

BAB I

PENDAHULUAN

1.1 Latar Belakang

Mangga atau mempelam termasuk kedalam tanaman sejenis buah, demikian pula nama pohonnya yang biasa disebut buah mangga. Tanaman ini memiliki nama ilmiah yakni *Mangifera Indica*. Sebuah pohon mangga dapat digolongkan kedalam tumbuhan yang cukup tinggi dengan struktur batangnya (*habitus*) termasuk kelompok *arboreus*, yaitu tumbuhan berkayu yang mempunyai tinggi pohon bisa mencapai lebih dari 5 meter. Asal usul penamaan "mangga" berasal dari bahasa Malaya yakni *maanga*. Istilah ini kemudian dibawa ke Eropa oleh orang-orang Portugis dan pengucapannya berubah menjadi mangga (bahasa Portugis), *mango* (bahasa Inggris) dan lain-lain. Di tanah air pohon mangga ini dikenal pula dalam berbagai bahasa daerah, seperti pohon pelem atau poh dan lain – lain.

Terdapat puluhan jenis atau bahkan lebih tanaman mangga yang ada di Indonesia dan bahkan tersebar di seluruh dunia. Berbagai macam jenis tanaman mangga itu tentu memiliki ciri fisik serta kandungan yang berbeda di masing – masing jenis buahnya. Salah satu cara mengetahui jenis tanaman mangga adalah dengan mengidentifikasi daunnya. Tentunya pada masing – masing jenis tanaman mangga juga memiliki jenis daun dengan ciri fisik atau tekstur yang berbeda satu sama lain. Dengan melihat permasalahan diatas maka dari itu diperlukan penelitian untuk memprediksikan atau mendeteksi jenis daun mangga menggunakan sebuah *tools* pengolahan citra dengan metode tertentu.

Algoritma *K-Nearest Neighbors* atau KNN adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data pembelajaran diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi data

pembelajaran. Sebuah titik pada ruang ini ditandai kelas c jika kelas c merupakan klasifikasi yang paling banyak ditemui pada k buah tetangga terdekat titik tersebut. Dekat atau jauhnya tetangga biasanya dihitung berdasarkan jarak Euclidean.

Pada fase pembelajaran, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi dari data pembelajaran. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk data test (yang klasifikasinya tidak diketahui). Jarak dari vektor yang baru ini terhadap seluruh vektor data pembelajaran dihitung, dan sejumlah k buah yang paling dekat diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak dari titik-titik tersebut.

Berdasarkan uraian diatas, maka penelitian ini bertujuan untuk memecahkan suatu masalah yakni deteksi jenis daun mangga menggunakan fitur *Centroid Contour Distance* (CCD) dan metode *K-Nearest Neighbors*.

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan diatas, maka rumusan masalah yang ada didalam penelitian ini adalah bagaimana merancang suatu aplikasi atau sistem yang dapat mendeteksi jenis daun mangga menggunakan fitur *Centroid Contour Distance* (CCD) dan metode *K-Nearest Neighbors*.

1.3 Batasan Masalah

Dalam penelitian ini terdapat beberapa batasan masalah, yakni sebagai berikut :

- 1) Sistem yang dibangun pada aplikasi kali ini hanya difokuskan pada daun mangga gadung, jiwo dan manalagi
- 2) Citra masukan adalah sebuah citra dengan syarat:
 - a) Menggunakan foto berwarna (bukan citra *grayscale*).
 - b) Foto harus memiliki ukuran minimum 500 x 500 piksel dan maksimum 4000 x 4000 piksel.
 - c) Citra tidak berisi obyek selain daun mangga.

- d) Sebelum pengambilan foto daun terlebih dahulu dipotong tangkainya hingga ke pangkal.
 - e) Posisi pangkal daun dalam foto berada di sebelah kanan atau kiri & bagian yang nampak dalam foto adalah yang berwarna hijau terang
 - f) Daun yang difoto dalam kondisi tidak terlalu berbayang & utuh (bukan daun kering atau yang dalam kondisi sobek)
 - g) Background foto daun harus rata dan berwarna putih.
 - h) Satu citra berisi satu jenis daun.
 - i) Semakin baik kondisi maka data akan semakin akurat & tepat
- 3) Metode yang digunakan adalah *K-Nearest Neighbors* (KNN)
- 4) Variabel yang digunakan adalah bentuk *Centroid Contour Distance* (CCD).

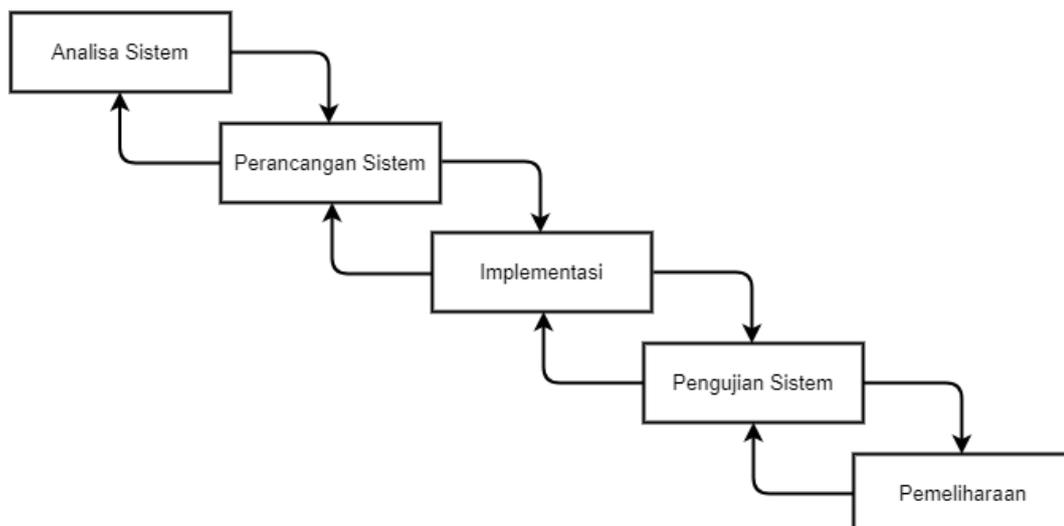
1.4 Tujuan

Tujuan dari penelitian ini adalah membuat suatu aplikasi atau sistem yang mampu mendeteksi jenis daun mangga menggunakan fitur *Centroid Contour Distance* (CCD) dan metode *K-Nearest Neighbors*.

1.5 Metode Penelitian

1.5.1. Metode Pengembangan Sistem

Penelitian ini menggunakan metode *waterfall* karena tahap demi tahap yang dilalui harus menunggu selesainya tahap sebelumnya dan berjalan berurutan. Metode ini merupakan metode yang sering digunakan oleh penganalisa sistem pada umumnya. Inti dari metode *waterfall* adalah pengerjaan dari suatu sistem dilakukan secara berurutan atau secara linear. Jadi setiap tahap harus diselesaikan terlebih dahulu secara penuh sebelum diteruskan ke tahap berikutnya untuk menghindari terjadinya pengulangan tahapan.



Gambar 1.1 Diagram Alir Secara Waterfall

a. Analisa sistem

Pada analisa kebutuhan sistem membahas beberapa kebutuhan dan atau persyaratan terkait dengan input, proses dan output, Kebutuhan atau persyaratan ini diperoleh berdasarkan data yang diperoleh dari petani.

b. Perancangan sistem dan perangkat lunak

Proses perancangan atau desain meliputi penyusunan diagram aliran data (DFD), flowchart sistem, dan struktur tabel di database serta perancangan antar muka program (*user interface*) untuk keperluan interaksi sistem dengan pengguna (*user*). Secara umum desain perangkat lunak mempertimbangkan dari sisi pengguna, yaitu tampilan antar muka dan proses perhitungan atau komputasi dengan benar.

c. Implementasi.

Pada tahap ini perancangan perangkat lunak direalisasikan sebagai serangkaian program atau unit program. Proses penulisan program dilakukan secara modular sesuai dengan pembagian blok-blok program pada langkah perancangan atau desain sistem menggunakan bahasa pemrograman tertentu.

d. Pengujian sistem

Unit program atau program individual diintegrasikan menjadi sebuah kesatuan sistem dan kemudian dilakukan pengujian. Pengujian ini ditujukan untuk menguji keterhubungan dari tiap-tiap fungsi perangkat lunak untuk menjamin bahwa persyaratan sistem telah terpenuhi dan untuk memastikan tidak ada error pada program yang dibuat.

e. Operasi dan pemeliharaan

Sistem yang diterapkan kepada *user* dan pemeliharaan mencakup koreksi dari beberapa kesalahan yang tidak diketemukan pada tahap sebelumnya, perbaikan atas implementasi unit sistem dan pengembangan pelayanan sistem, sementara persyaratan-persyaratan baru ditambahkan serta menambahkan jika ada perkembangan terbaru pada program.

1.5.2. Pengumpulan Data

Data yang diambil untuk menentukan gambar daun mangga pada penelitian hanya diperoleh dari data primer

a. Data Primer

- Melakukan pemotretan obyek daun langsung di lapangan

- Studi literature melalui buku

1.5.3. Studi Pustaka

- a. Mempelajari sistem pengolahan citra dari buku, jurnal, paper maupun penelitian penelitian sebelumnya
- b. Mempelajari teori yang berhubungan dengan segala aspek teknik yang digunakan dalam pengolahan citra dan ekstraksi fitur.
- c. Mengusulkan satu sistem deteksi gambar daun untuk membedakan background dengan obyek.

1.5.4. Analisis dan Perancangan Sistem

- a. Merancang sistem perbaikan kualitas citra sebagai langkah awal untuk mendapatkan kualitas citra terbaik untuk diproses di tahap selanjutnya.
- b. Merancang metode untuk mendeteksi warna dan garis tepi.
- c. Merancang metode untuk ekstraksi fitur menjadi datasheet yang akan diolah menggunakan algoritma klasifikasi. Pada penelitian ini penulis menggunakan fitur *Centroid Contour Distance* untuk analisa fitur bentuk pada daun mangga.
- d. Merancang sistem pelatihan dan klasifikasi menggunakan *K-Nearest Neighbors*.

1.5.5. Implementasi

Melakukan implementasi hasil perancangan pada suatu program aplikasi. Bahasa pemrograman yang digunakan adalah Python serta menggunakan *library* OpenCV untuk pemrosesan citra pada komputer personal dengan sistem operasi Microsoft Windows 8.1.

1.5.6. Pengujian dan Analisis Hasil

Sistem diuji untuk mengetahui hasil keluaran dari sistem. Citra masukan adalah citra yang diperoleh dari hasil pemotretan sendiri. Citra masukan tersebut disimpan dalam folder yang memisahkan antara beberapa jenis daun, sehingga mempermudah proses pengujian. Uji coba dilakukan untuk menjawab tujuan dan

permasalahan dalam penelitian ini. Uji coba yang dilakukan adalah uji coba pelatihan dan klasifikasi dengan metode K-NN

1.5.7. Penulisan Laporan Tugas Akhir

Menulis semua hasil dari perancangan dan pembuatan dari sistem identifikasi gambar jenis mangga berdasarkan warna dan bentuk daun.

1.6 Sistematika Penulisan

Sistematika penulisan yang akan dibuat dalam tugas akhir ini disusun beberapa bab, yang dijelaskan sebagai berikut :

BAB I Pendahuluan.

Bab ini menjelaskan permasalahan yang ada beserta solusi yang digunakan untuk menyelesaikan permasalahan tersebut. Penjelasan tersebut akan di bagi-bagi menjadi sub bab yaitu latar belakang, perumusan masalah, batasan masalah, tujuan dari pembuatan tugas akhir, metodologi penelitian, dan sistematika penulisan beserta jadwal pelaksanaan.

BAB II Tinjauan Pustaka

Bab ini membahas tentang teori-teori yang berhubungan dengan fakta atau kasus yang sedang dibahas. Disamping ini juga dapat disajikan mengenai pendapat yang berhubungan dan benar-benar bermanfaat sebagai bahan untuk melakukan analisis terhadap fakta atau kasus yang sedang diteliti pada bab IV.

BAB III Landasan Teori

Bab ini membahas tentang dasar teori yang digunakan dalam penelitian tugas akhir ini. Adapun yang dibahas pada bab ini adalah teori yang berkaitan dengan pembangunan deteksi jenis daun mangga menggunakan fitur *Centroid Contour Distance* (CCD) dan metode *K-Nearest Neighbors*.

BAB IV Analisis dan Sistem

Bab ini berisi analisis terhadap masalah yang diselesaikan dan perancangan sistem yang dibuat. Sistematika yang disajikan meliputi penyampaian masalah, spesifikasi kebutuhan sistem (fungsional dan non fungsional), data yang dioalah

dan contoh perhitungan, desain sistem dengan DFD (dan flowchart), desain antarmuka perangkat lunak

BAB V Implementasi Sistem

Bab ini berisi implementasi/realisasi dari desain di-bab 4 dan implementasi sistem dengan menyajikan antarmuka perangkat lunak koding program yang dimasukkan pada antarmuka juga perlu dimasukkan.

BAB VI Pengujian dan Pembahasan Hasil

Bab ini berisi tentang pengujian yang dilakukan, tampilan program dan analisa hasilnya. Bagian pertama bab ini dapat berisi printscreen antar muka ketika diuji dengan variasi masukkan juga perlu ditampilkan. Model pengujian juga disampaikan, misalnya pada kasus klasifikasi bagaimana pemecahan antara data uji pada kasus *clustering* menggunakan beberapa *cluster*. Jumlah pengujian yang dilakukan pada kasus klasifikasi sebaiknya minimal 30 dengan sejumlah variasi pengujian. Di bagian kedua, disajikan hasil-hasil pengujian dalam bentuk tabel atau grafik, dan analisis keseluruhan pengujian yang sudah dilakukan.

BAB VII Penutup

Bab ini berisi kesimpulan dan saran. Kesimpulan berisi tentang ringkasan hasil implementasi dan pengujian atau analisa, kesimpulan harus dilakukan dengan tajam dan jelas. Sedangkan saran berisi tentang usulan-usulan terhadap penyelesaian lebih lanjut permasalahan yang dikaji.

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terkait

Dari hasil penelitian sebelumnya dengan judul :

Eko Prasetyo, 2014, *Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab*. Meskipun riset ilmu data mining masih terbilang muda, tetapi penerapan dan manfaat yang didapat sangat luas dan besar. Bidang – bidang seperti bisnis, medis, pendidikan, astronomi, meteorologi dan sebagainya merupakan bidang – bidang yang sangat membutuhkan data mining. Untuk itulah buku ini dibuat agar menjadi panduan yang jelas, lengkap dan sistematis disertai contoh aplikasi MATLAB yang mudah pula untuk diikuti. Secara garis besar, buku ini diawali dengan membahas konsep – konsep dasar data mining, meliputi pendahuluan, pemrosesan awal dan fitur. Bagian kedua membahas metode – metode klasifikasi meliputi Decision Tree, Artificial Neural Network, Support Vector Machine dan Nearest Neighbour. Bagian ketiga membahas metode – metode analisis cluster, meliputi partisi dan hirarki. Di bagian akhir membahas evaluasi data mining. Pembahasan topik – topik didalamnya disertai contoh penerapan dalam program menggunakan MATLAB.

Wishnu Hardi, 2015, *Mengukur Kinerja Search Engine: Sebuah Eksperimentasi Penilaian Precision And Recall Untuk Informasi Ilmiah Bidang Ilmu Perpustakaan Dan Informasi*. Bagi mereka yang bekerja di sektor informasi, search engine merupakan bagian dari aktivitas kerja sehari-hari. Seorang pustakawan rujukan, misalnya, akan memprioritaskan penggunaan search engine sebagai sarana penelusuran informasi. Pustakawan yang bergelut dengan penelusuran sumber daya informasi online, setidaknya harus memiliki dua keahlian ketika bekerja dengan search engine. Ketergantungan yang sangat tinggi terhadap search engine sering tidak diimbangi dengan pemahaman bahwa sebenarnya search engine tidak dilengkapi filter dalam pengumpulan informasi yang diinginkan oleh pengguna. Dalam dunia sistem temu kembali informasi, cara

yang paling efektif untuk menguji kehandalan sebuah search engine adalah dengan meneliti tingkat precision (ketepatan) and recall (ketepatan)-nya.

Febri Liantoni, Hendro Nugroho, 2015, *Klasifikasi Daun Herbal Menggunakan Metode Naïve Bayes Classifier Dan K-Nearest Neighbor.*

Penelitian ini membahas tentang Perkembangan ilmu tanaman telah mengalami kemajuan yang pesat, khususnya ilmu mengenai tanaman herbal. Tanaman herbal memiliki banyak manfaat bagi kehidupan manusia yaitu sebagai penyediaan oksigen, bahan makanan, obat-obatan, maupun bahan kosmetika. Untuk mengetahui jenis-jenis tanaman herbal dapat dilakukan dengan proses klasifikasi. Klasifikasi tanaman herbal dapat dilakukan dengan cara mengidentifikasi bentuk citra daun dari tanaman tersebut. Proses klasifikasi tanaman herbal memerlukan ekstraksi fitur dari bentuk dari tanaman tersebut. Pada penelitian ini, fitur invariant moment dan fitur geometri digunakan untuk ekstraksi fitur daun herbal. Fitur yang digunakan berjumlah 10 fitur. Ada beberapa macam metode klasifikasi yang biasa digunakan. Pada penelitian ini metode klasifikasi yang digunakan adalah metode Naïve Bayes Classifier dan *K-Nearest Neighbor* (KNN). Metode Naïve Bayes Classifier merupakan metode Bayesian Learning yang paling cepat dan sederhana. Sedangkan metode KNN dapat melakukan klasifikasi dengan cepat berdasarkan jarak terdekat diantara objek data. Berdasarkan hasil uji coba yang dilakukan, penggunaan metode Naïve Bayes Classifier didapatkan nilai akurasi sebesar 75%, sedangkan dengan menggunakan metode *K-Nearest Neighbor* didapatkan nilai akurasi sebesar 70,83%. Hal ini menunjukkan bahwa kinerja metode Naïve Bayes Classifier lebih baik dibandingkan metode KNN.

Hanang Wijayanto, 2013, *Klasifikasi Batik Menggunakan Metode K-Nearest Neighbour Berdasarkan Gray Level Co-Occurrence Matrices (GlcM).*

Batik adalah sejenis kain tertentu yang dibuat khusus dengan motif-motif yang khas yang langsung dikenali masyarakat umum. Salah satu permasalahan pada batik adalah batik memiliki motif dan warna yang sangat beragam, sehingga sulit untuk klasifikasi batik ke dalam kelas tertentu. Penelitian ini dilakukan untuk mengklasifikasikan batik ke dalam kelas-kelas berdasarkan daerah asal batik

sehingga mempermudah dalam pengenalan batik dan pemahaman tentang batik. Metode yang digunakan adalah gray level co-occurrence matrices untuk ekstraksi ciri tekstur, sedangkan untuk menentukan kedekatan antara citra uji dengan citra latih menggunakan metode *k-nearest neighbor* berdasarkan fitur tekstur dari citra batik yang diperoleh. Fitur-fitur tekstur dicari menggunakan metode GLCM berdasarkan sudut 0° , 45° , 90° dan 135° . Metode evaluasi menggunakan confusion matrix untuk mengukur tingkat akurasi dalam proses klasifikasi. Dari hasil penelitian yang menggunakan 100 citra batik dengan 5 kelas yaitu Cirebon, Jakarta, Pekalongan, Solo dan Yogyakarta menghasilkan tingkat akurasi tertinggi pada uji coba 3 sudut 0 sebesar 57,50 % dan terendah pada uji coba 6 sudut 90 sebesar 20 %.

Abdurrasyid Hasim, 2014, *Perbandingan Algoritma Centroid Contour Gradient Dan Centroid Contour Distance Untuk Pengenalan Bentuk Daun.* Penelitian ini membandingkan algoritma CCG (*Centroid Contour Gradient*) dan CCD (*Centroid Contour Distance*) untuk ekstraksi fitur dalam pengenalan bentuk daun. CCG dan CCD adalah algoritma untuk merepresentasikan bentuk dengan pendekatan berbasis kontur (contour-based). CCG menghitung nilai gradient antar titik sepanjang tepi daun pada setiap interval sudut tertentu sedangkan CCD menghitung jarak titik tengah terhadap titik-titik tepi. Bentuk daun yang digunakan dalam penelitian ini adalah ellipsis, cordate, ovate, dan lanceolate. Data yang digunakan sebanyak 200 citra daun dengan jumlah citra masing-masing kelas sebanyak 50. Probabilistic Neural Network digunakan untuk mengklasifikasi bentuk daun. Didapatkan akurasi terbaik CCD sebesar 96.67%, jauh lebih besar dibanding akurasi terbaik CCG sebesar 60.00%.

Mazid kamal, Ruri Suko Basuki, 2013, *Segmentasi Citra Daun Tembakau Berbasis Deteksi Tepi Menggunakan Algoritma Canny.* Kualitas tembakau memiliki peran penting dalam menentukan kualitas produksi rokok. Tembakau terdiri dari berbagai kelas atau bahan baku yang dicampur untuk membuat rokok. Klasifikasi daun tembakau bertujuan mengukur dan menganalisis kualitas tembakau yang dapat dikelompokkan ke dalam kelas tertentu. Tetapi

seorang siswa kelas karakteristik yang sering membuat kesalahan yang disebabkan oleh kelelahan, keadaan emosi, visi dan pencahayaan. Karena faktor-faktor ini dilakukan segmentasi citra daun tembakau berdasarkan deteksi tepi menggunakan metode tembakau Canny. Citra awalnya dilakukan pre-processing untuk melakukan segmentasi citra untuk mendapatkan tepi. Selanjutnya akan di ekstraksi untuk diakui berdasarkan ukuran, bentuk dan tekstur untuk diklasifikasikan dan dibuat aplikasi untuk menentukan tingkat daun tembakau. Klasifikasi kelas umumnya dimulai dari proses akuisisi data, pra-pemrosesan dan pasca-pemrosesan. Deteksi tepi Canny dapat mendeteksi tepi sebenarnya dengan tingkat kesalahan minimum, menghasilkan tepi gambar yang optimal.

Febri Liantoni, 2015, *Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode K-Nearest Neighbor*. Tanaman adalah bagian terpenting dalam kehidupan di bumi sebagai pemasok oksigen untuk bernafas, bahan makanan, bahan bakar, obat-obatan dan banyak lagi. Tanaman dapat diklasifikasikan berdasarkan bentuk daunnya. Proses klasifikasi diperlukan fitur ekstraksi data yang baik, sehingga perlu memperbaiki proses fitur pada level pre-processing. Menggabungkan filter median dan erosi gambar digunakan untuk memperbaiki proses fitur. Sedangkan untuk ekstraksi fitur digunakan metode invarian moment. Dalam penelitian ini digunakan klasifikasi daun berdasarkan bentuk daun daun. Metode *K-Nearest Neighbor* (KNN) digunakan untuk proses klasifikasi daun. Metode KNN dipilih karena metode ini dikenal cepat dalam data pelatihan, efektif untuk data pelatihan yang besar, sederhana dan mudah dipelajari. Pengujian hasil klasifikasi daun dari citra yang ada pada dataset telah dibangun untuk mendapatkan nilai akurasi sebesar 86,67%..

Eny Maria, Yulianto, Yunita Putri Arinda, Jumiatty, Palma Nobel, 2018, *Segmentasi Citra Digital Bentuk Daun Pada Tanaman Di Politani Samarinda Menggunakan Metode Thresholding*. Untuk mengenal jenis-jenis bentuk tulang daun, telah dibuat sistem untuk membandingkan garis tulang daun beserta garis tepi daun menggunakan metode thresholding. Prosesnya dimulai dengan menginput citra digital daun atau tanaman, selanjutnya dikonversi ke citra

grayscale. Kemudian dilakukan proses segmentasi terhadap citra grayscale. Selanjutnya, dipilih hasil segmentasi dan ditandai dengan proses penajaman garis tepi menggunakan operator LOG. Proses terakhir adalah membuat histogram terhadap hasil proses penajaman garis tepi. Hasil segmentasi berhasil membandingkan dan menggolongkan bentuk tulang daun yang diambil menggunakan kamera ponsel pada tanaman di Politeknik Pertanian Negeri Samarinda dengan cara thresholding dengan hasil segmentasi citra daun yang telah digolongkan berdasarkan bentuk tulang daunnya dengan cara thresholding pula. Keseluruhan proses ini dilakukan dengan menggunakan MATLAB 2008.

Nur Hidayatullah, 2016, *Identifikasi Jenis Mangga Berdasarkan Warna Dan Bentuk Daun Menggunakan Metode ANN Voted Perceptron.* Bagi sebagian masyarakat, membedakan jenis bibit mangga merupakan hal yang sulit dilakukan mengingat bentuk daun yang hampir sama. Masyarakat baru bisa mengenali jenis mangga ketika mangga sudah berbuah dan itu masih harus menunggu minimal 6 bulan untuk berbuah. Voted Perceptron adalah salah satu metode klasifikasi linier yang arsitekturnya menggunakan layer input dan layer output dengan nilai vektor bobot tertentu untuk memaksimalkan margin antara dua kelas data. Penelitian ini menggunakan nilai Mean, momen nth, dan standart deviasi dari fitur warna green serta nilai compactness dan circularity untuk deskripsi bentuk dari citra daun. Dengan data latih sejumlah 30 data untuk masing masing kelas dengan perbandingan 20 data latih dan 10 sebagai data uji. Hasil penelitian ini didapat nilai akurasi sebesar 60% dan error sebesar 40% serta nilai precision untuk kelas mangga curut sebesar 0, kelas mangga gadung sebesar 0.6428, kelas mangga manalagi sebesar 0, kelas mangga madu sebesar 0.5263, kelas mangga golek sebesar 0.7778, kelas tidak terdaftar sebesar 0.625. Sedangkan nilai recall untuk kelas mangga curut sebesar 0, kelas mangga gadung sebesar 0.9, kelas mangga manalagi sebesar 0, kelas mangga madu sebesar 1.0, kelas mangga golek sebesar 0.7, kelas tidak terdaftar sebesar 1.0.

Feri Wibowo, Dimara Kusuma Hakim, Sigit Sugiyanto, 2018, *Pendugaan Kelas Mutu Buah Pepaya Berdasarkan Ciri Tekstur Glcm*

Menggunakan Algoritma K-Nearest Neighbors. Proses klasifikasi mutu buah pepaya dengan cara konvensional menggunakan visual mata manusia memiliki kelemahan di antaranya yaitu membutuhkan tenaga lebih banyak untuk memilah, tingkat persepsi manusia yang berbeda, tingkat konsistensi manusia dalam menilai mutu buah tidak menjamin karena manusia dapat mengalami kelelahan. Penelitian ini bertujuan merancang dan menyusun program pengolahan citra digital dan algoritma *k-Nearest Neighbor* untuk klasifikasi pematuan buah pepaya (Carica Papaya L) Calina IPB-9 ke dalam tiga kelas mutu yaitu kelas Super, A, dan B. Fitur tekstur yang diekstrak meliputi nilai energy, entropy, kontras, homogeneity, invers difference moment, variance, dan dissimilarity yang didapatkan berdasarkan GLCM (gray level cooccurrence matrices). Fitur-fitur tersebut dijadikan sebagai input pada algoritma *k-Nearest Neighbor* untuk menghitung jarak. Hasil pengujian menggunakan jumlah k tetangga 9 menunjukkan tingkat akurasi sebesar 88,88%.

Tabel 2.1 Perbandingan Tinjauan Pustaka

No	Nama Peneliti	Judul Penelitian	Pendekatan Penyelesaian Masalah	Tujuan Penelitian
1	Eko Prasetyo (2014)	Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab	Menjelaskan tentang data mining dan banyak metode, salah satunya <i>K-Nearest Neighbour (KNN)</i>	Menjadi bahan referensi dan panduan yang jelas, lengkap dan sistematis dalam bidang data mining disertai contoh aplikasi MATLAB yang mudah untuk diikuti
2	Wishnu Hardi (2015)	Mengukur Kinerja Search Engine: Sebuah Eksperimentasi Penilaian Precision And Recall Untuk Informasi Ilmiah Bidang Ilmu Perpustakaan Dan Informasi.	Menggunakan metode klasifikasi pengujian <i>Precision And Recall</i>	Mengukur kinerja Search Engine dengan menggunakan Penilaian Precision And Recall untuk informasi ilmiah bidang ilmu perpustakaan dan informasi

3	Febri Liantoni, Hendro Nugroho (2011)	Klasifikasi Daun Herbal Menggunakan Metode Naïve Bayes Classifier Dan <i>K-Nearest Neighbor</i>	Menggunakan metode <i>Naïve Bayes Classifier</i> Dan <i>K-Nearest Neighbour</i>	Dapat melakukan klasifikasi tanaman herbal dengan cara mengidentifikasi bentuk citra daun dari tanaman tersebut
4	Hanang Wijayanto (2013)	Klasifikasi Batik Menggunakan Metode <i>K-Nearest Neighbour</i> Berdasarkan Gray Level Co-Occurrence Matrices (GlcM)	Menggunakan metode <i>K-Nearest Neighbour (KNN)</i>	Dapat mengklasifikasikan batik ke dalam kelas-kelas berdasarkan daerah asal batik sehingga mempermudah dalam pengenalan batik dan pemahaman tentang batik
5	Abdurrasyid Hasim (2014)	Perbandingan Algoritma <i>Centroid Contour Gradient</i> Dan <i>Centroid Contour Distance</i> Untuk Pengenalan Bentuk Daun	Membandingkan Algoritma <i>Centroid Contour Gradient</i> dan <i>Centroid Contour Distance</i>	Dapat mengimplementasikan dan membandingkan kinerja algoritma <i>Centroid Contour Gradient</i> dan <i>Centroid Contour Distance</i> dalam mengenali bentuk daun

6	Mazid kamal, Ruri Suko Basuki (2013)	Segmentasi Citra Daun Tembakau Berbasis Deteksi Tepi Menggunakan Algoritma Canny	Menggunakan metode <i>Algoritma Canny</i>	Dapat mengklasifikasikan daun tembakau dan menganalisis kualitas tembakau sehingga dapat dikelompokkan ke dalam kelas tertentu
7	Febri Liantoni (2015)	Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode <i>K-Nearest Neighbor</i>	Menggunakan metode <i>K-Nearest Neighbour (KNN)</i>	Dapat melakukan klasifikasi daun berdasarkan bentuk daun dan perbaikan fitur citra menggunakan metode KNN
8	Eny Maria, Yulianto, Yunita Putri Arinda, Jumiaty, Palma Nobel (2018)	Segmentasi Citra Digital Bentuk Daun Pada Tanaman Di Politani Samarinda Menggunakan Metode Thresholding	Menggunakan metode <i>Thresholding</i>	Untuk mengenal jenis-jenis daun pada tanaman di Politani dengan membandingkan garis tulang daun beserta garis tepi daun menggunakan metode thresholding

9	Nur Hidayatullah (2016)	Identifikasi Jenis Mangga Berdasarkan Warna Dan Bentuk Daun Menggunakan Metode ANN Voted Perceptron	Menggunakan metode <i>ANN Voted Perceptron</i>	Menggunakan <i>Metode voted perceptron</i> agar dapat membedakan daun mangga, mengingat bentuk fisik daun yang hampir sama
10	Feri Wibowo, Dimara Kusuma Hakim, Sigit Sugiyanto (2018)	Pendugaan Kelas Mutu Buah Pepaya Berdasarkan Ciri Tekstur Glcm Menggunakan Algoritma <i>K-Nearest Neighbors</i>	Menggunakan metode <i>K-Nearest Neighbour (KNN)</i>	Untuk merancang dan menyusun program pengolahan citra digital dan algoritma <i>K-Nearest Neighbour</i> untuk klasifikasi pemutuan buah pepaya

BAB III

LANDASAN TEORI

3.1. Citra Digital

Citra adalah suatu representasi (gambaran), kemiripan, atau imitasi dari suatu objek. Citra terbagi 2 yaitu ada citra yang bersifat analog dan ada citra yang bersifat digital. Citra analog adalah citra yang bersifat kontinu seperti gambar pada monitor 20elewise, foto sinar X, hasil CT scan dll. Sedangkan pada citra digital adalah citra yang dapat diolah oleh computer (T.Sutoyo, 2009).

Sebuah citra digital dapat mewakili oleh sebuah matriks yang terdiri dari M kolom N baris, dimana perpotongan antara kolom dan baris disebut piksel (piksel = *picture element*), yaitu elemen terkecil dari sebuah citra. Piksel mempunyai dua parameter, yaitu koordinat dan intensitas atau warna. Nilai yang terdapat pada koordinat (x, y) adalah f(x, y), yaitu besar intensitas atau warna dari piksel di titik itu. Oleh sebab itu, sebuah citra digital dapat ditulis dalam bentuk matriks berikut.

$$(x,y) = \begin{bmatrix} f(0,0) & \cdots & f(0, m) \\ \vdots & \ddots & \vdots \\ f(n, 0) & \cdots & f(n, m) \end{bmatrix} \dots\dots\dots(3.1)$$

Berdasarkan gambaran tersebut, secara matematis citra digital dapat dituliskan sebagai fungsi intensitas f (x, y), dimana harga x (baris) dan y (kolom) merupakan koordinat posisi dan f(x, y) adalah nilai fungsi pada setiap titik (x, y) yang menyatakan besar intensitas citra atau tingkat keabuan atau warna dari piksel di titik tersebut. Pada proses digitalisasi (sampling dan kuantitas) diperoleh besar baris M dan kolom N hingga citra membentuk matriks M x N dan jumlah tingkat keabuan piksel G (T.Sutoyo, 2009).

3.1.1. Pengolahan Citra

Pengolahan citra digital adalah sebuah disiplin ilmu yang mempelajari hal-hal yang berkaitan dengan perbaikan kualitas gambar (peningkatan kontras, transformasi warna, restorasi citra), transformasi gambar (rotasi, translasi, skala, transformasi geometrik), melakukan pemilihan citra ciri (*feature images*) yang optimal untuk tujuan analisis, melakukan proses penarikan informasi atau deskripsi objek atau pengenalan objek yang terkandung pada citra, melakukan kompresi atau reduksi data untuk tujuan penyimpanan data, transmisi data, dan waktu proses data. Input dari pengolahan citra adalah citra, sedangkan outputnya adalah citra hasil pengolahan (T.Sutoyo, 2009).

3.1.1.1. Deteksi Tepi

Deteksi tepi (*Edge Detection*) pada suatu citra adalah suatu proses yang menghasilkan tepi – tepi dari objek - objek citra, tujuannya adalah

1. Untuk menandai bagian yang menjadi detail citra.
2. Untuk memperbaiki detail dari citra yang kabur, yang terjadi karena error atau adanya efek dari proses akuisisi citra.

Ide dasar dibalik deteksi tepi adalah untuk mencari tempat di dalam citra dimana intensitas berubah secara tepat, menggunakan satu dari dua kriteria umum berikut:

1. Mencari tempat di mana turunan pertama intensitas lebih besar dalam jarak daripada *threshold* yang ditetapkan.
2. Mencari tempat dimana turunan kedua dari intensitas mempunyai *zero crossing* (Prasetyo, 2012)

a) Deteksi Tepi Sobel

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

$$\text{a. } G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \quad \text{b. } G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

Gambar 3.1 Deteksi Tepi Sobel

Deteksi tepi sobel menggunakan mask dalam gambar (a) dan (b) untuk memperkirakan secara digital turunan pertama G_x dan G_y . Dengan kata lain, gradien pada titik pusat ketentagaan dihitung oleh detektor sobel sebagai berikut :

$$g = [G_x^2 + G_y^2]^{1/2}$$

$$= \{ [(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)]^2 + [(z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)]^2 \}^{1/2}$$

Maka dapat dikatakan bahwa piksel pada lokasi (x, y) adalah piksel tepi jika $g \geq T$ pada lokasi tersebut, dimana T adalah *threshold* yang di tentukan (Prasetyo, 2011).

b) Deteksi Tepi Prewitt

-1	-1	-1
0	0	0
1	1	1

-1	0	1
-1	0	1
-1	0	1

$$\text{c. } G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \quad \text{d. } G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

Gambar 3.2 Deteksi Tepi Prewitt

Detektor tepi Prewitt menggunakan mask dalam gambar (c) dan (d) untuk memperkirakan secara digital turunan pertama G_x , dan G_y .

Parameter pada fungsi ini sama dengan parameter Sobel. Detektor Prewitt agak lebih mudah untuk diimplementasikan secara komputasi dari pada detector Sobel. Detektor ini cenderung memproduksi hasil yang lebih ber-*noise* (dalam Sobel diatasi dengan nilai 2 untuk *smoothing*) (Prasetyo, 2011).

c) Deteksi Tepi Robert

-1	0
0	-1

0	-1	
1	0	

e. $G_x = z_9 + z_5$

f. $G_x = z_8 + z_6$

Gambar 3.3 Deteksi Tepi Robert

Detektor tepi Robert menggunakan mask dalam gambar (e) dan (f) untuk memperkirakan secara digital turunan pertama G_x , dan G_y .

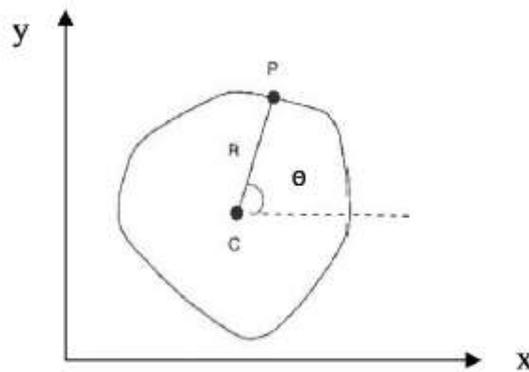
Parameter pada fungsi ini sama dengan parameter Sobel. Detektor Robert adalah satu dari detektor tepi yang paling tua dalam pengolahan citra digital, Detektor ini yang paling sederhana. Detektor ini tidak lebih banyak digunakan dalam implementasi *hardware* di mana kesederhanaan dan kecepatan adalah faktor dominan (Prasetyo, 2011).

3.2. Centroid Contour Distance

Untuk merepresentasikan bentuk suatu objek, terdapat dua teknik pendekatan yaitu berbasis kontur (*contour-based*) dan berbasis wilayah (*region-based*). Pendekatan berbasis kontur hanya memanfaatkan informasi yang terdapat pada kontur tepi sedangkan pendekatan berbasis wilayah melibatkan seluruh bagian dari suatu objek (Zhang dan Lu 2004). Algoritma *Centroid Contour Gradient* (CCG) dan *Centroid Contour Distance* (CCD) merupakan algoritma untuk merepresentasikan bentuk dengan pendekatan berbasis kontur. Untuk mengekstraksi fitur, kedua algoritma tersebut hanya melibatkan bagian kontur dari objek. Namun dalam kesempatan ini penulis hanya akan menggunakan fitur

Centroid Contour Distance untuk bisa merepresentasikan daun mangga yang akan diuji.

CCD menghitung jarak antara tiap titik tepi terhadap titik tengah. Diilustrasikan pada Gambar 3.4 bahwa CCD mendeteksi titik pada tepi objek pada sudut tertentu. Titik P pada tepian ditentukan oleh titik tengah C, jarak antara titik P dan titik tengah C, dan sudut θ . Dalam kesempatan ini penulis menggunakan sudut θ sebesar 1 derajat. Berarti proses pencarian fungsi CCD pada masing – masing daun dilakukan sebanyak 360 ($N=360$). Pada persamaan 3.2 juga ditunjukkan fungsi *Centroid-distance* (R_i), yaitu jarak antara titik tepi (x_i, y_i) dari titik tengah (x_c, y_c) (Pahalawatta 2008).



Gambar 3.4 Ilustrasi fitur Centroid Contour Distance

$$R_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad i = (1,2,3,\dots,N) \dots\dots\dots(3.2)$$

- R_i = Fungsi CCD ke i (dilakukan sebanyak 360 fungsi)
- (X_i, Y_i) = Titik pixel pada bagian tepi daun
- (X_c, Y_c) = Titik pixel pada bagian pusat daun

3.3. Pre Processing

3.3.1. Normalisasi

Normalisasi adalah proses penskalaan nilai atribut dari data sehingga berada pada *range* tertentu. Tanpa dilakukan normalisasi, bisa jadi fitur x yang akan mendominasi fungsi biaya pada klasifikator. Setelah dinormalisasi, semua fitur akan berada pada jangkauan yang sama sehingga proporsi pengaruh pada fungsi biaya dalam klasifikator menjadi seimbang. Untuk menskalakan dalam jangkauan [0,1] dapat digunakan persamaan berikut :

$$\hat{Z}_{ik} = \frac{Z_{ik} - \min(Z_k)}{\max(Z_k) - \min(Z_k)} \dots\dots\dots(3.3)$$

Keterangan :

\hat{Z}_{ik} = nilai atribut / kriteria setelah normalisasi

Z_{ik} = nilai atribut / kriteria sebelum normalisasi

$\min(Z_k)$ = nilai terkecil

$\max(Z_k)$ = nilai terbesar

Sementara untuk menskalakan dalam jangkauan [1,-1] digunakan persamaan berikut ini :

$$\hat{Z}_{ik} = \frac{2Z_{ik} - (\max(Z_k) + \min(Z_k))}{\max(Z_k) - \min(Z_k)} \dots\dots\dots(3.4)$$

3.4. K-Nearest Neighbour

K-Nearest Neighbor Classifier (K-NN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Pengertian klasifikasi yaitu proses pengelompokan, artinya mengumpulkan benda / entitas yang sama serta memisahkan benda/entitas yang tidak sama (Prasetyo, 2012). Dekat atau jauhnya tetangga biasanya dihitung dengan jarak *Euclidean*. Teknik ini sederhana dan

dapat memberikan akurasi yang baik terhadap hasil klasifikasi. Data pembelajaran diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data.

Ruang dimensi dibagi menjadi bagian-bagian berdasarkan klasifikasi data pembelajaran. Nilai k yang terbaik untuk algoritma ini tergantung pada data; secara umum nilai k yang tinggi akan mengurangi efek *noise* pada klasifikasi, akan tetapi membuat batasan antara setiap klasifikasi menjadi lebih kabur. Nilai k yang bagus dapat dipilih dengan optimasi parameter, misalnya dengan menggunakan *cross-validation*. Kasus khusus di mana klasifikasi diprediksikan berdasarkan data pembelajaran yang paling dekat (dengan kata lain, k = 1) yang biasanya disebut algoritma *nearest neighbor*.

$$d(i, j) = \sqrt{(x_{i_1} - x_{j_1})^2 + (x_{i_2} - x_{j_2})^2 + \dots + (x_{i_p} - x_{j_p})^2} \dots \dots \dots (3.5)$$

- Keterangan :
- d (i, j) = nilai jarak
 - x_i = nilai – nilai pada fitur 1
 - x_j = nilai – nilai pada fitur 2

Dalam algoritma ini, nilai k yang terbaik itu tergantung pada jumlah data. Ukuran nilai k yang besar belum tentu menjadi nilai k yang terbaik begitu juga sebaliknya (Prasetyo, 2012). Langkah – langkah untuk menghitung algoritma KNN :

1. Menentukan parameter K sebagai banyaknya jumlah tetangga terdekat dengan objek baru.
2. Menghitung jarak antar objek/data baru terhadap semua objek/data yang telah di *training*.
3. Dilakukan *Sort* atau mengurutkan hasil perhitungan tersebut mulai dari jarak terkecil hingga paling besar.
4. Tentukan tetangga terdekat berdasarkan jarak minimum ke K.
5. Tentukan kategori dari tetangga terdekat dengan objek/data.
6. Gunakan kategori mayoritas sebagai klasifikasi objek/data baru.

3.5. Contoh Klasifikasi dengan K-Nearest Neighbour

Berikut ini adalah contoh proses pelatihan untuk mencari nilai klasifikasi dari data uji yang telah didapatkan melalui proses perhitungan *Centroid Contour Distance* (CCD) sebelumnya. Pada umumnya data klasifikasi K-NN hanya ada dua yakni X1 dan X2 namun dalam penelitian kali ini bisa dipastikan data nya akan berjumlah banyak. Oleh karena itu untuk pelatihan juga ditetapkan data sejumlah 10.

Setelah diperoleh data sebanyak 10 fungsi melalui proses ekstraksi fitur CCD untuk selanjutnya diklasifikasi dan didapatkan nilai tetangga terdekatnya untuk selanjutnya dapat diketahui termasuk kedalam parameter K yang mana (bagus, sedang atau tidak bagus). Penggunaan parameter tidak terpaku hanya pada kualitas namun juga bisa untuk membedakan, contohnya daun mangga.

Dibawah ini terdapat contoh tabel data latih (sebanyak 8 data) yang diambil dari beberapa sampel daun mangga. Nilai K direpresentasikan sebagai data hasil dari perhitungan CCD. Selain itu terdapat variable parameter untuk mempermudah klasifikasi.

Contoh data uji adalah dengan nilai

$$K_1 = 6, K_2 = 5, K_3 = 8, K_4 = 8, K_5 = 10, K_6 = 12, K_7 = 9, K_8 = 7, K_9 = 6, K_{10} = 4$$

Parameter dijadikan variable : Bagus \rightarrow (A), Sedang \rightarrow (B), Tidak Bagus \rightarrow (C)

Data	K_1	K_2	K_3	K_4	K_5	K_6	K_7	K_8	K_9	K_{10}	Kategori
1	5	7	7	9	8	11	10	7	6	4	A
2	6	6	8	10	11	12	8	6	5	3	B
3	4	6	7	9	10	11	9	7	6	4	A
4	5	7	6	7	8	10	9	8	6	4	C
5	7	7	8	8	9	10	8	8	6	5	A
6	6	8	7	10	11	12	10	8	7	5	C
7	5	7	6	7	10	12	12	9	7	6	B
8	4	5	7	9	8	9	10	7	5	3	B

Tabel 3.1 Data Training

- Langkah pertama yang harus dilakukan adalah menentukan parameter K. Misalnya jumlah tertangga terdekat adalah $K = 5$. Penentuan parameter K sejumlah 5 dilakukan untuk akurasi lebih tinggi dan menghindari adanya mayoritas kategori yang sama.
- Selanjutnya hitung jarak antara data baru dengan semua data training. Pada penelitian kali ini penulis menggunakan *Euclidean Distance*.

Data 1

$$\sqrt{(5 - 6)^2 + (7 - 5)^2 + (7 - 8)^2 + (9 - 8)^2 + (8 - 10)^2 + (11 - 12)^2 + \dots + (4 - 4)^2}$$

Data 2

$$\sqrt{(6 - 6)^2 + (6 - 5)^2 + (8 - 8)^2 + (10 - 8)^2 + (11 - 10)^2 + (12 - 12)^2 + \dots + (3 - 4)^2}$$

Data 3

$$\sqrt{(4 - 6)^2 + (6 - 5)^2 + (7 - 8)^2 + (9 - 8)^2 + (10 - 10)^2 + (11 - 12)^2 + \dots + (4 - 4)^2}$$

Data 4

$$\sqrt{(5 - 6)^2 + (7 - 5)^2 + (6 - 8)^2 + (7 - 8)^2 + (8 - 10)^2 + (10 - 12)^2 + \dots + (4 - 4)^2}$$

Data 5

$$\sqrt{(7 - 6)^2 + (7 - 5)^2 + (8 - 8)^2 + (8 - 8)^2 + (9 - 10)^2 + (10 - 12)^2 + \dots + (5 - 4)^2}$$

Data 6

$$\sqrt{(6 - 6)^2 + (8 - 5)^2 + (7 - 8)^2 + (10 - 8)^2 + (11 - 10)^2 + (12 - 12)^2 + \dots + (5 - 4)^2}$$

Data 7

$$\sqrt{(5 - 6)^2 + (7 - 5)^2 + (6 - 8)^2 + (7 - 8)^2 + (10 - 10)^2 + (12 - 12)^2 + \dots + (6 - 4)^2}$$

Data 8

$$\sqrt{(4 - 6)^2 + (5 - 5)^2 + (7 - 8)^2 + (9 - 8)^2 + (8 - 10)^2 + (9 - 12)^2 + \dots + (3 - 4)^2}$$

Perhitungan tahap 1 mencari selisih antara data training dan data uji

$$Data 1 = \sqrt{1 + 4 + 1 + 1 + 4 + 1 + 4 + 1 + 1 + 0}$$

$$Data 2 = \sqrt{0 + 1 + 0 + 4 + 1 + 0 + 1 + 1 + 1 + 1}$$

$$Data 3 = \sqrt{4 + 1 + 1 + 1 + 0 + 1 + 0 + 0 + 0 + 0}$$

$$Data 4 = \sqrt{1 + 4 + 4 + 1 + 4 + 4 + 0 + 1 + 0 + 0}$$

$$Data 5 = \sqrt{1 + 4 + 0 + 0 + 1 + 4 + 1 + 1 + 0 + 1}$$

$$Data 6 = \sqrt{0 + 9 + 1 + 4 + 1 + 0 + 1 + 1 + 1 + 1}$$

$$Data 7 = \sqrt{1 + 4 + 4 + 1 + 0 + 0 + 9 + 4 + 1 + 4}$$

$$Data 8 = \sqrt{4 + 0 + 1 + 1 + 4 + 9 + 1 + 0 + 1 + 1}$$

Tahap 2 selanjutnya dilakukan perpangkatan 2 kemudian dijumlah untuk mencari jarak *Euclidian* masih masing titik

- Setelah mendapatkan jarak dari data baru maka dilakukan *sort* atau perankingan dengan data training untuk bisa menentukan tetangga terdekat berdasarkan jarak minimum K.
- Dari hasil perankingan maka dapat diperoleh kategori dari data yang dikelompokkan, mulai dari ranking 1 sampai 5, bisa dilihat tabel 3.2 dan 3.3 dibawah ini.

Data	K ₁	K ₂	K ₃	K ₄	K ₅	K ₆	K ₇	K ₈	K ₉	K ₁₀	Kategori	Jarak
1	5	7	7	9	8	11	10	7	6	4	A	3.605551
2	6	6	8	10	11	12	8	6	5	3	B	3.162278
3	4	6	7	9	10	11	9	7	6	4	A	2.828427
4	5	7	6	7	8	10	9	8	6	4	C	4.358899
5	7	7	8	8	9	10	8	8	6	5	A	3.605551
6	6	8	7	10	11	12	10	8	7	5	C	4.358899
7	5	7	6	7	10	12	12	9	7	6	B	5.291503
8	4	5	7	9	8	9	10	7	5	3	B	4.690416

Tabel 3.2 Sebelum perankingan

Data	K_1	K_2	K_3	K_4	K_5	K_6	K_7	K_8	K_9	K_{10}	Kategori	Jarak
3	4	6	7	9	10	11	9	7	6	4	A	2.828427
2	6	6	8	10	11	12	8	6	5	3	B	3.162278
1	5	7	7	9	8	11	10	7	6	4	A	3.605551
5	7	7	8	8	9	10	8	8	6	5	A	3.605551
4	5	7	6	7	8	10	9	8	6	4	C	4.358899
6	6	8	7	10	11	12	10	8	7	5	C	4.358899
8	4	5	7	9	8	9	10	7	5	3	B	4.690416
7	5	7	6	7	10	12	12	9	7	6	B	5.291503

Tabel 3.3 Setelah perangkingan (ditandai warna biru)

- Terakhir dari data yang diperoleh beserta kategorinya maka dapat diperoleh nilai mayoritas (kategori terbanyak) untuk selanjutnya data uji dapat diklasifikasi termasuk dalam jenis daun mangga manalagi, jiwo ataukah gadung.