

BAB I

PENDAHULUAN

1.1 Latar Belakang

Saat ini di berbagai bidang seperti keamanan, perdagangan, pemerintahan, pendidikan dan rumah sakit, memiliki banyak koleksi citra digital. Banyak dari koleksi yang dimiliki tersebut merupakan hasil dari melakukan digitalisasi foto analog, diagram, gambar, lukisan. Selain itu peningkatan citra digital dipengaruhi oleh meningkatnya ketersediaan media untuk mengambil citra (Nagaraja & C.J, 2015). Penggunaan *database* citra digital adalah cara yang dapat digunakan untuk penyimpanan koleksi tersebut. Menjadi tantangan pada penelitian baru adalah bagaimana *transfer*, pemrosesan, pengarsipan, dan temu kembali pada *database* citra digital ini. Di antara tantangan tersebut, temu kembali citra digital dari *database* adalah topik penelitian yang penting dan paling menarik karena penerapannya dan berbagai masalah penelitian (Nagaraja & C.J, 2015). Dalam jurnal (Khrisne & Yusanto, 2018) para ahli mengemukakan dua teknik dalam proses temu kembali citra, yaitu *text-based image retrieval* (berbasis teks) dan *content-based image retrieval* (berbasis gambar).

Pertama *text-based image retrieval*, adalah teknik yang digunakan untuk pencarian citra berbasis teks. Teknik pencarian citra berdasarkan teks mulai digunakan untuk pencarian citra setelah periode 1970 (Kusrini, 2015). Pencarian citra dengan kueri teks dapat dilakukan jika citra mengandung deskripsi teks dan pengguna dapat mengingat deskripsi dari citra. Selain itu muncul sebuah kendala pada saat memberikan teks untuk mendeskripsikan sebuah citra adalah subyektifitas

yang mempengaruhi seseorang dalam mendeskripsikan gambar berdasarkan sudut pandang / perspektif. Sehingga tidak sedikit teks dari gambar tidak sesuai dengan deskripsi gambarnya (Sukafona & Thalib, 2018).

Teknik yang kedua adalah *Content-Based Image Retrieval* (CBIR) merupakan teknik temu kembali citra berbasis konten atau data. Pada teknik ini pencarian citra dilakukan dengan pencocokan atribut pada citra. Atribut - atribut citra mencirikan gambar pada citra, sehingga pencarian citra dapat dilakukan dengan membandingkan atribut – atribut tersebut (Layona, 2014). Tujuan dari CBIR atau temu kembali citra berbasis konten adalah untuk ekstraksi satu set fitur seperti ciri warna, tekstur dan bentuk yang dapat secara efektif menggambarkan ciri citra dan kemudian menggunakan ciri tersebut untuk menemukan kembali citra. Citra memiliki ciri warna, bentuk, dan tekstur yang dapat membedakannya dengan gambar yang lain. Setiap ciri citra didapatkan dengan proses ekstraksi ciri (Susilo, 2006). Secara garis besar proses CBIR terdiri dari : (1) *Preprocessing*, (2) Ekstraksi Ciri Citra, (3) Penghitungan Jarak Kedekatan.

Poin yang menjadi dasar perancangan sistem CBIR adalah pemilihan fitur yang menunjukkan efektifitas untuk merepresentasikan ciri dari sebuah citra. Fitur adalah atribut yang dapat membedakan citra dari citra yang lain. Salah satu karakteristik yang penting digunakan untuk melakukan analisa permukaan suatu objek adalah tekstur. Fitur tekstur (*texture*) adalah salah satu fitur atau ciri yang mempresentasikan pola dari permukaan sebuah obyek dalam suatu citra (Hartadi, Santoso, & Hidayatno, 2011). Fitur warna adalah salah satu ciri citra yang paling baik digunakan dalam suatu sistem CBIR, karena warna telah diuji memiliki

hubungan yang dekat dengan objek dalam suatu citra (Acharya, 2007). Selain itu segmentasi citra merupakan salah satu tahapan yang dapat digunakan dalam proses pengenalan pola suatu citra. Segmentasi adalah teknik untuk memecah atau membagi suatu citra menjadi beberapa daerah ataupun zona pada suatu citra dimana setiap zona memiliki kemiripan atribut (Trong-Ton Pham, 2015).

Beberapa penelitian CBIR yang telah dilakukan menggunakan fitur warna dan tekstur yang menjadikan deskripsi ciri dari citra adalah sebagai berikut. Penelitian oleh (Shahbahrami, 2008) menggunakan fitur warna dengan metode *Color Moment*, *Color Histogram*, dan *Color Coherence Vector (CCV)*. Sedangkan untuk ekstraksi fitur tekstur yang pertama digunakan metode *Co-occurrence Matrix* dan yang kedua *Discrete Wavelet Transform*. Dari ekstraksi fitur tekstur, metode *Co-Occurrence Matrix* memiliki tinggi nilai *recall* yang sama dengan metode *CCV* pada fitur warna dibandingkan dengan metode *DWT*, namun metode *DWT* memiliki nilai presisi yang sedikit lebih tinggi dari metode *Co-Occurrence Matrix* walaupun nilai presisi yang didapatkan masih lebih rendah dari metode *Color Moment*. *Color Moments* digunakan untuk ekstraksi ciri warna karena memiliki tingkat akurasi yang konsisten dan ukuran citra tidak mempengaruhi hasil dari ekstraksi karena ada distribusi probabilitas warna.

Penelitian oleh (Sadeq, Zahra, & Brooghani, 2017) menggunakan 1570 citra database yang dibagi menjadi 10 kelas. Adapun metode ekstraksi ciri warna yang digunakan adalah *Color Histogram* dan *Color Moment*. Sedangkan metode ekstraksi ciri tekstur yang digunakan adalah *Gray Level Co-Occurance Matrix (GLCM)* dan *Wavelet Transform*. Adapun proses pencocokan citra adalah dengan

menggabungkan seluruh vektor ciri dari masing-masing metode dan mengukur kemiripan ciri dengan menggunakan *Minkowski Distance*.

Penelitian oleh (Khrisne & Yusanto, 2018) menggunakan *Metode Block Truncation Algorithm* dan *Grid Partitioning*, pada penelitian ini di terapkan 2 metode yaitu *Block Truncation* yaitu algoritma untuk memetakan fitur warna dari sebuah citra yang selanjutnya akan diekstraksi dengan metode *Color Moment*. Yang kedua adalah tehnik *Grid Partitioning* yang berfungsi untuk memecah gambar menjadi tiga (3) bagian, yaitu bagian atas, tengah dan bawah. *Dataset* yang digunakan pada penelitian ini adalah *Wang Dataset*. Dengan menggunakan *grid* mendapatkan beberapa keuntungan yaitu mempersmpit area ekstraksi ciri. Dari hasil penilitian mendapatkan kinerja dari *grid partitioning* persegi panjang lebih baik daripada metode *Region Segmentation*.

Penelitian oleh (Atmaja, 2018) menggunakan kombinasi fitur warna, tekstur, bentuk dan penambahan teknik pengklasteran *K-Means*. Dalam penelitiannya penulis melakukan kombinasi fitur warna, tekstur, dan bentuk. Data yang digunakan adalah *Wang data set* sebagai data latih dan uji. Dari hasil penelitian nilai *precision*, *recall*, dan *accuracy* yang didapat dari pengujian yang dilakukan pada saat dilakukan proses klaster $k=10$ masing-masing mencapai 41% untuk nilai *precision*, untuk nilai *recall* mencapai 79%, dan *accuracy* 58%.

Berdasarkan uraian di atas dan penelitian sebelumnya yang sudah dilakukan, maka penelitian ini bertujuan untuk melakukan kombinasi beberapa metode dalam CBIR. Metode *Grid Partitioning* digunakan untuk melakukan segmentasi zona bertujuan untuk mempersempit proses ekstraksi pada citra sehingga mengurangi

waktu komputasi dalam proses ekstraksi (Trong-Ton Pham, 2015). Metode *Color Moments* digunakan untuk ekstraksi ciri warna karena memiliki tingkat akurasi yang konsisten dan ukuran citra tidak mempengaruhi hasil dari ekstraksi karena ada distribusi probabilitas warna. Untuk ciri tekstur digunakan metode *Co-occurrence Matrix*. Data yang digunakan dalam penelitian ini adalah *Wang Dataset*, yang dapat di unduh pada <http://wang.ist.psu.edu/~jwang/test1.tar>. Untuk pengujian dari kinerja CBIR akan menggunakan pengukuran *Precision*, *Recall*, dan *Accuracy*.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan pada penelitian ini, maka dapat diidentifikasi beberapa masalah yang menjadi fokus penelitian ini yaitu :

1. Peningkatan citra digital disegala bidang seperti keamanan, perdagangan, pemerintahan, dan kesehatan membutuhkan solusi dalam hal transfer data, pemrosesan, pengarsipan, dan temu kembali. Temu kembali citra merupakan topik penelitian yang penting dan paling menarik karena penerapannya dan berbagai masalah penelitian (Nagaraja & C.J, 2015).
2. Temu kembali citra menggunakan teks sangat tergantung kepada pengguna dalam mendeskripsikan sebuah citra baik untuk penyimpanan pada *database* ataupun untuk temu kembali citra. Karena subyektifitas yang mempengaruhi seseorang dalam mendeskripsikan gambar berdasarkan sudut pandang / perspektif.
3. *Content-Based Image Retrieval* (CBIR) atau temu kembali citra berbasis citra merupakan salah satu solusi. Penelitian dalam bidang CBIR telah banyak dilakukan. Kombinasi berbagai macam metode menghasilkan nilai

presisi yang berbeda dan menjadi referensi untuk penerapan dalam mesin pencari gambar serta penelitian selanjutnya, namun masih ada kombinasi metode lain yang yang berpeluang untuk diteliti.

1.3 Pembatasan Masalah

Adapun penulisan penelitian ini dibatasi permasalahannya agar tidak menyimpang dari masalah yang telah dirumuskan. Batasan-batasan tersebut antara lain :

1. Kombinasi metode yang digunakan diantaranya *Grid Partitioning* (GP) dengan *Color Moments* (CM) atau GP-CM, *Grid Partitioning* (GP) dengan *Co-occurrence Matrix* (GLCM) atau GP-GLCM, dan kombinasi *Grid Partitioning* (GP) dengan *Color Moments* (CM) serta *Co-occurrence Matrix* (GLCM) atau GP-CM-GLCM.
2. Citra digital yang dipergunakan dalam penelitian ini adalah *Wang dataset* diambil dari <http://wang.ist.psu.edu/~jwang/test1.tar>.
3. Citra terdiri dari 1.000 gambar digital yang secara umum dibagi menjadi 10 kelas utama dengan melakukan penyimpanan datanya secara bertahap berdasarkan kelas citra sampai dengan 1000 citra pada database. Adapun kelas dari gambar digital antara lain: (1) Afrika, (2) Pantai, (3) Gedung, (4) Bus, (5) Dinosaur, (6) Gajah, (7) Bunga, (8) Kuda, (9) Gunung, (10) Makanan (Wang, Li, & Wiederhold, 2001).
4. Tingkat kemiripan (*similarity*) antara citra *query* (uji) dengan citra *database* (latih) diukur dengan menggunakan perhitungan jarak dari *vektor* tiap-tiap fitur, menggunakan metode *Manhattan Distance*.

5. Keakuratan sistem dalam menemukan kembali citra yang relevan antara citra *query* (uji) dengan citra *database* (latih) diukur menggunakan *precision*, *recall*, *accuracy*.

1.4 Perumusan Masalah

Berdasarkan latar belakang dan identifikasi masalah yang telah diuraikan pada penelitian ini, maka dapat dirumuskan beberapa masalah yang menjadi fokus penelitian ini yaitu :

1. Bagaimana implementasi sistem *Content-Based Image Retrieval* menggunakan metode *Grid Partitioning*, *Color Moments*, dan *Co-occurrence Matrix*?
2. Bagaimana nilai *precision*, *recall*, dan *accuracy* dari *Content-Based Image Retrieval* menggunakan metode *Grid Partitioning*, *Color Moments*, dan *Co-occurrence Matrix*?

1.5 Tujuan Penelitian

Berdasarkan permasalahan yang telah dirumuskan, adapun tujuan yang diharapkan dari penelitian yang dibuat penulis ini adalah sebagai berikut.

1. Mengimplementasikan sistem *Content-Based Image Retrieval* menggunakan metode *Grid Partitioning*, *Color Moments*, dan *Co-occurrence Matrix*.
2. Menghitung nilai *precision*, *recall*, dan *Accuracy* dari *Content-Based Image Retrieval* menggunakan metode *Grid Partitioning*, *Color Moments*, dan *Co-occurrence Matrix* yang harapannya bisa menjadi referensi untuk penerapan pada mesin pencari gambar dan penelitian selanjutnya.