

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Dalam penelitian terdahulu diharapkan peneliti dapat membedakan penelitian yang ingin dilakukan dengan penelitian yang sudah pernah dilakukan oleh peneliti sebelumnya serta didapatkan kaitanya antara penelitian terdahulu dengan penelitian yang akan dilakukan.

Berikut perbandingan penelitian terdahulu yang telah dilakukan :

1. Penelitian dengan judul Analisis dan Implementasi *Content Based Image Retrieval* Menggunakan Metode ORB, yang telah diteliti oleh Muhammad Mirza, Tjokorda Agung Budi W., Siti Sa'adah, pada tahun 2015. Pada penelitian ini disimpulkan bahwa semakin rendah *threshold* maka semakin banyak pula *keypoint* yang bisa diperoleh pada metode oFAST, banyaknya jumlah *keypoint* pada setiap citra sebagai fitur sangat mempengaruhi akurasi ARP, Metode ORB *invariant* terhadap citra *noise Gaussian* dengan akurasi yang cukup baik, namun tidak terlalu baik dengan citra *resize*, dan Citra *query rotated* memiliki *size* yang berbeda dengan citra original, sehingga memungkinkan mendapatkan *keypoint* lebih banyak daripada citra original.
2. Penelitian dengan Judul *Content Based Image Retrieval* Batik Tradisional Yogyakarta Menggunakan Ekstrasi Ciri Berdasarkan Tekstur *Filter Gabor Wavelets 2D*, yang telah diteliti oleh Alfonsus Stefan Arwanda, dan Nazori Agani, pada tahun 2013. Pada penelitian ini disimpulkan bahwa program ini dapat berfungsi untuk mencari data citra berdasarkan citra masukkan yang diberikan oleh pengguna, yang kemudian akan ditampilkan lima belas citra yang memiliki jarak kemiripan yang hampir mendekati dan diurutkan berdasarkan jarak kemiripan yang terkecil hingga terbesar, tingkat kesesuaian terhadap data citra yang diambil dipengaruhi oleh pemahaman yang dimiliki oleh pengguna terhadap citra masukkan yang ingin dicari tersebut, dan kesesuaian terhadap hasil proses pencarian yang dilakukan oleh program ini bergantung pada banyaknya koleksi yang dimiliki oleh sistem.

3. Penelitian dengan judul *Content Based Image Retrieval* Citra Sidik Jari Menggunakan Metode *Wavelet Daubechies* Dan Diagram *NOHIS-Tree*, yang telah diteliti oleh Agung Satrio Buwono, pada tahun 2014. Pada penelitian ini disimpulkan bahwa performa *Wavelet Daubechies* sebagai metode ekstraksi fitur pada aplikasi *Content Based Image Retrieval* citra sidik jari terbagi menjadi 4 hasil dengan masing-masing memiliki akurasi yang berbeda-beda, performa CBIR menggunakan metode *NOHIS-Tree* mempunyai pencarian lebih cepat daripada CBIR menggunakan metode *brute force*.
4. Penelitian dengan judul *Content-Based Image Retrieval* Menggunakan Metode *Block Truncation Algorithm* dan *Grid Partitioning*, yang telah diteliti oleh Duman Care Khrisne, dan Mohamad David Yusanto, pada tahun 2015. Pada penelitian ini disimpulkan bahwa penggunaan metode *Grid Partitioning* (GP) dalam proses optimalisasi hasil temu kembali gambar nilai presisi berhasil bertambah, dan nilai *recall* menjadi lebih kecil dibandingkan dengan penelitian sebelumnya. Terjadi hubungan tarik-ulur (*trade-off*) antara *presisi* dengan *recall*, dengan kondisi nilai *presisi* yang tinggi dan nilai *recall* yang rendah. Sehingga dapat diartikan aplikasi dapat mengembalikan gambar dan hampir semuanya relevan, tetapi sejumlah besar gambar relevan lain terabaikan.
5. Penelitian dengan Judul *Content Based Image Retrieval* Menggunakan *Moment Invariant*, Tekstur Dan *Backpropagation*, yang telah diteliti oleh Ni G.A.P Harry Saptarini, dan Rocky Yefrenes Dillak, pada tahun 2012. Pada penelitian ini disimpulkan bahwa Metode yang dikembangkan mampu melakukan *query* terhadap *image* dalam *database* citra dengan *precision* sebesar 75%.
6. Penelitian dengan Judul Purwarupa Sistem *Content Based Image Retrieval* Pada Pencarian Produk Sepatu, yang telah diteliti oleh Baldri, Septia Rani, dan Izzati Muhimmah, pada tahun 2018. Pada penelitian ini disimpulkan bahwa sistem yang dibangun sudah dapat memberikan hasil pencarian dengan menampilkan citra-citra yang relevan. Untuk ke depannya, data citra yang digunakan sebagai basis data perlu ditambah lagi jumlahnya. Selain itu juga

perlu dieksplorasi pengaruh metode ekstraksi fitur yang berbeda terhadap tingkat presisi sistem.

7. Penelitian dengan Judul *Content Based Image Retrieval (CBIR) Berdasarkan Fitur Low Level : Literature Review*, yang telah diteliti oleh Rahmad Hidayat, Agus Harjoko, dan Anny Kartika Sari, pada tahun 2012. Pada penelitian ini disimpulkan bahwa kedepan proses tinjauan terhadap makalah CBIR dapat dilakukan dengan menggunakan klasifikasi yang lebih detail, seperti klasifikasi berdasarkan fitur-fitur *low level* pada gambar secara lokal. Ekstraksi fitur lokal pada gambar dilakukan berdasarkan objek yang terdapat dalam gambar yang telah melalui proses segmentasi.

2.2 Segmentasi Citra

Segmentasi citra merupakan langkah awal dalam klasifikasi citra dengan metode berbasis objek untuk mengelompokkan piksel yang mempunyai kesamaan struktur, bertujuan untuk membuat struktur individual menjadi region atau wilayah individual (Syahbana, 2013). *ECognition* yang mempunyai beberapa algoritma segmentasi seperti *chessboard*, *quadtrees*, *multiresolution*, dan *spectral difference* adalah salah satu perangkat lunak yang menyediakan proses segmentasi. Segmentasi yang sering digunakan adalah segmentasi multiresolusi.

2.3 Segmentasi Citra Landsat

Satelit *Landsat* pertama diluncurkan tahun 1972. Yang terakhir *Landsat 7* diluncurkan pada 15 April 1999. Dihasilkan jutaan citra pada instrumen satelit-satelit *Landsat*. Program *Landsat* terus berlanjut dengan diluncurkannya *Landsat 8* pada tanggal 13 Februari 2013 yang memiliki 11 saluran. Satelit *Landsat* dirancang untuk keperluan berbagai bidang seperti kehutanan, pertanian, geologi, perencanaan penggunaan lahan, dan lain sebagainya.

2.4 Segmentasi Berbasis Warna

Kebanyakan algoritma yang selama ini dikerjakan untuk mendeteksi vegetasi adalah berdasarkan segmentasi berbasis warna, dalam hal ini warna di

kuatisasikan (*quantized*) lebih dahulu. Menurut (Irianto, Y. Suhendro., 2008), kuatisasi adalah proses pengurangan warna dalam setiap band 1 - nya menjadi berjumlah 17 dari aslinya yang berjumlah 255. *Pixel* dengan *chanle* Merah (*red*) mempunyai intensitas antara 0 sampai dengan 15 yang diwakili oleh satu nilai, intensitas 15 – 30 untuk yang lain, intensitas 30 - 45 untuk yang lain lagi, dan seterusnya. Setelah itu, *region* atau *pixel* diberi label ke semua *pixel* yang saling berhubungan (satu terhubungan atau terkait dengan *pixel* sebelahnya melalui 8 *pixel* yang terhubungan) yang mempunyai nilai kuantisasi yang sama. *Region* atau daerah yang mempunyai kurang dari 100 *pixel* ditandai dan dibuang, semua *region* diproses dengan operasi morfologi.

2.5 Monitoring

Monitoring adalah siklus kegiatan yang mencakup pengumpulan, peninjauan ulang, pelaporan, dan tindakan atas informasi suatu proses yang sedang diimplementasikan (Mercy, 2005). Monitoring sering digunakan dalam *checking* antara kinerja dan target yang ditentukan. Ketika suatu proses sedang berlangsung, disitulah biasanya monitoring dilakukan. Pelaku monitoring umumnya adalah *self* monitoring (pelaku proses) ataupun atasan pekerja.

Fungsi dasar monitoring adalah *compliance* monitoring, yaitu memastikan proses sesuai rencana dan *performance* monitoring, yaitu mengetahui perkembangan organisasi dalam pencapaian target yang diharapkan (Mercy, 2005). Umumnya, *output* monitoring berupa *progress report* proses.

2.6 Deteksi Perubahan Daratan

Menurut Chen, H., Y. Xu, H. Shum, S-C. Zhu, and N. Zheng, (2001) setiap jengkal tanah di permukaan bumi adalah unik dan dicirikan oleh vegetasi yang menutupinya. *Land use* dan *land cover* secara nyata belum dapat digunakan untuk membedakan karakteristik dari permukaan bumi. Apa yang dapat kita jumpai pada permukaan daratan atau tanah dapat berupa tanaman pertanian, perkembangan kota, penebangan kayu, dan penambangan. Sedangkan *land cover* dapat berupa hutan, tanah gambut, padang rumput, jalan, dan daerah perkotaan.

Istilah *land cover* pada awalnya diartikan sebagai suatu vegetasi seperti hutan, rerumputan hanya saja kemudian arti tersebut meluas yang meliputi bangunan-bangunan, keanekaragaman hayati, dan perairan (Delac, K., Mislav Grgic, and Sonja Grgic, 2009). Sementara itu (El-Bakry, H., M., and Qiangfu Zhao, 2005) mengemukakan empat aspek penting dalam monitoring sumber daya alam, sebagai berikut: deteksi perubahan-perubahan yang sedang terjadi, identifikasi sifat perubahan yang terjadi, pengukuran luas daerah yang berubah, dan evaluasi model perubahan yang terjadi.

Beberapa teknik telah dikerjakan oleh para ahli untuk mendeteksi perubahan-perubahan tersebut, sebagai contohnya (Elisa Back, Danielle Ropar, and Peter Mitchell, 2007), dan (Freeman, W.T., E. Pasztor and O. Carmichael. *Learning low-level vision*, 2000) meringkas sebelas *algorithm* dalam deteksi perubahan, seperti: *Mono-temporal change delineation*, *Delta or post classification comparisons*, *multidimensional temporal feature space analysis*, *image differencing*, *composite analysis*, *multi temporal linear data transformation*, *image regression*, *change vector analysis*, *multi temporal biomass index*, *background subtraction*, dan *image rationing*. Sementara itu metoda atau teknik *matching citra* digital berdasarkan fitur *local invariant* telah diaplikasikan yang berkaitan dengan permasalahan-permasalahan dibidang *computer vision* seperti *image retrieval*, *object recognition*, *video data mining*, dan *image based localization*.

2.7 Klasifikasi Objek

Analisa citra digital pada dasarnya terdiri dari segmentasi (*image segmentation*), klasifikasi, dan interpretasi. Pada awalnya pekerjaan segmentasi dan klasifikasi dilakukan dengan metoda statistik dan pendekatan *fuzzy logic*. Metoda statistik digunakan dengan teori kemungkinan yang dapat digunakan dengan cara yang sederhana dan konsisten. Klasifikasi yang akan digunakan dalam penelitian ini adalah berdasarkan pendekatan *object-oriented*. Pendekatan atau teknik klasifikasi *object-oriented* tidak sama dengan klasifikasi berbasis *pixel* dan statistik, klasifikasi ini diharapkan dapat menjelaskan hubungan dari sematik

diantara objek-objek nyata di bumi dengan objek-objek yang ada dalam *aerial photography* dan atau citra *Landsat*. (Byung-Joo Oh., 2005) pernah menyatakan bahwa relasi atau kaitan ini dapat meningkatkan keakuratan hasil dari klasifikasi objek-objek yang ada dipermukaan dataran. Objek yang mempunyai banyak informasi di-ekstrak dengan menggunakan algoritma *segmentasi multi resolution*.

Banyak algoritma segmentasi yang menggunakan parameter-parameter tertentu, segmentasi multi-resolusi juga memerlukan parameter-parameter seperti skala, warna, bentuk, kepadatan, dan kehalusan. Ukuran dan bentuk dari citra suatu objek dapat dipengaruhi oleh nilai parameter yang berbeda, perlu juga diketahui bahwa untuk menemukan ukuran dan bentuk objek yang sesuai untuk klasifikasi membutuhkan banyak waktu dan sangat subjektif.

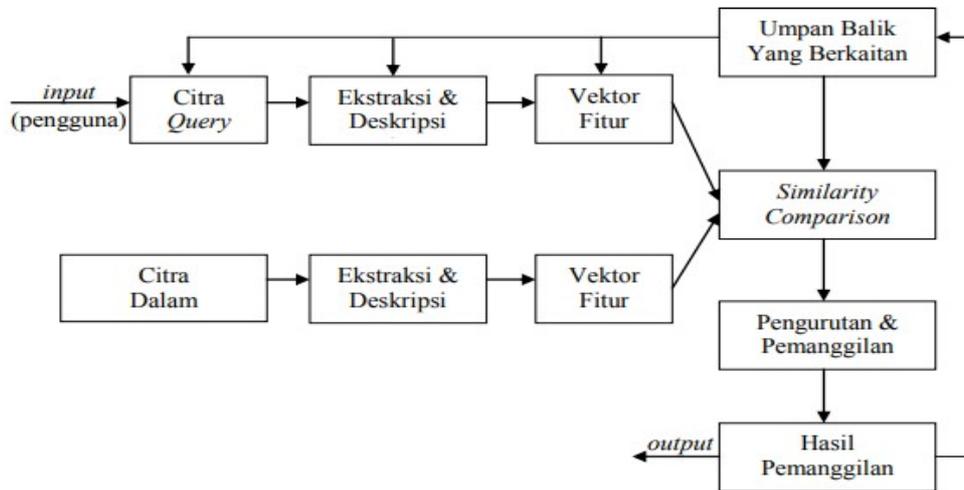
Dalam penelitian ini, objek yang terdapat dalam suatu citra di-ekstrak dengan menggunakan nilai *default parameter* tanpa memperhatikan ukuran dan bentuk. Kemudian berdasarkan objek yang penting dalam suatu citra, suatu awal proses dan deskripsi statistik kemudian dihitung. Contoh klasifikasi lain, adalah klasifikasi yang di control dikerjakan dengan menggunakan algoritma *backpropagation neural network*, kemudian klasifikasi yang tidak dikontrol dikerjakan dengan algoritma *Kohonen's self-organizing feature map* (SOFM), (Castelli, V., and L. D. Bergman (Eds.) 2006).

2.8 Sistem Content Based Image Retrieval (CBIR)

Content Based Image Retrieval (CBIR) didefinisikan sebagai proses pemanggilan citra dari *database* atau penyimpanan citra digital lainnya sesuai dengan konten citra tersebut (Chaudari,2012). Dibagi tiga fitur yaitu warna, tekstur dan bentuk. Fitur warna paling sering digunakan karena lebih mudah dalam pencarian informasi.

Tahap awal adalah melakukan proses ekstraksi dan deskripsi citra pada *database* oleh pengguna sehingga menghasilkan vektor fitur yang akan membentuk fitur *database*. Vektor fitur merupakan hasil proses ekstraksi. Kemudian mengukur jarak kesamaan vektor fitur citra *query* dengan vektor fitur citra dalam *database* dan akan diurutkan atau biasa disebut dengan *Similarity*

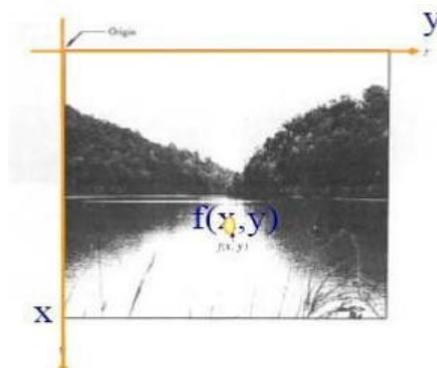
Comparison. Citra dengan jarak kesamaan vektor fitur tertinggi akan di tampilkan sebagai *output* (Long, et al. 2003). Diagram sistem pemanggilan citra berdasarkan konten :



Gambar 2.1 Diagram Sistem *Content Based Image Retrieval* (CBIR)
(Sumber: Long, et al. 2003)

2.9 Citra Digital

Citra digital dapat didefinisikan sebagai fungsi dua variabel $f(x,y)$ dimana x dan y adalah koordinat spasial dan nilai $f(x,y)$ yang merupakan intensitas citra pada koordinat tersebut. Teknologi dasar untuk menciptakan dan menampilkan warna pada citra digital berdasarkan pada penelitian terdahulu bahwa sebuah warna merupakan kombinasi dari tiga warna dasar, yaitu merah, hijau, dan biru (*Red, Green, Blue - RGB*). Sistem kordinat pada sebuah citra digital dapat dilihat pada Gambar 2.6.



Gambar 2.2 Sistem Koordinat Citra Digital

Dalam penelitian ini format *file* citra yang digunakan adalah format *file* citra JPEG (.jpg).

2.9.1 Format File Citra JPEG

JPEG atau *Joint Photographic Experts Group* adalah format citra yang sering digunakan untuk menyimpan citra-citra dengan ukuran lebih kecil. JPEG mampu menampilkan warna dengan kedalaman 24-bit *true color*, mengkompresi citra dengan sifat *lossy* dan biasanya digunakan untuk menyimpan citra-citra hasil foto. Jika ingin menampilkan citra dengan detail rumit dan bergradasi *file* citra dengan format ini dapat digunakan. *File* citra JPEG dapat menghasilkan citra yang hampir sama seperti aslinya. *File* citra JPEG mampu menghasilkan 16 juta warna. Ukuran *file* citra JPEG pada umumnya lebih besar dari GIF.

2.10 Precision Recall

Menurut Kurniawan (2010), *Recall* adalah perbandingan jumlah dokumen relevan yang terambil sesuai dengan *query* yang diberikan dengan total kumpulan dokumen yang relevan dengan *query*. *Precision* adalah perbandingan jumlah dokumen yang relevan terhadap *query* dengan jumlah dokumen yang terambil dari hasil pencarian. *Precision* dapat diartikan sebagai ketepatan atau kecocokan (antara permintaan informasi dengan jawaban terhadap permintaan itu). Sedangkan istilah *recall* dibidang sistem temu kembali informasi (*information retrieval*) berkaitan dengan kemampuan menemukan kembali informasi yang sudah tersimpan (Pendit 2008).

Sedangkan *Precision* dapat diartikan sebagai kepersisan atau kecocokan (antara permintaan informasi dengan jawaban terhadap permintaan itu). Jika seseorang mencari informasi di sebuah sistem, dan sistem menawarkan beberapa dokumen, maka kepersisan ini sebenarnya juga adalah relevansi. Artinya, seberapa persis atau cocok dokumen tersebut untuk keperluan pencari informasi, bergantung pada seberapa relevan dokumen tersebut bagi si pencari.

2.11 Regresi Linier

Regresi merupakan suatu alat ukur yang digunakan untuk mengukur ada atau tidaknya korelasi antar variable. Istilah regresi berarti ramalan atau taksiran pertama yang kali diperkenalkan oleh Sir Francis Galton pada tahun 1877.

Analisis regresi terbagi menjadi dua yaitu regresi linier dan Nonlinier. Analisis regresi linear terdiri dari analisis regresi linear sederhana dan analisis regresi linear berganda. Perbedaan antar keduanya terletak pada jumlah variabel independennya. Regresi linear sederhana hanya memiliki satu variabel independen, sedangkan regresi linear berganda mempunyai banyak variabel independen. Analisis regresi Nonlinier adalah regresi eksponensial.

2.11.1 Kriteria Data Regresi Linier

Terdapat dua syarat yang harus dipenuhi oleh data dalam menggunakan analisis regresi linier yaitu:

1. Data

Data harus terdiri dari dua jenis variabel, yaitu dependen dan independen. Selain itu data berupa kuantitatif dan variabel berupa kategori, seperti SD, SMA, SMK, dll.

2. Asumsi

Setiap data diasumsikan variabel dependen terdistribusi secara normal. Selain itu, antara variabel dependen dan independen harus memiliki hubungan linier dengan observasi harus saling bebas.