

BAB II

TINJAUAN PUSTAKA

Dalam bab ini penulis akan memaparkan beberapa sumber penelitian maupun teori yang dapat pendukung terkait dengan penelitian yang didapat dengan melakukan studi pustaka berbagai sumber literasi sebagai bahan landasan teori dalam melakukan penelitian ini. Sebelum membahas materi teori yang terkait, pada bab ini akan dibahas terlebih dahulu penelitian terkait yang berhubungan dengan pembuatan system ini dan yang menjadi acuan penelitian :

2.1 Penelitian Terdahulu

Penelitian ini menerapkan acuan keterkaitan dengan teori dari beberapa penelitian terdahulu. Dari penelitian yang sudah dilakukan sebelumnya penulis tidak menjumpai judul penelitian yang sama dengan judul yang diangkat penulis. Akan tetapi penulis mengambil beberapa penelitian terdahulu dan digunakan sebagai bahan refrensi kajian penelitian penulis. Berikut ialah beberapa penelitian terdahulu yang terkait dengan penelitian penulis.

Penelitian menurut Nurul Fuad (2017) dengan judul pemanfaatan metode k-nearest neighbor pada klasifikasi image berdasarkan pola fitur dan tekstur Pencarian gambar berbasis konten dapat menggunakan Content Based Image Retrieval (CBIR). CBIR bekerja dengan mengukur kesamaan gambar kueri dengan semua gambar dalam database sehingga biaya kueri adalah berbanding lurus dengan jumlah gambar dalam database. Membatasi jangkauan pencarian gambar dengan cara klasifikasi merupakan salah satu cara untuk mengurangi biaya query pada CBIR. Penerapan K-Nearest Neighbor Metode ini bertujuan untuk mengklasifikasikan citra serta mengukur tingkat akurasi dan waktu pengerjaan klasifikasi. Dari hasil akhir sistem ini serta dari hasil uji coba yang telah dilakukan, dapat ditarik kesimpulan bahwa fitur gambar, metode ekstraksi fitur dan jumlah neighbor yang melakukan voting mempengaruhi tingkat akurasi klasifikasi gambar dengan metode KNearest Neighbor. Pola Fitur yang memberikan tingkat akurasi paling tinggi adalah fitur warna dan tekstur

sedangkan metode ekstraksi fitur yang memberikan tingkat akurasi tertinggi adalah kombinasi Color Histogram 256 bin dan Edge Histogram Descriptor 2x2. Tingkat akurasi klasifikasi berfluktuasi tergantung jumlah neighbor yang melakukan voting namun cenderung menurun seiring bertambahnya jumlah neighbor yang melakukan voting. Adapun klasifikasi yang membutuhkan waktu paling singkat adalah klasifikasi dengan fitur tekstur, disusul kemudian oleh klasifikasi dengan fitur warna dan yang membutuhkan waktu paling lama adalah klasifikasi berdasarkan fitur warna dan tekstur. Dalam tiap percobaan digunakan 100 gambar testing. Adapun hasil pengujian dapat dilihat dalam grafik berikut: Fitur gambar yang digunakan dalam klasifikasi akan mempengaruhi tingkat akurasi klasifikasi. Klasifikasi dengan berdasarkan fitur warna dan tekstur memiliki tingkat akurasi rata-rata yang paling tinggi yaitu sebesar 75,81%. Metode ekstraksi fitur warna berpengaruh terhadap tingkat akurasi klasifikasi. Adapun metode ekstraksi fitur warna yang memberikan tingkat akurasi klasifikasi tertinggi adalah Color Histogram 32 bin (CH 32) yaitu sebesar 76,04 %.

Penelitian menurut Nugraha Candra Yuda Setiawan (2016) dengan judul “Penerapan metode *naïve bayes* untuk menentukan jenis ikan cupang hias”. Permasalahan yang dibahas sulitnya untuk mengenali ikan cupang hias beserta jenisnya. Penelitian ini menggunakan metode *naïve bayes* bertujuan untuk melakukan klasifikasi data pada kelas tertentu, kemudian pola tersebut dapat digunakan untuk menentukan spesifikasi jenis ikan cupang hias. Hasil dari penerapan metode *naïve bayes* menghasilkan cara yang bisa dilakukan oleh penghobi pemula. Sehingga lebih mengenal dan memiliki pengetahuan yang bisa membantu untuk memilah jenis ikan cupang hias dengan menentukan spesifikasi jenis ikan cupang yang terbaik.

Penelitian menurut Vina Purnama Sari dengan judul “Identifikasi *Betta Fish* berdasarkan ekstraksi bentuk menggunakan parameter *Eccentricity* dan *Metric*”. Permasalahan yang dibahas sulitnya membedakan bentuk spesies *Betta Fish* yang memiliki bentuk hampir serupa sehingga diperlukan identifikasi ikan secara otomatis agar tidak terjadi kerugian biaya akibat kesalahan mengenali spesies *Betta Fish*. Penelitian ini bertujuan untuk mengidentifikasi empat spesies yaitu

Halfmoon, Double tail, Serit dan Plakat. Metode yang digunakan *Thersholding* untuk proses segmetasi, parameter *Eccentricity* dan *Metric* untuk ekstrasi bentuk, dan klasifikasi *Probabilistic Neural Network*. *Thresholding* digunakan untuk memisahkan objek dengan background serta mengubah citra menjadi citra biner. *Eccentricity* merupakan nilai perbandingan antara jarak foci ellips minor dengan foci ellips mayor suatu objek. *Metric* merupakan nilai perbandingan antara luas dan keliling objek. *Probabilistic Neural Network* adalah algoritma jaringan syaraf tiruan yang menggunakan pelatihan *supervised* dalam pengambilan keputusan dan klasifikasi. *Dataset* yang digunakan berjumlah 160 citra *Betta Fish*. Jumlah citra setiap spesies adalah 40 citra. Proses ekstaksi bentuk parameter *Eccentricity* dan *Metric* dan klasifikasi *Probabilistic Neural Network* mencapai tingkat akurasi sebesar 71.875%.

Penelitian menurut Yuliana (2016) dengan judul “Pengenalan citra ikan cupang hias menggunakan algoritma canny dan template matching”. Permasalahan yang dibahas sulitnya mengetahui jenis-jenis ikan cupang hias. Citra yang berupa ikan cupang hias dijadikan sebagai inputan program, citra tersebut kemudian diproses menggunakan deteksi tepi canny bertujuan untuk mendapatkan garis tepian pada citra ikan cupang hias. Setelah didapat citra yang telah terdeteksi tepi canny langkah selanjutnya adalah pengenalan menggunakan metode template matching correlation. Dari uji coba yang dilakukan, didapat bahwa aplikasi ini dapat mengenali citra ikan cupang hias dengan tingkat keakuratan 52,5%. Banyaknya jumlah data template sangat mempengaruhi dalam keakuratan pengenalan citra ikan cupang hias.

Penelitian menurut Julia Elprida Hutagalung, Mhd Ihsan Pohan, Yuli Happy Marpaung (2020) dengan judul ”Identifikasi kesegaran ikan nila menggunakan teknik citra digital”. Berdasarkan uji coba dan analisis hasil pengujian pada metode klasifikasi ikan menggunakan ORB dan KNN, dapat diuraikan beberapa kesimpulan. Algoritma ORB dapat mendeteksi keypoint dan menghasilkan fitur yang invariant terhadap perubahan sudut rotasi pada citra ikan. Terbukti dengan pengujian klasifikasi ikan dengan perbedaan rotasi, hasil tetap menunjukkan nilai akurasi yang tinggi. Klasifikasi ikan menggunakan KNN dari

hasil penelitian ini sangat akurat. Dari hasil uji coba pengujian, walaupun memiliki komputasi yang lebih lama, metode yang diusulkan memiliki tingkat keberhasilan dalam mengklasifikasi seluruh sampel data ikan dengan tingkat akurasi sangat bagus.

Penelitian menurut Taftyani Yusuf Prahudaya dan Agus Harjoko (2017) dengan judul “Metode klasifikasi mutu jambu biji menggunakan KNN berdasarkan fitur warna dan tekstur”. Kajian ini dapat diambil kesimpulan bahwa metode klasifikasi KNN yang diusulkan dapat melakukan klasifikasi jambu biji dengan akurasi 91.25% pada nilai ketetangaan terbaik $K=3$. Fitur yang diekstraks untuk klasifikasi jambu biji adalah RGB, luas cacat, *energy*, *homogeneity*, dan *contrast*. Pada penelitian ini masih terdapat beberapa hal yang perlu disempurnakan antara lain penggunaan kamera yang dapat dikendalikan secara langsung sehingga proses akuisisi data lebih cepat, penggunaan konveyor untuk jambu biji agar dapat diterapkan di industri.

Penelitian menurut (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 juga dilakukan oleh Zubair Anis dan Muslikh (2017) dengan judul “Identifikasi Jamur Menggunakan Metode *K-Nearest Neighbor* dengan Ekstraksi Ciri Morfologi”. Dalam proses pengklasifikasian jamur menggunakan metode KNN menggunakan ekstraksi ciri morfologi menghasilkan tingkat akurasi tertinggi yang berada pada nilai $k=60$ yaitu mencapai 99%. dalam range antara

$k=8000$ hingga $k=60$, tingkat akurasi mengalami kenaikan yang berarti semakin tinggi nilai k , maka akurasi semakin menurun.

Penelitian yang dilakukan Saidah, dkk (2019), dengan judul identifikasi kualitas beras menggunakan metode *K-Nearest Neighbor dan Support Vector Machine*. Dalam penelitiannya hasil analisis dan pengujian simulasi yang sudah dilakukan, dihasilkan beberapa kesimpulan sebagai berikut: intensitas cahaya yang paling baik digunakan adalah dalam kondisi terang. Warna latar yang paling baik digunakan untuk mengetahui kualitas grade beras yaitu ialah warna latar hitam. Faktor resize mempengaruhi akurasi dan ukuran citra yang terbaik. Pada penelitian ini jenis K-NN yang paling baik digunakan adalah euclidian distance dengan nilai $K=1$, memperoleh akurasi sebesar 96,67 %. Parameter pengujian SVM adalah kernel option dan kernel type, berturut-turut memperoleh akurasi sistem optimal saat kernel option 7 dan type kernel polynomial dengan akurasi 96,6 % untuk OAA dan OAO.

Penelitian yang dilakukan oleh Farokhah Lia (2020), dengan judul “Implementasi *K-Nearest Neighbor* untuk klasifikasi bunga dengan ekstraksi fitur warna RGB”. Metode Klasifikasi KNN dengan metode ekstraksi fitur warna RGB kurang cocok untuk mengklasifikasi label objek dengan kemiripan warna walaupun bentuknya berbeda namun masih bisa dipakai dalam mengklasifikasikan objek dengan perbedaan warna yang mencolok di salah satu sisi walaupun bentuk objeknya memiliki struktur bentuk yang mirip. Untuk meningkatkan akurasi rata-rata klasifikasi label pada bunga yang memiliki kesamaan warna perlu dilakukan uji coba dengan metode klasifikasi KNN dengan tambahan ekstraksi fitur dan parameter lainnya atau metode lain yang lebih handal.

penelitian selanjutnya dilakukan oleh Hidayat, dkk (2020), dengan judul “Segmentasi pasar kopi menggunakan metode KNN di Indonesia”. Pada identifikasi citra kopi menggunakan algoritma *K-Nearest Neighbor* dengan menggunakan 100 data citra yang terdiri dari gambar kopi dan non kopi. Data uji sebanyak 20 citra dan memperoleh akurasi sebesar 83.3%. Hasil dari segmentasi pasar kopi dengan data yang diambil dari instagram berdasar hastag (#) kopi dan

hashtag (#) daerah menghasilkan bahwa peminat kopi tertinggi adalah Surabaya. Berdasarkan hasil diatas dapat disimpulkan bahwa dengan menggunakan algoritma *K-Nearest Neighbor* serta menggunakan ekstraksi fitur histogram dapat diterapkan dengan baik pada identifikasi citra kopi dan non kopi.

2.2 Klasifikasi

Klasifikasi merupakan suatu metode yang digunakan untuk proses pengelompokan data. Klasifikasi juga dapat diartikan suatu metode pengelompokan data atau objek baru kedalam suatu kelas maupun kategori yang sesuai dengan variabel-variabel tertentu. Data mining merupakan teknik dari klasifikasi, yang digunakan untuk melihat suatu kelompok data yang sebelumnya sudah didefinisikan. Atribut ini dijadikan variabel untuk menentukan suatu kelas objek baru. Proses klasifikasi ini memiliki tujuan untuk menentukan suatu kelas objek yang kelasnya belum diketahui secara akurat (nugraga, 2018).

Dalam proses klasifikasi memiliki dua fase, ialah fase learning dan fase testing. Fase learning ialah sebagian data yang kelas datanya sudah diketahui dan dijadikan untuk model yang akan dibuat. Sebaliknya fase testing memiliki arti model fase yang telah terbentuk dan diuji dengan sebagian data lainya untuk mengetahui tingkat akurasi model tersebut. Jika model memiliki tingkat akurasi yang cukup maka model bisa dipakai untuk memprediksi suatu kelas data yang masih belum diketahui. (Juhrah, 2015).

2.3 Citra Digital

Citra digital merupakan dua kedalaman ruang yang dapat ditampilkan oleh layar komputer sebagai sekumpulan nilai digital yang disebut *pixel element* (liantoni, 2015).

2.3.1 Jenis-jenis Citra

Ada banyak cara yang digunakan untuk menyimpan citra digital kedalam memori penyimpanan. Proses penyimpanan menentukan jenis citra yang

terbentuk. Menurut (Sutoyo, 2019) ada beberapa jenis citra yang sering dipakai ialah citra warna, citra grayscale dan citra biner.

- a. Citra Biner (Monokrom). memiliki dua warna, ialah hitam dan putih. Untuk menyimpan kedua warna ini membutuhkan space 1 bit pada memory penyimpanan.
- b. Citra Grayscale (Skala Keabuan). Banyaknya warna bergantung kepada jumlah bit yang disediakan pada memory penyimpanan untuk memuat kebutuhan dari warna ini. 4 warna diwakili dengan citra 2 bit, 8 warna diwakili citra 3 bit, begitu juga seterusnya. Karena dengan semakin besarnya jumlah bit warna yang tersedia didalam memori akan semakin halus pula gradasi warna yang didapatkan.
- c. Citra Warna (True Color). Setiap piksel di citra
Setiap piksel yang terdapat pada citra warna merupakan kombinasi perwakilan dari tiga warna dasar yaitu Red Green Blue. Setiap warna memakan space penyimpanan 8 bit = 1 byte, setiap warna memiliki gradasi sebanyak 255 warna. Yang artinya setiap piksel memiliki kombinasi warna sejumlah $28 \times 28 \times 28 = 224 = 16$ juta lebih warna. Dengan jumlah warna yang sangat banyak dan hampir mencakup seluruh warna yang ada dialam format ini dinamakan true color.

2.3.2 Elemen-elemen Citra Digital

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

- a. Kecerahan (Brightness)
Mengambarkan suatu intensitas cahaya yang dikeluarkan melalui piksel citra sehingga bisa ditangkap oleh indra penglihatan.
- b. Kontras (Contrast)
Merupakan komposisi terang dan gelap pada suatu citra, Citra dengan kualitas baik memiliki komposisi yang sama antara terang dan gelap.
- c. Kontur (Countour)

Merupakan suatu keadaan yang disebabkan oleh alterasi ketajaman pada piksel-piksel yang berdekatan.

d. Warna (Colour)

Warna merupakan persepsi yang akan ditangkap oleh sistem visual melalui panjang gelombang cahaya yang dipantulkan melalui objek.

e. Bentuk (Shape)

Merupakan suatu properti intrinsik sebuah objek 3 dimensi, dan merupakan properti intrinsik utama untuk sistem visual manusia.

f. Tekstur (Texture)

Identik dengan penyaluran spesial berasal dari derajat keabuan pada sekelompok piksel-piksel yang saling berdekatan.

2.3.3 Pengolahan Citra

Pengolahan citra ialah suatu proses dalam mengolah piksel-piksel pada suatu citra digital yang bertujuan untuk mendapatkan hasil sesuai yang di inginkan. Membangun citra agar mudah ditafsirkan dan difahami oleh manusia. Citra dapat berupa gambar, video atau sesuatu yang sifatnya digital dan dapat dimasukkan dalam memori penyimpanan. Setiap citra bisa didapatkan dengan cara akuisisi citra, yaitu suatu proses yang dilakukan untuk memperoleh citra.

2.3.4 Teknik Pengolahan Citra

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

a. Image Enhancement

Suatu metode yang digunakan dalam proses perbaikan citra dengan cara perbaikan kualitas citra, baik dari tingkat kecerahan ataupun kontras.

b. Image restoration

Suatu metode yang digunakan dalam proses perbaikan model citra, untuk menjadikan bentuk model citra yang sesuai.

c. *Color Image Processing*

Suatu metode yang dikaitkan dengan citra warna, baik itu dalam bentuk image enhancement, image restoration, ataupun bentuk lainnya.

d. *Wavelet dan Multiresolution processing*

Suatu proses citra yang menyatakan dalam berapa resolusi.

e. *Image Compression*

Suatu metode yang dipakai untuk merubah suatu ukuran data yang ada didalam citra.

f. *Morphological Processing*

Suatu metode yang digunakan untuk mendapatkan sebuah informasi yang menyatakan deskripsi suatu bentuk dari sebuah citra.

g. *Segmentation*

Suatu metode yang digunakan untuk memisahkan dan membedakan suatu objek yang ada didalam suatu citra. Contoh kecil memisahkan objek dengan background.

h. *Object Recognition*

Suatu metode yang digunakan untuk mengenali objek apapun yang ada pada suatu citra.

2.4 Citra Warna RGB

Citra warna RGB merupakan suatu program yang digunakan untuk menampilkan warna di layar. RGB termasuk salah satu struktur dari model warna. Suatu gambar yang memiliki model warna rgb akan mempunyai 3 layer warna dasar dalam proses pembentukan warna, ialah *Red*, *Green*, *Blue*. Secara matematis, model warna RGB hampir sama dengan sistem koordinat kartesius, dikarenakan ketiga sumbunya bernama R, G dan B. Ruang warna pada model RGB akan dibagi, divisualisasikan sebagai kubus. Bisa dilihat pada gambar

dibawah ini. pada gambar tersebut dapat didefinisikan semua nilai warna setelah melalui proses normalisasi yang dilakukan oleh sistem dengan hasil rentan nilai sejumlah $[0,1]$. Nilai pada warna RGB yang paling utama terletak di sumbu utama pada sumbu kordinat. Sedangkan semua sudut kubus selain kubus utama, akan melakukan proses perpaduan antara warna utama yang akan merepresentasikan nilai warna sekunder. Warna sekunder tersebut terdiri dari 3 warna yaitu *magenta*, *yellow* dan *cyan*. Pada titik kordinat pusat memiliki nilai $(0,0,0)$ yang merepresentasikan warna *black* (hitam). Sedangkan warna *white* (putih) terletak pada sudut yang paling jauh dari titik koordinat atau titik nol $(1,1,1)$. Rentang nilai yang ditarik dari titik $(0,0,0)$ sampai dengan $(1,1,1)$ merupakan diagonal ruang yang menghasilkan gradasi warna keabuan (Pamungkas, 2015). Berikut merupakan gambar citra warna RGB.



Gambar 2.1 Gambar warna citra RGB

2.5 Fitur *Grayscale*

Grayscale Merupakan suatu citra yang hanya memiliki warna dengan tingkat keabuan. Hal ini dikarenakan fitur warna *grayscale* hanya memerlukan sedikit informasi di setiap pikselnya dibandingkan dengan citra yang memiliki warna. Warna RGB (*Red*, *Green* dan *Blue*) merupakan warna abu-abu pada citra *grayscale* yang mempunyai ketajaman yang sama. Sehingga citra *grayscale* hanya memerlukan suatu nilai dengan intensitas tunggal dibandingkan dengan citra berwarna yang memerlukan 3 intensitas disetiap pikselnya. Intensitas yang dihasilkan citra *grayscale* akan disimpan pada 8bit *integer* dimana yang nantinya

akan memberikan 256 kemungkinan yang dimulai dari level rendah atau level 0 sampai dengan level tinggi atau level 255 dimana hitam memiliki nilai 0 dan Putih memiliki nilai 255 dan nilai diantaranya ialah nilai derajat keabuan, dan merupakan Tingkat keabuan atau grayscale level (Anggraini, Hidayat, & Darana, 2017).

2.6. Segmentasi Citra

Segmentasi citra ini adalah salah satu bagian dari pengolahan citra dimana proses yang dilakukannya mengenalkan suatu objek ke sistem. Mereka juga mempunyai maksud untuk membelah citra jadi beberapa wilayah yang jenisnya dikelompokkan berdasarkan tingkatan keabuan dari *pixel* dengan tingkatan yang ada disebelahnya. Lalu dalam prosesi segmentasi suatu citra nantinya dilanjutkan ke proses pada tingkatan yang lebih tinggi terhadap suatu citra. dibawah ini merupakan proses segmentasi menurut pendapat Giannakopoulos (2008).

- a. Menentukan citra *Red, Green, Blue* (RGB) yang menjadi objek deteksi dengan acuan nilai warna grayscale atau hasil dari proses data training serta nilai toleransi grayscale yang dipakai.
- b. Mengubah citra RGB menjadi *grayscale*.
- c. Lakukan proses filtering warna pada citra dengan nilai (T) dan nilai toleransi sebagai acuannya. Nilai x dijadikan sebaai warna *grayscale* pada piksel yang ada maka warna yang tidak masuk kedalam rentang $T - tol < x < T + tol$ akan diberi warna hitam.
- d. Mengubah kembali citra menjadi RGB dan tampilkan hasil *filtering* (Nur Khasanah, Harjoko, & Candradewi, 2016).

2.7 Gray Level Co-Occurrence Matrix (GLCM)

GLCM adalah *matriks* yang dibangun menggunakan histogram tingkat kedua. Fitur ini mempunyai keteraturan pola tertentu yang terbentuk dari susunan piksel pada citra. Suatu citra dikatakan mempunyai informasi tekstur apabila memiliki pola atau karakteristik pada citra yang muncul secara berulang dengan interval

jarak dan arah tertentu. Berikut merupakan beberapa contoh fitur tekstur yang umum digunakan yaitu: entropi, energi, kontras, homogenitas, nilai rata-rata (mean) intensitas, dan deviasi standar. Fitur tersebut telah diusulkan oleh Haralick dan rekan-rekannya pada tahun 1973 (Lusiana, Al Amin, Hartono, & Kristianto, 2019).

Metode *Grey Level Co-Occurrence Matrix* (GLCM) termasuk dalam metode statistic dalam perhitungan menggunakan distribusi derajat keabuan (histogram) dengan mengukur tingkat kekontrasan, granularitas, dan kekasaran suatu daerah dari hubungan ketetanggaan antar piksel di dalam citra. GLCM merupakan metode untuk mengekstraksi tekstur orde kedua. Orientasi sudut pada GLCM dinyatakan dalam derajat, standarnya 0° , 45° , 90° , dan 135° . Nilai dari hubungan derajat keabuan akan ditransformasikan ke matriks *co-occurrence* dengan ukuran window 3×3 , 5×5 , 7×7 , 9×9 , dan seterusnya. (Abdul, 2012).

Pada penelitian ini hanya digunakan beberapa parameter saja persamaan dari parameter tersebut dinyatakan sebagai berikut :

- a. Nilai *Energy* menunjukkan ukuran sifat homogenitas citra pada penyebaran derajat keabuan. Nilai *Energy* yang tinggi muncul pada saat tekstur citra cenderung seragam.

$$Energy = \sum_i \sum_j \{p(i, j)\}^2 \dots \dots \dots 2.1$$

- b. Contrast menunjukkan ukuran penyebaran(momen inersia) elemen- elemen matrik citra. Apabila posisinya jauh dari diagonal utama, nilai kekontrasan besar. Secara visual, nilai kekontrasan merupakan ukuran variasi antar derajat keabuan suatu wilayah citra

$$Contrast = \sum_i \sum_j (i - j)^2 p(i, j) \dots \dots \dots 2.2$$

- c. *Correlation* mengukur ketidakmiripan merupakan suatu tekstur dimana nilainya akan besar bila acak dan bernilai kecil jika seragam.

$$Correlation = \frac{\sum_i \sum_j p(i, j) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \dots \dots \dots 2.3$$

- d. Nilai *Homogeneity* menunjukkan kehomogenan citra yang berderajat keabuan sejenis. Citra homogen akan memiliki Homogeneity yang besar.

$$Homogeneity = \sum_i \sum_j \frac{1}{1+(i+j)^2} p(i, j) \dots \dots \dots 2.4$$

(Idestio, Barsyah Dwi, 2013)

2.8 Histogram of Oriented Gradients (HOG)

Rekayasa fitur dilakukan untuk mencari ciri dari citra wajah, yakni dengan menggunakan metode HOG. *Histogram of Oriented Gradients* (HOG) adalah teknik ekstraksi fitur dalam pengolahan citra yang mengelompokkan nilai gradien piksel menurut orientasi arah pada setiap bagian lokal dari citra. Tampilan dan bentuk objek lokal seringkali dapat dikarakterisasi cukup baik oleh distribusi gradien intensitas lokal atau arah tepian walaupun tidak diketahui persis posisi gradien atau tepian yang sesuai. Hal inilah yang menjadi ide dasar dari teknik ekstraksi fitur yang diajukan oleh Dalal dan Triggs (2005).

Distribusi gradien atau histogram inilah yang menjadi suatu fitur untuk pencocokan kesamaan (*similarity matching*) (Utaminigrum et al., 2017), sehingga dapat dilatihkan ke dalam suatu pembelajaran mesin. Proses HOG diawali dengan menghitung nilai gradien terhadap sumbu X dan Y (G_x , G_y) dari hasil selisih nilai piksel citra (B_u) pada indeks sebelumnya dengan nilai piksel citra pada indeks sesudahnya sesuai sumbu masing-masing seperti persamaan (2) dan (3).

$$G_x(x, y) = B_u(x + 1, y) - B_u(x - 1, y) \quad (2) \dots\dots\dots 2.5$$

$$G_y(x, y) = B_u(x, y + 1) - B_u(x, y - 1) \quad (3) \dots\dots\dots 2.6$$

Keterangan :

G_x , : citra gradien sumbu X dan Y

: citra awal

Citra gradien diubah ke dalam besaran gradien (φ) yaitu jarak Euclidean posisi X dan Y, dan orientasi (θ) yaitu inverse tangen dari posisi X dan Y. Proses ini identik dengan pengubahan koordinat Cartesian dengan nilai gradien sebagai nilai posisi pada sumbu X dan Y menjadi koordinat Polar seperti persamaan (4) dan (5).

$$= \sqrt{G_x^2 + G_y^2} \quad (4) \dots\dots\dots 2.7$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (5) \dots\dots\dots 2.8$$

Keterangan :

: besaran gradien (gradient magnitude)

: orientasi gradien (gradient orientation)

Kelompok orientasinya dilakukan pada setiap bagian lokal citra atau disebut sel. Satu sel terdiri dari piksel-piksel sesuai ukuran yang bisa ditentukan, dan dari satu sel pula dapat dibentuk histogram dengan jumlah kelompok bin yang bisa ditentukan. Pembentukan kelompok dapat dilakukan dengan mengindahkan arah derajat atau tidak (signed atau unsigned). Penentuan besaran gradien dan orientasi dilakukan dengan metode interpolasi trilinear untuk mencari derajat keanggotaannya, yang terdiri dari interpolasi bilinear terhadap sel dan interpolasi linier terhadap bin histogram, sehingga menghindari efek aliasing, yakni nilai gradient akan tetap sesuai bila direkonstruksi.

Normalisasi dilakukan di tiap grup sel yang disebut blok. Normalisasi tiap blok atau bagian lokal citra dilakukan karena beragam tingkat besaran gradien dapat terjadi dalam citra, yang dipengaruhi oleh variasi pencahayaan dan kontras latar belakang dan latar depan. Proses normalisasi dalam satu blok dapat dilakukan dengan cara seperti salah satu dari persamaan (4).

$$L1 - norm : f = \frac{v}{(\|v\|_1 + \epsilon)} \dots\dots\dots 2.9$$

$$L1 - sqrt : f = \sqrt{\frac{v}{(\|v\|_1 + \epsilon)}}$$

$$L2 - Hys : L2 - norm, clip(0,0.2), renorm$$

$$L2 - norm : f = \frac{v}{\sqrt{\|v\|_2^2 + \epsilon^2}}$$

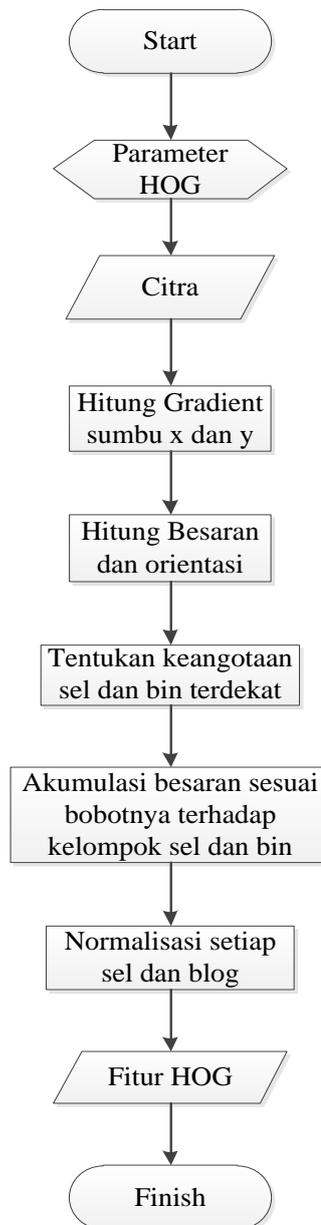
Keterangan :

f : vektor histogram telah ternormalisasi

v : vektor histogram belum ternormalisasi

$\|v\|_k$: normalisasi vektor, jenis normalisasi $k = 1$ atau $k = 2$

ϵ : konstanta nilai kecil (10^{-n} , $n > 0$), untuk menghindari pembagian dengan nol.



Gambar 2.2 Flowchart *Histogram of Oriented Gradients*

Dengan menyesuaikan parameter jarak antar blok (stride) ataupun tumpang tindihnya (overlap), banyak blok mampu memiliki salah satu vektor dari sel yang sama, tetapi dengan hasil yang berbeda sebab normalisasi yang terpengaruh sel pada blok tersebut. Hasil normalisasi tersebut ialah akhir dari metode ini, menciptakan vektor fitur yang mewakili wujud suatu citra. Jumlah vektor fitur yang dihasilkan dapat dihitung dari faktor jumlah bin, jumlah sel yang dihasilkan,

serta berapa banyak blok yang dapat terjalin dari pemilihan dimensi blok serta langkah blok. Secara garis besar, alur metode *Histogram of Oriented Gradients* digambarkan pada Gambar 2.2.

2.9 *K-Nearest Neighbors* (KNN)

Algoritma *K-Nearest Neighbor* (K-NN) merupakan metode klasifikasi yang menentukan label (class) dari sebuah objek baru berdasarkan mayoritas class dari jarak terdekat k dalam kelompok data latih. (Wijaya & Ridwan, 2019)

Algoritma *K-Nearest Neighbor* (K-NN) merupakan algoritma untuk mengklasifikasi yang bekerja mengambil sejumlah K data terdekat (ketetanggaanya) sebagai acuan untuk menentukan kelas dari data baru. KNN ini cara mengklasifikasinya yaitu berdasarkan *similarity* atau kemiripan atau yang paling dekat terhadap data lainnya. Algoritma KNN data yang paing dekat disebut atau data poin yang berada berdekatan dikatakan “neighbor” atau “tetangga” (Afifah, 2020).

Secara umunya algoritma KNN memiliki cara kerja sebagai berikut.

1. Menentukan jumlah tetangga (K) digunakan sebagai acuan pertimbangan penentuan kelas.
2. Menghitung data yang baru dimasukan ke masing-masing data poin dataset.
3. Mengambil beberapa jumlah K data terhadap jarak terdekat, lalu ditentukan kelas dari data yang baru.

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \dots\dots\dots(2.10)$$

Dengan keterangan :

x_1 :sampel data

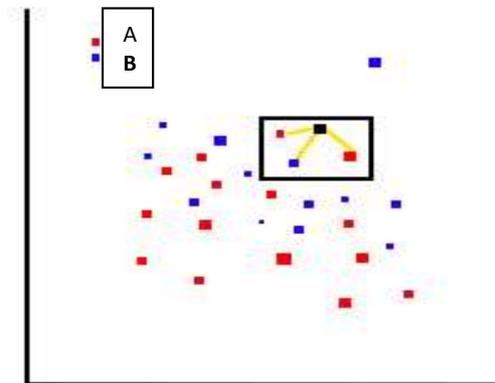
x_2 :data uji

d :jarak

p :dimensi data

i :variable data

Berikut gambar 2.3 ilustrasi KNN dibawah ini.



Gambar 2.3 Ilustrasi KNN

Dari gambar ilustrasi diatas, ada beberapa jumlah poin yang terdiri dari 2 poin yaitu poin A berwarna merah dan poin B berwarna biru. Misalkan ada data baru yang berwarna hitam yang akan diprediksi kelasnya dengan algoritma KNN. Dari contoh diatas, nilai K yang digunakan adalah 3. Setelah dihitung dari jarak warna hitam ke data poin lainnya, ada 3 titik terdekat yaitu 2 titik warna merah dan 1 titik warna biru. Seperti yang sudah digambarkan dalam ilustrasi kotak warna hitam, maka dapat disimpulkan data baru poin hitam masuk kekelas A data poin warna merah.

Contoh Cara Kerja Algoritma KNN

Sebagai contohnya, misalkan kita memiliki data *customer* dengan tabel 2.1 sebagai berikut.

Tabel 2.1 Data *Customer*, Menentukan Kelas Data Baru

Age	Income	Class
29	350	A
51	430	B
33	290	A
24	255	A
40	410	B
45	380	B
34	390	?

Dari data diatas ada 6 data yang sudah memiliki kelas dan yang 1 data baru dengan blok warna hijau yang harus ditentukan kelasnya. Pada tabel diatas terdapat dua kelas yaitu kelas A dan B, Age dan Income adalah variable yang nilainya tidak bisa dipengaruhi oleh variable lain dan akan dibuat untuk menghitung jarak. Sedangkan class merupakan variable yang nilainya dapat dipengaruhi oleh variable lain (Age dan Income).

Menentukan Nilai K

Pertama yang harus dilakukan adalah menentukan nilai K terlebih dahulu. Dalam penentuan nilai K tidak ada rumus pastinya. Namun ada tips yang perlu dipertimbangkan dalam menentukan nilai K yaitu jika jumlah kelas genap maka nilai K-nya ganjil dan jika jumlah kelasnya ganjil maka nilai K-nya genap. Mengapa demikian, karna logikanya seperti ini, misal ada kelas genap kelas A dan B dan mengambil nilai K-nya genap misal 4 maka ada kemungkinan hasil 4 tetangga terdekat 2 kelas A dan 2 kelas B, sehingga hasilnya akan sama. Namun jika diambil nilai K-nya ganjil 3 atau 5 maka otomatis akan ada yang lebih banyak jumlahnya.

Menghitung jarak data baru dengan masing-masing data lainnya

Data yang dihitung ada 6 data dengan menghitung jarak dengan menggunakan metode Euclidean distance, sebagai berikut.

Data 1

$$dis = \sqrt{(34 - 29)^2 + (390 - 350)^2} = 40.31$$

Data 2

$$dis = \sqrt{(34 - 51)^2 + (390 - 430)^2} = 43.46$$

Data 3

$$dis = \sqrt{(34 - 33)^2 + (390 - 290)^2} = 100.01$$

Data 4

$$dis = \sqrt{(34 - 24)^2 + (390 - 255)^2} = 135.37$$

Data 5

$$dis = \sqrt{(34 - 40)^2 + (390 - 410)^2} = 20.88$$

Data 6

$$dis = \sqrt{(34 - 45)^2 + (390 - 380)^2} = 14.48$$

Setelah selesai dihitung dan dicari nilai K-nya atau jarak terdekat, maka diambil 3 data dengan jarak terdekat. Dalam perhitungan diatas terdapat 3 data terdekat yaitu data 6, data 5 dan data 1. Untuk lebih jelasnya bisa dilihat tabel 2.2 sebagai berikut.

Tabel 2.2 Mencari Hasil Jarak Terdekat

Data	Age	Income	Jarak dengan data baru
6	29	350	14.48
5	51	430	20.88
1	33	290	40.31
2	24	255	43.46
3	40	410	100.01
4	45	380	135.37

Dari data diatas dapat disimpulkan bahwa data baru masuk kedalam data kelas B karena dari 3 tetangga atau nilai K terdekat ada 2 yang masuk kelas B dan 1 masuk kelas A, di bawah ini bisa dilihat tabel 2.3 hasil penentuan kelas dengan algoritma KNN.

Tabel 2.3 Hasil Menentukan Kelas Data Baru

Age	Income	Class
29	350	A
51	430	B
33	290	A
24	255	A
40	410	B
45	380	B
34	390	A

2.10 Matlab

Matlab merupakan singkatan dari Matrices Laboratory, Matlab sendiri dikembangkan oleh MathWork dan masuk dalam bahasa pemrograman tingkat tinggi. Matlab dikembangkan sebagai bahasa pemrograman dan sekaligus juga

sebagai alat visualisasi yang memiliki kemampuan untuk menyelesaikan beberapa kasus yang berhubungan langsung dengan keilmuan Matematika, seperti dibidang rekayasa teknik, fisika, statistika, komputasi dan juga modeling. Pengguna dapat memanfaatkan matlab untuk melakukan analisis data, mengembangkan algoritma dan juga membuat model maupun aplikasi. Dengan bahasa, tools, dan fungsi-fungsi built-in membuat pengguna lebih mudah dalam mengeksplorasi berbagai pendekatan sehingga memperoleh solusi lebih cepat dibandingkan dengan menggunakan spreadsheets atau bahasa pemrograman tradisional seperti C/C++ atau JavaTM.

Sebagai standar variabel elemen, *Matrices Laboratory* atau Matlab menggunakan konsep Array/Matriks tanpa memerlukan pendekatan array seperti bahasa pemrograman lainnya. Array sendiri merupakan sekumpulan data yang diatur dalam baris dan kolom yang diwakili oleh sebuah nama. Setiap nilai data dalam sebuah array dapat diakses dengan cara memasukan nama yang sesuai dengan array tersebut beserta posisi array dalam baris berapa dan kolom berapa. Dalam ruang lingkup pendidikan matlab menjadi alat pemrograman standart dalam bidang Matematika dan Rekayasa. Sedangkan dalam ruanglingkup industri matlab sendiri menjadi pilihan paling produktif yang digunakan untuk riset, komputasi dan analisa. Matlab sendiri memiliki beberapa keunggulan dibandingkan dengan bahasa pemrograman lain yaitu sebagai berikut:

2.10.1 Keunggulan Matlab

1. Matlab handal dalam komputasi yang terkait dengan array ataupun matriks. Keandalan dalam proses komputasi ini dapat dilihat dari tidak diperlukanya pendefinisian ukuran array/matriks dalam sebuah variabel. Ukuran atau dimensi dari variabel bertipe array secara otomatis akan menyesuaikan dengan array yang sudah ditetapkan dengan catatan ukuran variabel bertipe array ini sifatnya dinamis, dengan demikian maka tidak perlu memikirkan berapa ukuran variabel yang dibutuhkan dalam proses perhitungan.
2. Matlab merupakan Platform yang berdiri sendiri.

3. Mempunyai fungsi-fungsi bawaan (*Predefined Function*).
4. Memiliki perangkat yang independen untuk menampilkan grafik maupun gambar.
5. Berbasis GUI (Graphical User Interface).

2.10.2 Karakteristik Matlab

1. Bahasa pemrograman yang berdasarkan pada matriks (Baris dan Kolom).
2. Menyediakan banyak toolbox yang digunakan dalam aplikasi-aplikasi khusus seperti: Simulink, Neural Network, State Flow, Data Acquisition Toolbox, Image Acquisition Toolbox, Signal Processing Blockset, Communications Blockset, Fuzzy Logic dan lain sebagainya.
3. Waktu pengembangan program matlab lebih cepat dibandingkan dengan pemrograman lain seperti Fortran dan C/C++.
4. Dalam penulisan source code program, tidak harus mendeklarasikan array terlebih dahulu.

2.11 Ikan Cupang

Ikan cupang (*Ctenops vittatus*) merupakan anggota dari familia Anabantidae yang mempunyai labirin. Labirin merupakan alat pernafasan tambahan pada ikan.

Klasifikasi ikan cupang menurut Saanin (1968, 1984):

Phylum	: Chordata
Classis	: Pisces
Ordo	: Labyrinthici
Familia	: Anabantidae
Genus	: Ctenops
Species	: Ctenops vittatus

Ikan cupang merupakan ikan yang memiliki bentuk sirip ekor yang berbeda-beda seperti ekor bertipe mahkota (Crown tail), ekor penuh (Full tail) dan bertipe Slayer, dengan sirip panjang dan berwarna-warni. Keindahan bentuk sirip dan warna sangat menentukan nilai estetika dan nilai komersial ikan cupang. Secara

umum ikan cupang memiliki tubuh yang bervariasi, mulai dari pipih hingga silinder yang bersisik kasar, pangkal ekor terlihat lebar sehingga tubuhnya terlihat kokoh dan kuat, serta terdiri dari sirip pectoral (sirip insang), sirip dorsal (sirip punggung), sirip ventral (sirip perut), sirip kaudal (sirip ekor), dan sirip anal.

Ikan cupang jantan berwarna lebih cerah, siripnya terlihat mengembang dengan indah dan bentuk tubuh lebih panjang dan ramping, sedangkan cupang betina warna tubuh cenderung pucat, sirip tidak selebar cupang jantan, dan bentuk tubuh pendek dan gemuk (Atmadjaja, 2009). Ikan cupang juga memiliki alat pernafasan tambahan yang disebut dengan labirin sehingga ikan cupang dapat mengambil dan menyimpan oksigen lebih banyak. Oleh sebab itu, ikan cupang mampu hidup di perairan yang relative tenang dan sedikit oksigen. Perairan yang tenang cenderung memiliki kadar oksigen terlarut dan sedikit karena airnya tidak mengalir. Air yang mengalir cenderung mudah terpecah bagian permukaannya sehingga oksigen dapat dengan mudah masuk ke badan air (Atmadjaja, 2009).

2.11.1 Ikan Cupang *Halfmoon*

Ikan Cupang *halfmoon* sendiri ditujukan pada ikan cupang dengan tampilan sirip berbentuk seperti 'setengah bulan'. Dibandingkan diadu atau dipertarungkan, cupang jenis ini lebih banyak dirawat menjadi ikan hias. Variasi warna yang dimiliki *halfmoon* sangat beragam, sehingga terdapat tipe jenis dengan tampilan sirip atau ekor yang berbeda-beda.

Berikut ciri-ciri umum *Halfmoon*:

1. Bentuk ekor panjang dan terlihat membulat. Ekor bisa mengembang sampai 180 derajat saat ikan ingin 'pamer' menunjukkan kelebihanannya, juga saat ikan berhadapan dengan musuh atau merasa terancam.
2. Mengembangnya ekor bisa melebihi 180 derajat dan disebut dengan OHM (*Over Half Moon*).
3. Sementara mengembangnya ekor cupang *halfmoon* yang kurang dari 180 derajat disebut dengan Delta.
4. Tampilan ekor yang mengembang membuat cupang jenis *halfmoon* memiliki tipe jenis yang beragam.

5. Kebanyakan spesies *halfmoon* adalah hasil perkawinan silang yang mana membuat warnanya jadi bervariasi.
6. Umumnya *halfmoon* memiliki warna sirip terang yang membuatnya mudah dikenali.
7. Karena beberapa tipe jenisnya disesuaikan masing-masing ciri fisik dan tingkat kemampuan bertahan atau menyerang, ikan cupang ini dibedakan menjadi dua macam yaitu untuk dipertandingkan dan sebagai ikan hias biasa.



Gambar 2.4 Gambar ikan cupang *halfmoon*

2.11.2 Ikan Cupang Plakat

Ikan cupang Plakat merupakan jenis ikan cupang petarung, sesuai dengan namanya yang berasal Thailand. Plakat memiliki arti tarung.

Ciri ikan cupang plakat adalah sirip dan ekornya yang pendek, namun dengan ukuran tubuh yang cukup besar dan terlihat kekar. Sayangnya, warna pada tubuh ikan cupang plakat tidak begitu mencolok. Karena merupakan ikan petarung, plakat akan bergerak sangat agresif saat melihat keberadaan ikan cupang lainnya.



Gambar 2.5 Gambar ikan cupang plakat

2.11.3 Ikan Cupang Crown Tail / Serit

Atau yang juga disebut cupang serit. Ini adalah jenis cupang pertama yang dibudidayakan di Indonesia. Bentuk siripnya seperti jarum yang membentuk serit atau sisir rambut. Sampai saat ini Indonesia dikenal sebagai penghasil cupang jenis ini. Sejarahnya, cupang ini dikembangkan oleh Ahmad Yusuf, peternak asal Jakarta Timur pada 1997. Dia menyilangkan cupang lokal dengan indukan impor Thailand berwarna biru dan memiliki ekor delta dan bergerigi.



Gambar 2.6 Gambar ikan *cupang crown tail*/serit