

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Menelusuri informasi menggunakan teknologi pada masa kini, telah dilakukan oleh berbagai macam pihak untuk berbagai kepentingan. Kecanggihan yang dimiliki oleh teknologi, membuat informasi yang tersedia dalam *database* tidak hanya berupa teks, tetapi dapat berupa gambar. Gambar adalah sesuatu yang dapat mewakili suatu objek dengan karakter warna maupun bentuk yang sesuai dengan aslinya (McAndrew, 2004). Gambar dapat mewakili informasi seperti manusia, hewan, tumbuhan, grafik, kegiatan sehari-hari dan sebagainya. Informasi yang disajikan dalam bentuk gambar dapat membantu manusia dalam mengenali bentuk atau suatu tempat yang belum pernah dilihatnya. Dalam perlombaan sebuah potret gambar, keaslian gambar adalah nilai utama dalam perlombaan. Karena keaslian gambar, menunjukkan bahwa potret gambar tersebut bukan milik orang lain atau hasil plagiat. Tetapi bisa saja potret gambar yang dibuat, ternyata sudah pernah dibuat oleh orang lain. Untuk mengatasi hal tersebut, tidak sedikit orang yang melakukan pencarian potret gambar untuk membandingkannya dengan yang terdapat pada internet. Pencarian gambar pada *database* disebut juga sebagai *Image Retrieval*. Pencarian gambar pada *database* dapat dilakukan berdasarkan text dan gambar. Penelusuran gambar berdasarkan text, dinilai tidak efisien, karena jika menggunakan text, penelusur harus menyebutkan karakteristik yang dimiliki gambar tersebut seperti bentuk, warna, tekstur, dan sebagainya. Hal tersebut tentu tidak efisien, karena sistem akan sulit untuk menentukan kata kunci pencarian.

Pada dasarnya komputer tidak dapat melihat gambar sama seperti manusia, tetapi komputer dapat membaca gambar yang telah dikonversi menjadi gambar analog dua dimensi atau disebut citra. *Content-Based Image Retrieval* adalah sebuah *engine* untuk melakukan pencarian citra berdasarkan gambar pada *database*. *Content-Based Image Retrieval* (CBIR) adalah salah satu penelitian pada bidang *computer vision* dalam mencari dan mendeteksi sebuah citra digital dalam *database* (Fahui, Zhang, & Feng, 2003). CBIR Citra Bunga bekerja dengan menggunakan citra sebagai *input query* atau kunci masukan dan menghasilkan citra yang memiliki kesamaan (Patheja P.S., 2012). Dalam merancang CBIR Citra Bunga, terdapat tiga

pendekatan yang umumnya dilakukan, yaitu *preprocessing*, ekstraksi fitur, dan *similarity measure*. Proses ekstraksi fitur adalah proses membaca citra untuk mengetahui nilai-nilai yang dimiliki citra. Sedangkan untuk melakukan perbandingan citra serupa, dilakukan pada proses *similarity measure*.

Penelitian terhadap CBIR Citra Bunga telah berkembang sejak ditemukan pada tahun 1992 (Kato, 1992). Permasalahan terjadi ketika memilih pendekatan ekstraksi fitur yang akan digunakan pada data uji. Pendekatan menggunakan ekstraksi fitur warna akan menghasilkan nilai yang optimal, apabila digunakan untuk mendeteksi warna pada citra. Tetapi pendekatan menggunakan ekstraksi fitur warna tidak akan menghasilkan nilai optimal, apabila digunakan untuk mendeteksi bentuk pada citra. Berdasarkan permasalahan tersebut, pemilihan pendekatan metode ekstraksi fitur, masih menjadi permasalahan yang melatarbelakangi beberapa penelitian pada CBIR Citra Bunga. Untuk mengetahui tingkat keberhasilan dari pendekatan yang dipilih dalam merancang sistem CBIR Citra Bunga, dilakukan perhitungan tingkat ketepatan pada CBIR Citra Bunga. Perhitungan tingkat ketepatan dihitung berdasarkan nilai *precision* yang didapatkan dari pengujian sistem CBIR Citra Bunga. Nilai *precision* pada CBIR Citra Bunga, menunjukkan kinerja sistem dalam menghasilkan dokumen yang relevan. Nilai *precision* dipengaruhi oleh faktor pemilihan metode seperti pemilihan *preprocessing*, ekstraksi fitur dan metode pengukuran tingkat kemiripan. Selain hal tersebut, data uji yang digunakan pada pengujian, mempengaruhi nilai *precision* yang dihasilkan.

Permasalahan pemilihan ekstraksi fitur untuk mendapatkan nilai *precision* yang lebih baik, sejatinya bukanlah kajian yang baru. Sudah banyak peneliti yang melakukan pendekatan berbeda, diantaranya penelitian yang dilakukan oleh Dinda P. Balqis (2006), Hiremath & Pujari (2007), Singha Manimala & K.Hemachandran (2012), dan Singh & Hemachandran (2012). Pada penelitiannya, pendekatan ekstraksi fitur yang digunakan sangat berbeda, sehingga nilai yang dihasilkan berbeda. Tetapi terdapat juga penelitian dengan hasil nilai yang berbeda tetapi pendekatan ekstraksi fitur yang digunakan sama, diantaranya penelitian yang dilakukan oleh Kaur, Banga, & Banga (2013), dan Nirapure & Reddy (2013). Pada penelitian tersebut nilai rata-rata *precision* yang didapat dari dua penelitian adalah

84,4%, dan 26,42%. Pada penelitiannya, digunakan pendekatan ekstraksi fitur warna dengan metode *Hue Saturation Value* (HSV). Hal yang membuat nilai yang dihasilkan berbeda-beda yaitu karena penggunaan metode HSV pada masing-masing penelitian berbeda. Penelitian dengan pendekatan ekstraksi fitur yang sama menggunakan metode *Gray Level Co-occurrence Matrix* (GLCM) dilakukan juga oleh V Sebastian, Unnikrishnan, & Balakrishnan (2012), dan B, Ramamurthy & Chandran, R,K (2012). Pada penelitian tersebut nilai rata-rata *precision* yang didapat dari dua penelitian adalah 72% dan 57,4%. Nilai yang dihasilkan pada masing-masing penelitian tentu berbeda, karena terdapat perbedaan penggunaan metode. Selain dilakukan secara tunggal, pendekatan ekstraksi fitur dapat digabungkan. Penelitian dengan ekstraksi fitur gabungan, telah dilakukan oleh Singh & Hemachandran (2012) dan Singha Manimala & K.Hemachandran (2012). Pada penelitian tersebut nilai rata-rata *precision* yang didapat dari dua penelitian adalah 71% dan 76%. Selain pemilihan ekstraksi fitur, pendekatan *preprocessing*, *similarity measure*, dan pemilihan data uji citra yang digunakan pada penelitian sebelumnya dapat menjadi faktor pembeda. Penelitian terdahulu yang telah dipaparkan, menggunakan data uji Citra Bunga.

Pada penelitian ini, penelitian yang dilakukan bukanlah penelitian yang ditujukan untuk melanjutkan penelitian sebelumnya, tetapi berdasarkan penelitian sebelumnya, penelitian ini akan menerapkan pendekatan ekstraksi yang sama tetapi dengan penggunaan metode yang berbeda. Poin lebih yang dimiliki oleh penelitian sebelumnya, akan diterapkan pada penelitian ini. Maka dari itu, penelitian ini akan melakukan pendekatan *preprocessing*, ekstraksi fitur, dan *similarity measure*. Kaur, Banga, & Banga (2013), pada penelitiannya membuktikan bahwa proses *preprocessing* dengan segmentasi dapat meningkatkan nilai *precision*. Maka proses *preprocessing* dilakukan dengan pendekatan segmentasi. Berdasarkan penelitian sebelumnya, ekstraksi fitur warna dan tekstur dapat memberikan nilai *precision* tertinggi 76% pada data uji Citra Bunga. Maka dari itu, pendekatan ekstraksi fitur yang digunakan adalah HSV, GLCM dan gabungan kedua fitur. Data uji Citra Bunga yang akan digunakan adalah *Oxford Flower 17*. Pendekatan *similarity measure* yang digunakan pada penelitian sebelumnya adalah *Euclidean Distance*

(ED), maka, penelitian ini akan menggunakan ED sebagai pendekatan *similarity measure*.

Pada penelitian ini, segmentasi dilakukan dengan tiga tahap, pertama proses perhitungan *thresholding*, proses penghilangan *noise*, morfologi, dan *masking*. Ekstraksi fitur yang digunakan adalah ekstraksi fitur warna *Hue Saturation Value* (HSV), ekstraksi fitur tekstur *Gray Level Co-occurrence Matrix* (GLCM), dan gabungan dari kedua fitur. Pendekatan ekstraksi fitur warna HSV dilakukan dengan mengubah nilai HSV menjadi histogram. Sedangkan, ekstraksi tekstur GLCM, akan digunakan berdasarkan 4 orientasi dan 4 properti dari GLCM. Orientasi pada sudut  $0^{\circ}$   $45^{\circ}$   $90^{\circ}$   $135^{\circ}$ , dan properti yang digunakan, yaitu *contrast*, *correlation*, *energy*, *homogeneity*. Fitur gabungan digunakan berdasarkan aturan standar pada MPEG-7, yaitu dengan merubah nilai hasil ekstraksi fitur warna dan tekstur kedalam bentuk vektor (Manjunath, Ohm, & Vasudevan, 2001). Nilai olahan citra yang telah diekstraksi akan diubah kedalam bentuk vektor. Nilai citra yang telah berbentuk vektor akan diukur tingkat dengan metode ED. Citra yang telah diukur tingkat kemiripannya, akan disusun berdasarkan tingkat kemiripan dan dimunculkan sebagai 10 besar citra termirip dengan *query*. Nilai *precision* pada penelitian ini akan diuji berdasarkan pengujian pada 10 query citra untuk setiap kelas bunga.

Dengan metode yang diusulkan pada pengembangan CBIR Citra Bunga, diharapkan pada penelitian ini, nilai *precision* yang didapat, mampu memberikan nilai yang lebih baik dari penelitian sebelumnya.

## 1.2 Rumusan Masalah

Dalam penelitian ini, terdapat rumusan masalah yang melatarbelakangi penelitian, diantaranya:

1. Bagaimana cara mengimplementasikan *HSV Feature Color Extraction*, dan ekstraksi fitur tekstur *Gray Level Co-occurrence Matrix* kedalam *Content Based Image Retrieval* Citra Bunga?
2. Bagaimana nilai *precision* dari hasil pengujian pada sistem *Content Based Image Retrieval* dengan metode yang telah diusulkan?

### 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah disusun pada bagian rumusan masalah, terdapat tujuan yang ingin dicapai penelitian, diantaranya:

1. Mengimplementasikan *HSV Feature Color Extraction*, dan ekstraksi fitur *Gray Level Co-occurrence Matrix* ke dalam *Content Based Image Retrieval*.
2. Mendapatkan nilai *precision* dari hasil pengujian pada sistem *Content Based Image Retrieval* dengan metode yang telah diusulkan.

### 1.4 Batasan Masalah

Berdasarkan penjelasan pada latar belakang, rumusan, dan tujuan, penelitian ini memiliki batasan masalah, yaitu data uji citra yang digunakan pada pengujian sistem *Content Based Image Retrieval* adalah *Oxford Flower* dengan kategori 17 kelas bunga, *Oxford Flower 17 (OF17)*.

### 1.5 Sistematika Penulisan

Untuk mempermudah penulisan penelitian ini, dibuat sebuah sistematika penulisan yang terdiri dari:

#### **BAB 1 : PENDAHULUAN**

Pada bab ini penulis menjelaskan tentang latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, dan sistematika penulisan. Latar belakang menjelaskan tentang permasalahan yang terjadi pada penelitian CBIR Citra Bunga. Rumusan masalah, merumuskan permasalahan dalam bentuk poin berdasarkan permasalahan pada CBIR Citra Bunga. Tujuan penelitian, ditujukan untuk menjawab permasalahan yang telah dirumuskan. Batasan masalah, menjelaskan data uji yang digunakan pada CBIR Citra Bunga. Sistematika penulisan akan menguraikan penjelasan penulisan yang dilakukan pada setiap bab.

#### **BAB 2 : KAJIAN PUSTAKA**

Pada bab ini penulis menjelaskan kajian pustaka yang terdiri dari *Content-Based Image Retrieval*, *Feature Extraction* yang terdiri dari; *HSV Color Feature Extraction*; *Gray Level Co-occurrence Matrix (GLCM)*, *Euclidean*

*distance*, *Oxford Flower Dataset*, *Mean Average Precision (MAP)*, *Segmentasi thresholding*, *Matrix Laboratory (MATLAB)*, dan Metode pengembangan perangkat lunak.

### **BAB 3 : METODOLOGI PENELITIAN**

Pada bab ini penulis menjelaskan prosedur yang akan dilakukan dalam penelitian. Prosedur yang dijelaskan, mencakup desain penelitian, dan alat dan baha yang digunakan pada penelitian ini. Desain penelitian akan menggambarkan secara struktural bagaimana penelitian dilakukan. Desain penelitian mencakup tahap awal penelitian, studi literatur, sistem CBIR Citra Bunga yang diusulkan, pengembangan perangkat lunak, eksperimen dan dokumentasi. Alat dan bahan mencakup alat berupa software dan hardware yang diperlukan dalam penelitian ini, dan bahan yang mencakup dataset serta kajian literatur untuk membantu penelitian dari sisi teori.

### **BAB 4 : HASIL PENELITIAN DAN PEMBAHASAN**

Pada bab ini akan diuraikan pengumpulan data yang terdiri dari daftar citra uji yang digunakan untuk pengujian, pengembangan perangkat lunak yang terdiri dari perancangan, implementasi kode dan pengujian/eksperimen, hasil dari pengujian akan diuraikan pada bagian hasil dan pembahasan eksperimen. Eksperimen dilakukan berdasarkan 6 skenario, yang terbagi dari kedalam citra tersegmentasi dan citra tidak tersegmentasi. Hasil eksperimen akan menguraikan perhitungan yang dilakukan pada setiap skenario. Pembahasan eksperimen akan menguraikan keadaan hasil perhitungan pada pengujian dan menyimpulkan hasil dari penelitian.

### **BAB 5 : KESIMPULAN DAN SARAN**

Pada bab ini akan diuraikan kesimpulan yang akan menjawab rumusan masalah penelitian, dan menguraikan saran yang dapat dilakukan untuk penelitian selanjutnya.