dan Sekar Jati (motif daun jati).



**Gambar 2.2 Motif Rancak tenggul**



**Gambar 2.3 Motif Pari Sumilak**



**Gambar 2.4 Motif Gatra Rinonce**



**Gambar 2.5 Motif Sekar Jati**