

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Search engine adalah sebuah program komputer yang dapat mengambil berbagai dokumen ataupun data dari sebuah *database* ataupun jaringan komputer khususnya internet. *Search engine* dapat mencari berbagai dokumen yang dianggap sesuai dengan kata kunci atau *keyword* yang kita inginkan. Dokumen yang dapat ditemukan oleh *search engine* dapat bermacam-macam bentuknya, salah satunya adalah dokumen dalam format gambar atau *image*.

Pencarian gambar (*image retrieval*) telah dikembangkan sejak tahun 1970. Pencarian gambar dapat dibagi menjadi 2 metode, yaitu : *text-based image retrieval* (TBIR) dan *content-based image retrieval* (CBIR). *Text-based image retrieval* adalah proses pencarian gambar dengan menggunakan teks. Proses pencarian gambar dengan menggunakan teks memiliki 2 kelemahan. Pertama, proses pencatatan biasanya dilakukan secara manual dan adanya perbedaan interpretasi antara user yang satu dengan *user* yang lain dalam hal pendeskripsian atau interpretasi gambar ke dalam kata-kata. Kemudian dikembangkanlah *content-based image retrieval*.

Content-based image retrieval (CBIR) adalah pencarian gambar dengan membandingkan komponen visual dari gambar seperti warna, tekstur, bentuk, atau pola gambar tersebut. Ekstraksi ciri yang dibutuhkan untuk pencarian berbeda-beda tergantung dari domain gambar yang akan dicari. Dengan menggunakan ekstraksi ciri yang tepat, maka hasil pencarian bisa lebih baik. Untuk gambar flora dan fauna, pencarian tidak cukup hanya dengan menggunakan ekstraksi warna. Hal ini disebabkan karena terdapat objek berbeda yang memiliki komposisi warna yang sama. Sebagai contoh bunga matahari memiliki komposisi warna kuning yang hampir sama dengan buah pisang. Jika kita hanya menggunakan ekstraksi warna dalam proses pencarian besar kemungkinan pisang akan ter-*retrieve* ketika melakukan *query* bunga matahari. Akan tetapi bentuk pisang dan bunga matahari berbeda, sehingga dengan menambahkan ekstraksi sisi dalam proses pencarian dapat meningkatkan hasil pencarian citra. Untuk gambar dengan domain flora dan fauna, ekstraksi ciri yang dapat digunakan adalah ekstraksi sisi dan ekstraksi warna.

Dalam sistem CBIR yang akan dibangun, ekstraksi sisi diperoleh dengan menggunakan metode *multi-scale edge detection* dan ekstraksi warna diperoleh dengan menggunakan *color histogram*. Pada dasarnya, penentuan sisi atau *edge* dengan menggunakan *multi-scale edge detection* adalah dengan mencari

representasi *edge* sebuah gambar dengan menggunakan sebuah operator *edge detection* akan tetapi proses pencarian *edge* dilakukan beberapa kali dengan perbedaan nilai *Gaussian blur*. Penggunaan nilai *Gaussian blur* yang berbeda bertujuan untuk mengurangi *noise* yang ada pada citra sehingga hasil *edge* yang diperoleh menjadi lebih baik. *Color histogram* yang akan digunakan dibagi berdasarkan color space dan representasi histogramnya. Adapun histogram warna yang akan digunakan antara lain RGB dengan kuantisasi 512 warna, RGB channel, dan HSV dengan kuantisasi 162 warna. Dengan memanfaatkan *edge detection* dan *color histogram* diharapkan hasil pencarian gambar bisa menjadi lebih baik.

1.2 Perumusan Masalah

Adapun rumusan masalah yang akan diselesaikana dalam tugas akhir ini adalah sebagai berikut :

1. Bagaimanakah penerapan *multi-scale edge detection* dan *color histogram* dalam *content based image retrieval system*?
2. Bagaimanakah performansi hasil pencarian dengan menggunakan ekstraksi ciri *multi-scale edge detection* dan *color histogram* dalam proses pencarian gambar?
3. Parameter apakah yang dapat mempengaruhi *multi-scale edge detection* dan *color histogram* dalam proses pencarian gambar?

1.3 Tujuan

Adapun beberapa tujuan yang ingin dicapai dari pembuatan tugas akhir ini antara lain :

1. Membangun *content based image retrieval system* dengan mengimplementasikan *multi-scale edge detection* dan *color histogram*.
2. Menganalisa performansi sistem CBIR yang dibangun untuk melakukan pencarian gambar dengan parameter *precision*, *recall*, dan *mean average precision*.

1.4 Batasan Masalah

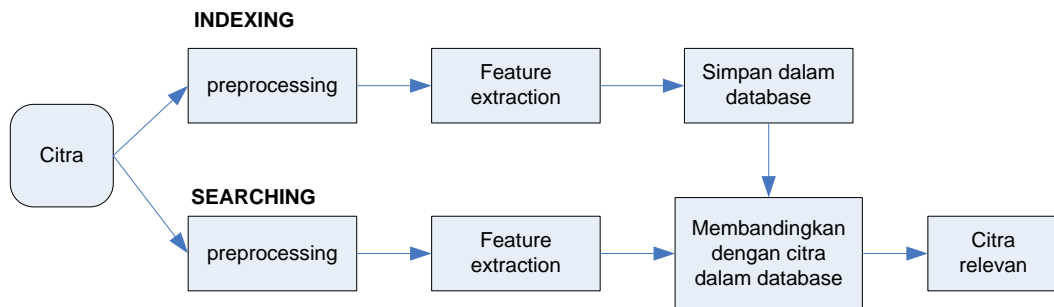
Dalam pengerjaan tugas akhir ini dibutuhkan batasan-batasan masalah agar tidak menyimpang dan mencegahn luasnya ruang lingkup permasalahan yang harus ditangani. Batasan-batasan masalah tersebut antara lain :

1. Penelitian yang dilakukan berfokus pada penerepan *multi-scale edge detection* dan *color histogram* untuk proses pencarian citra.
2. Hasil ekstraksi ciri akan direpresentasikan dalam bentuk histogram.
3. Citra yang menjadi masukan berupa citra berformat JPEG.
4. Citra yang akan dicari berupa gambar flora dan fauna yang diperoleh dari google dan flickr. Citra ini dikelompokkan ke dalam 9 kategori gambar.
5. Untuk analisis penerapan metode *multi-scale edge detection* dan *color histogram* dalam CBIR maka akan dibuat *search engine* yang bebasiskan *multi-scale edge detection* dan *color histogram*.
6. Adapun yang menjadi parameter performansi dari system CBIR yang dibangun adalah nilai *precision*, *recall*, dan *mean average precision* dari hasil pencarian dengan menerapkan *multi-scale edge detection* dan *color histogram*.

1.5 Metodologi Penyelesaian Masalah

Metodologi yang digunakan dalam penelitian tugas akhir ini adalah :

1. Study literatur
 - a. Mempelajari referensi dan sumber-sumber yang berkaitan dengan *content-based image* retrieval dan ekstraksi ciri. Adapun referensi yang digunakan adalah [9], [8], [3], [5], [10].
 - b. Mempelajari dan memahami metode *multi-scale edge detection* dan *color histogram* serta langkah penerepannya dalam sistem CBIR. Refrensi yang digunakan adalah [7], [15], [19], [16], [6], [18].
2. Analisis sistem
Dalam tahap ini dilakukan indentifikasi serta analisis kebutuhan sistem. Misalnya tahapan apa saja yang harus dilakukan serta metode pendukung seperti apa yang diperlukan dalam sistem.
3. Perancangan sistem
Sistem yang dirancang terdiri dari 2 proses utama, yaitu proses *indexing* dan proses *searching*. Proses *indexing* adalah proses mengekstrak ciri citra dan kemudian menyimpannya di dalam *database*. Proses *searching* adalah proses pencarian citra. Perancangan sistem dapat dilihat pada gambar 1.1.



Gambar 1.1 Perancangan sistem

4. Tahap implementasi

Pada tahap ini akan dilakukan implementasi dari rancangan system yang telah dibuat sebelumnya. Data-data yang menjadi masukan atau hasil akan disimpan untuk kemudian dianalisa. Adapun proses dalam tahap implementasi antara lain :

- a. Membuat program yang dapat menyimpan ekstraksi sisi dan warna citra.
- b. Membuat program yang dapat melakukan pencarian citra, berdasarkan ekstraksi sisi dan warna dari citra *query*.

5. Pengujian sistem

Dalam tahap ini akan dilakukan pengujian terhadap program yang telah diimplementasikan sebelumnya. Pengujian ini akan dilakukan dengan melakukan pencarian berdasarkan citra yang telah disiapkan.

6. Analisis hasil pengujian

Setelah sistem selesai diuji dengan melakukan pencarian beberapa citra, hasil yang diperoleh kemudian dianalisis dan ditentukan tingkat performansinya berdasarkan nilai *precision* dan *recall*. Dari hasil analisis tersebut akan diambil kesimpulan tentang penerapan *multi-scale edge detection* dan *color histogram* dalam proses pencarian citra.

7. Pembuatan laporan

Pada tahap ini, akan dilakukan penyusunan laporan akhir dan pengumpulan dokumentasi dengan mengikuti kaidah penulisan yang benar dan sesuai dengan ketentuan-ketentuan atau sistematika yang telah ditetapkan oleh institusi.