

SKRIPSI

**KLASIFIKASI KUALITAS BIJI BUAH KAKAO
BERDASARKAN WARNA DAN TEKSTUR MENGGUNAKAN
*VIDEO PROCESSING***

Disusun dan diajukan oleh:

**FADILLAH RAMADHANY
D121 181 315**



**PROGRAM STUDI SARJANA TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2024**

LEMBAR PENGESAHAN SKRIPSI

KLASIFIKASI KUALITAS BIJI BUAH KAKAO BERDASARKAN WARNA DAN TEKSTUR MENGGUNAKAN *VIDEO PROCESSING*

Disusun dan diajukan oleh

Fadillah Ramadhany
D121181315

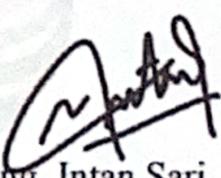
Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian
Studi Program Sarjana Program Studi Teknik Informatika
Fakultas Teknik Universitas Hasanuddin
Pada tanggal 5 Desember 2024
dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama,

Pembimbing Pendamping,


Prof. Dr. Ir. Indrabayu., S.T., M.T.,
M.Bus.Sys., IPM., ASEAN.Eng.
NIP 19750716 200212 1 004


Prof. Dr. Eng. Intan Sari Areni, S.T.,
M.T.
NIP 19750203 200012 2 002



Ketua Program Studi,


Prof. Dr. Ir. Indrabayu., S.T., M.T., M.Bus.Sys., IPM., ASEAN.Eng.
NIP 19750716 200212 1 004

PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini;

Nama : Fadillah Ramadhany

NIM : D121181315

Program Studi : Teknik Informatika

Jenjang : S1

Menyatakan dengan ini bahwa karya tulisan saya berjudul

KLASIFIKASI KUALITAS BIJI BUAH KAKAO BERDASARKAN WARNA DAN TEKSTUR MENGGUNAKAN *VIDEO PROCESSING*

Adalah karya tulisan saya sendiri dan bukan merupakan pengambilan alihan tulisan orang lain dan bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri.

Semua informasi yang ditulis dalam skripsi yang berasal dari penulis lain telah diberi penghargaan, yakni dengan mengutip sumber dan tahun penerbitannya. Oleh karena itu semua tulisan dalam skripsi ini sepenuhnya menjadi tanggung jawab penulis. Apabila ada pihak manapun yang merasa ada kesamaan judul dan atau hasil temuan dalam skripsi ini, maka penulis siap untuk diklarifikasi dan mempertanggungjawabkan segala resiko.

Segala data dan informasi yang diperoleh selama proses pembuatan skripsi, yang akan dipublikasi oleh Penulis di masa depan harus mendapat persetujuan dari Dosen Pembimbing.

Apabila dikemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan isi skripsi ini hasil karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Gowa, 5 Desember 2024

Yang Menyatakan



Fadillah Ramadhany

ABSTRAK

FADILLAH RAMADHANY. *Klasifikasi Kualitas Biji Buah Kakao Berdasarkan Warna Dan Tekstur Menggunakan Video Processing* (Dibimbing oleh **Indrabayu dan Intan Sari Areni**)

Penelitian ini dilatarbelakangi oleh kebutuhan untuk meningkatkan kualitas kakao di Indonesia. Meskipun Indonesia merupakan salah satu negara penghasil kakao terbesar di dunia, kualitas yang dihasilkan masih terbilang rendah. Hal ini disebabkan oleh subjektivitas dan inkonsistensi dalam pengklasifikasian biji kakao yang umumnya masih dilakukan secara manual. Oleh karena itu, diperlukan sistem otomatis untuk klasifikasi biji kakao guna meningkatkan kualitas biji kakao yang dihasilkan serta efisiensi dan akurasi di industri kakao. Penelitian ini bertujuan untuk mengimplementasikan algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan kualitas biji kakao berdasarkan fitur warna dan tekstur dengan menggunakan *video processing*, yang dibagi menjadi dua kategori, yaitu biji kakao baik dan biji kakao buruk. Data penelitian menggunakan video biji kakao yang bergerak di atas conveyor dengan variasi kecepatan dan jarak. Tahap *training* melibatkan 60 biji kakao baik dan 60 biji kakao buruk, dan tahap *testing* menggunakan 30 biji kakao baik dan 30 biji kakao buruk. Video diekstrak menjadi *frame* citra, kemudian melalui tahap *preprocessing* yang mencakup *resize*, augmentasi, konversi warna RGB ke HSV, *masking*, deteksi objek, dan *cropping*. Fitur yang digunakan meliputi nilai rata-rata HSV (*hue*, *saturation*, dan *value*) serta fitur GLCM (kontras, disimilaritas, energi, korelasi, homogenitas, dan ASM), dan diklasifikasi menggunakan SVM. Validasi dan pengujian model dilakukan menggunakan metrik akurasi, *precision*, dan *recall*. Hasil penelitian menunjukkan skenario terbaik diperoleh dengan menerapkan *resize* sebesar 50% ukuran citra asli, tanpa penajaman, menggunakan dua fitur dengan korelasi tertinggi yaitu *saturation* dan *value*, serta pengambilan data optimal pada kecepatan *conveyor* 80 RPM dan jarak kamera 5 cm. Berdasarkan uji coba skenario terbaik pada video klasifikasi single-objek biji kakao didapatkan akurasi klasifikasi sebesar 96,67% dengan presisi 93,75% dan *recall* 100%, dan pada pengujian video klasifikasi multi-objek didapatkan akurasi klasifikasi sebesar 86,67% dengan *precision* 84,38% dan *recall* 90%.

Kata Kunci: Klasifikasi, Biji Kakao, *Video Processing*, *Support Vector Machine*, HSV, GLCM

ABSTRACT

FADILLAH RAMADHANY. *Classification of Cocoa Bean Quality Based on Color and Texture Using Video Processing* (Supervised by **Indrabayu** and **Intan Sari Areni**)

This research is motivated by the need to improve the quality of cocoa in Indonesia. Although Indonesia is one of the largest cocoa-producing countries in the world, the quality produced is still relatively low. This is due to the subjectivity and inconsistency in the classification of cocoa beans which are generally still done manually. Therefore, an automated system is needed for cocoa bean classification to improve the quality of cocoa beans produced and efficiency and accuracy in the cocoa industry. This study aims to implement the Support Vector Machine (SVM) algorithm in classifying the quality of cocoa beans based on color and texture features using video processing, which is divided into two categories, namely good cocoa beans and bad cocoa beans. The research data uses videos of cocoa beans moving on a conveyor with variations in speed and distance. The training stage involves 60 good and 60 bad cocoa beans, and the testing stage uses 30 good and 30 bad cocoa beans. The video is extracted into image frames and then goes through a preprocessing stage that includes resizing, augmentation, RGB to HSV color conversion, masking, object detection, and cropping. The features used include the average HSV value (hue, saturation, and value) and GLCM features (contrast, dissimilarity, energy, correlation, homogeneity, and ASM), and are classified using SVM. Model validation and testing use accuracy, precision, and recall metrics. The results show that the best scenario is obtained by applying a resize of 50% of the original image size, without sharpening, using two features with the highest correlation, namely saturation and value, and optimal data collection at a conveyor speed of 80 RPM and a camera distance of 5 cm. Based on the best scenario trial on the video, a single-object classification accuracy of 96.67% was obtained with a precision of 93.75% and a recall of 100%. While in the multi-object video classification test, a classification accuracy of 86.67% was obtained with a precision of 84.38% and a recall of 90%.

Keywords: *Classification, Cocoa Beans, Video Processing, Support Vector Machine, HSV, GLCM*

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	i
PERNYATAAN KEASLIAN.....	ii
ABSTRAK	iii
ABSTRACT	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR	vi
DAFTAR TABEL.....	viii
DAFTAR SINGKATAN DAN ARTI SIMBOL	ix
DAFTAR LAMPIRAN.....	x
KATA PENGANTAR	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Ruang Lingkup.....	3
BAB II TINJAUAN PUSTAKA.....	5
2.1 Kakao	5
2.2 Visi Komputer.....	7
2.3 Citra.....	8
2.4 Pengolahan Citra Digital.....	10
2.5 Pengolahan Video Digital	12
2.6 Deteksi Objek.....	13
2.7 Ekstraksi Fitur	14
2.8 <i>Pearson Correlation</i>	17
2.9 Klasifikasi dengan <i>Support Vector Machine</i> (SVM)	17
2.10 <i>Confusion Matrix</i>	20
BAB 3 METODE PENELITIAN.....	23
3.1 Waktu dan Lokasi Penelitian	23
3.2 Benda Uji dan Alat.....	23
3.3 Tahapan Penelitian.....	24
3.4 Teknik Pengambilan Data.....	25
3.5 Perancangan Sistem	27
BAB 4. HASIL DAN PEMBAHASAN.....	69
4.1 Hasil <i>Training</i>	69
4.2 Hasil <i>Testing</i>	75
BAB 5. KESIMPULAN DAN SARAN.....	83
5.1 Kesimpulan	83
5.2 Saran.....	84
DAFTAR PUSTAKA	85

DAFTAR GAMBAR

Gambar 1 Citra Biner (a), Citra <i>Grayscale</i> (b), Citra Berwarna (c)	10
Gambar 2 Ruang Warna HSV	12
Gambar 3 Penentuan Hyperplane Terbaik	18
Gambar 4 Tahapan Penelitian	24
Gambar 5 Biji Kakao Baik (a) dan Biji Kakao Buruk (b).....	26
Gambar 6 Teknik Pengambilan data biji kakao pada <i>conveyor</i>	27
Gambar 7 Alur Perancangan Sistem	28
Gambar 8 Contoh hasil seleksi <i>frame</i> : <i>frame</i> yang digunakan (a), <i>frame</i> yang tidak digunakan (b).....	30
Gambar 9 Proses <i>resize</i> pada citra menjadi 50% ukuran citra asli	32
Gambar 10 Grafik Akurasi Skenario <i>Resizing</i>	35
Gambar 11 Grafik Akurasi Skenario <i>Sharpening</i>	39
Gambar 12 Nilai matriks 2 x 2 hasil konversi RGB ke HSV	41
Gambar 13 <i>Frame</i> biji kakao hasil konversi RGB ke HSV; Biji baik RGB (a), Biji baik HSV (b), biji buruk RGB (c), biji buruk HSV (d).....	41
Gambar 14 Cara Pengambilan nilai ambang HSV	42
Gambar 15 Nilai matriks konversi citra HSV ke citra Hitam Putih.....	43
Gambar 16 Contoh <i>Frame</i> Biji Kakao Baik (a) dan Biji Buruk (b) saat proses <i>Masking Image</i>	44
Gambar 17 Alur Deteksi Objek.....	44
Gambar 18 Contoh <i>Frame</i> biji kakao baik (a) dan buruk (b) hasil <i>bounding box</i> dan <i>crop</i>	45
Gambar 19 Alur Ekstraksi Fitur Warna	46
Gambar 20 Konversi citra HSV	47
Gambar 21 Nilai HSV	47
Gambar 22 Alur Ekstraksi Fitur Tekstur.....	49
Gambar 23 Pembentukan matriks GLCM jarak 1 dan sudut 0°	49
Gambar 24 Pembentukan matriks GLCM jarak 1 dan sudut 45°	49
Gambar 25 Pembentukan matriks GLCM jarak 1 dan sudut 90°	50
Gambar 26 Pembentukan matriks GLCM jarak 1 dan sudut 135°	50
Gambar 27 Hasil Ekstraksi Fitur.....	55
Gambar 28 <i>Heatmap Feature Correlation</i>	58
Gambar 29 Alur Sistem Proses Pelatihan SVM.....	62
Gambar 30 Memilih <i>Board</i> dan <i>Port</i>	65
Gambar 31 <i>Source code</i> kecepatan <i>conveyor</i>	66
Gambar 32 <i>Frame</i> Hasil klasifikasi kualitas biji kakao; Biji Kakao Baik (a) dan Biji Kakao Buruk (b).....	68
Gambar 33 Hasil Skenario Pengambilan Data.....	70
Gambar 34 <i>Confussion Matrix</i> Skenario 5cm 80 rpm	71
Gambar 35 Hasil Parameter Terbaik dari Metode <i>Grid Search</i>	72
Gambar 36 Desain corong <i>conveyor</i> baru untuk skenario pengambilan data multi-objek (a) ilustrasi dan (b) kondisi nyata	73
Gambar 37 <i>Confussion Matrix</i> Klasifikasi Multi-Objek Biji Kakao	74
Gambar 38 Data Kesalahan Klasifikasi Proses Pengujian.....	79
Gambar 39 Visualisasi Distribusi Nilai Value dan Saturation.....	80

Gambar 40 Frame hasil pengujian multi-objek biji kakao..... 82

DAFTAR TABEL

Tabel 1. Mutu Biji Kakao Berdasarkan SNI	6
Tabel 2. Interpretasi nilai korelasi pearson	17
Tabel 3. Skenario Pengambilan Data	27
Tabel 4. Hasil Pengambilan Data Video	29
Tabel 5. Hasil Ekstraksi <i>Frame</i>	30
Tabel 6. Hasil Augmentasi Flip <i>Horizontal</i>	31
Tabel 7. Hasil <i>Resizing</i>	32
Tabel 8. Hasil Training Skenario <i>Resize</i>	33
Tabel 9. Hasil Penajaman Citra.....	36
Tabel 10. Hasil Training Skenario <i>Sharpening</i>	37
Tabel 11. <i>Range</i> Nilai HSV.	42
Tabel 12. Nilai Korelasi Pearson Semua Fitur	56
Tabel 13. Skenario Jumlah Fitur	59
Tabel 14. Hasil Training Skenario Jumlah Fitur	60
Tabel 15. Hasil Skenario Pengambilan Data.....	69
Tabel 16. Hasil Testing Pada Video.....	76
Tabel 17. <i>Confusion Matriks</i> Pada Proses Pengujian.....	78
Tabel 18. Nilai Fitur Pada Data Kesalahan Klasifikasi.....	80
Tabel 19. <i>Confusion Matriks</i> Pada Proses Pengujian Multi-Objek.....	81

DAFTAR SINGKATAN DAN ARTI SIMBOL

Lambang/Singkatan	Arti dan Keterangan
AI	<i>Artificial Intelligence</i>
SVM	<i>Support Vector Machine</i>
HSV	<i>Hue, Saturation, Value</i>
RGB	<i>Red, Green, Blue</i>
GLCM	<i>Gray Level Co-occurrence Matrix</i>
FPS	<i>Frames Per Second</i>
RPM	<i>Revolutions Per Minute</i>
DC	<i>Direct Current</i>
FAO	<i>Food and Agriculture Organization</i>
SNI	Standar Nasional Indonesia
ASM	<i>Angular Second Moment</i>

DAFTAR LAMPIRAN

Lampiran 1 Contoh <i>Frame</i> Biji Kakao Pada Proses <i>Training</i>	89
Lampiran 2 Nilai Ekstraksi Fitur.....	90
Lampiran 3 Gambar Hasil <i>Confussion Matrix</i> Untuk Setiap Skenario.....	91
Lampiran 4 Video Pengambilan Data dan Hasil Testing.....	92
Lampiran 5 <i>Source Code</i>	93

KATA PENGANTAR

Alhamdulillah, puji syukur kehadirat Allah SWT atas berkat rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul “Klasifikasi Kualitas Biji Buah Kakao Berdasarkan Warna dan Tekstur Menggunakan *Video Processing*” sebagai salah satu persyaratan yang harus dipenuhi dalam menyelesaikan jenjang Strata-1 pada Program Studi Teknik Informatika Fakultas Teknik Universitas Hasanuddin. Sholawat serta salam kepada nabi Muhammad SAW yang telah menunjukkan dan mengajarkan akhlak mulia sehingga didapatkan kenyamanan dan keramahan dalam berhubungan dengan orang di sekitar.

Dengan segala kerendahan hati, penulis menyadari bahwa dalam penyusunan dan penulisan laporan skripsi ini tidak lepas dari bantuan, bimbingan serta dukungan dari berbagai pihak, dari masa perkuliahan sampai dengan masa penyusunan skripsi ini. Sehingga, pada kesempatan ini penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT. Atas semua karunia serta pertolongan-Nya yang tiada batas, yang telah diberikan kepada penulis di setiap langkah dalam penelitian hingga penulisan laporan ini.
2. Kedua orang tua penulis, Bapak Mahyudin Djaga dan Ibu Hasrah Sangintang yang selalu mendoakan untuk kebaikan penulis, selalu memberikan kasih sayang, cinta, dukungan dan motivasi. Terima kasih Bapak dan Ibu yang selalu sabar dan semangat tiada henti dalam menghadapi dan mendidik penulis.
3. Nenek penulis, Isra Saluki dan Almarhum Kakek penulis, Hasdik Sangintang yang sudah merawat penulis di masa kecil.
4. Bapak Prof. Dr. Ir. Indrabayu, S.T., M.T., M.Bus.Sys., IPM., ASEAN.Eng., selaku pembimbing I dan Ibu Prof. Dr. Eng. Intan Sari Areni, ST., M.T. sebagai pembimbing II yang telah meluangkan waktu, tenaga dan pikiran serta perhatian yang luar biasa untuk membimbing penulis dalam proses penyelesaian skripsi ini.
5. Segenap Dosen dan Staf Departemen Teknik Informatika Fakultas Teknik Universitas Hasanuddin yang telah memberikan ilmu dan pengetahuan baru, serta bantuan kepada penulis selama masa perkuliahan.
6. Segenap keluarga *AIMP Research Group* Universitas Hasanuddin yang telah memberikan banyak bantuan selama penelitian dan teman diskusi terkait progres penyusunan skripsi.
7. Teman-Teman SYNCHRONOUS yang banyak membantu selama kuliah dan dalam proses menyelesaikan skripsi ini.

8. Sahabat-sahabat Nur Hasanah, Tiara Yania, Putri Alisyah, Indah Mulya, Maghfirah Tenri, Fikri Khaikal, Rachmat Maulana, Rahmadani, Try Rezki, Dandi Garda, Aqilah Syahratunnisa, Elisa, dan Agunawan, yang senantiasa membantu, menemani, dan memberikan semangat kepada penulis untuk menyelesaikan penyusunan tugas akhir ini.
9. A. M. Ichsan Noer Abustan yang senantiasa mendukung dan menemani penulis di saat senang maupun sedih.
10. Serta pihak-pihak lain yang tidak sempat disebutkan dan tanpa sadar telah membantu penulis dalam menyelesaikan skripsi.

Dengan rasa syukur dan kerendahan hati, penulis memberikan rasa hormat yang tak terhingga, semoga Allah SWT. membalas semua kebaikan dari semua pihak yang telah banyak membantu penulis dalam menyelesaikan skripsi ini. Penulis menyadari skripsi ini masih banyak kekurangan dan jauh dari kata sempurna. Oleh karena itu, penulis mengharapkan segala masukan dan saran yang membangun sehingga skripsi ini dapat memberi manfaat bagi penulis dan pembaca. Akhir kata, semoga skripsi ini dapat dijadikan sebagai sumber ilmu pengetahuan dan bermanfaat bagi penulis dan pembaca pada umumnya.

Makassar, November 2024

Penulis

BAB I PENDAHULUAN

1.1 Latar Belakang

Di era revolusi industri 4.0, berbagai aktivitas sosial, pendidikan, ekonomi dan sebagainya selalu dikaitkan dengan penggunaan mesin-mesin otomatisasi yang terintegrasi (Fetia Wardhiani, 2019). Revolusi Industri 4.0 mendorong sistem otomatisasi di dalam semua proses aktivitas. Sektor pertanian juga tidak luput dari pengaruh industri 4.0. Menteri Pertanian, Andi Amran Sulaiman menerangkan “Melalui implementasi Industri 4.0 disektor pertanian, diharapkan proses usaha tani menjadi semakin efisien, peningkatan produktivitas, dan daya saing”. Menurut Amran, ada 5 teknologi utama yang mendukung penerapan Industri 4.0, antara lain *Internet of Things*, *Artificial Intelligence*, *Human-Machine Interface*, Teknologi robotik dan sensor, serta teknologi *3D printing*. Kesemuanya ini mentransformasi cara manusia berinteraksi hingga pada tingkat yang sangat fundamental, dengan tujuan meningkatkan efisiensi dan daya saing dalam industri (KEMENTAN, 2019).

Salah satu komoditas unggulan di sektor pertanian adalah komoditas kakao. Kakao memainkan peran penting dalam perekonomian negara dengan menghasilkan devisa, memberikan pendapatan kepada petani, menciptakan lapangan kerja, mendorong sektor agribisnis dan agroindustri, serta mendukung pengembangan wilayah (DITJENBUN, 2019). *Theobroma cacao L.* atau Kakao adalah tanaman yang salah satu manfaatnya adalah menjadi bahan baku pembuatan coklat. Cokelat terbuat dari biji kakao yang telah melalui beberapa proses pengolahan sehingga menghasilkan bentuk dan rasa yang sama dengan yang ada di pasaran. Menurut Badan Pusat Statisti, Indonesia merupakan produsen kakao terbesar ketujuh di dunia setelah Pantai Gading dan Ghana dengan produksi mencapai 641,7 ribu ton pada tahun 2023. Hal ini menjadikan kakao sebagai salah satu komoditas perkebunan unggulan di Indonesia (Hapsari, 2023).

Namun sayangnya, kualitas biji kakao dari Indonesia cukup rendah, hal ini disebabkan oleh pengolahan produk kakao Indonesia kebanyakan masih dilakukan secara tradisional (I Nyoman Wata, 2018). Dalam menentukan kualitas biji secara tradisional masih terdapat banyak kekurangan, karena dipengaruhi subjektivitas

karyawan seperti cepat lelah dan pengaruh fisik lainnya, sehingga pada kondisi tertentu proses pengklasifikasian tidak konsisten (Purnama, 2021). Kualitas biji kakao dapat dilihat dari penampakan fisik biji melalui warna dan teksturnya. Ciri-ciri biji kakao yang berkualitas baik yaitu berwarna coklat kemerahan yang dominan dengan tekstur berongga, tidak kopong dan cangkang biji tidak pecah. Untuk ciri biji kakao dengan kualitas yang buruk yaitu berwarna keabu-abuan (*slaty*) atau gelap kehitaman dengan tekstur yang pejal, gepeng, terdapat cacat pecah pada cangkang biji, atau biji saling melekat (Diansari, 2015).

Untuk meminimalisir terjadinya kesalahan-kesalahan saat penentuan kualitas biji kakao dan untuk mempercepat waktu penentuan kualitasnya, maka diperlukan suatu sistem yang mampu melakukan klasifikasi kualitas biji kakao secara otomatis dengan jumlah yang banyak. Salah satu caranya adalah dengan memanfaatkan teknologi pembelajaran mesin. *Support Vector Machine* (SVM) adalah salah satu algoritma yang sering digunakan pada proses klasifikasi karena keunggulan dalam akurasi dan kecepatan proses komputasi. Dari penelitian yang dilakukan oleh Selvia Ferdiana Kusuma, dkk pada tahun 2017 tentang klasifikasi kematangan buah mengkudu berdasarkan warna dan tekstur dengan membandingkan metode klasifikasi K-Nearest Neighbor dan SVM, didapatkan hasil klasifikasi dengan SVM memiliki akurasi yang lebih tinggi yaitu 85%, sementara K-NN memiliki akurasi 76,67% (Kusuma et al., 2017). Juga Penelitian yang dilakukan oleh Yudhi Aditya, Armin Lawi, Hartono dan Mursalin pada tahun 2018 terkait klasifikasi citra biji kakao berdasarkan fitur warna menggunakan SVM menghasilkan akurasi tinggi yaitu 99,281%.

Berdasarkan hal itulah, penulis melakukan penelitian yang berjudul “Klasifikasi Kualitas Biji Buah Kakao Berdasarkan Warna Dan Tekstur Menggunakan *Video Processing*”, agar dapat digunakan pada industri pengolahan biji kakao dalam mengklasifikasikan kualitas biji kakao yang baik dan tidak baik secara otomatis dengan menggunakan *video processing*. Pada penelitian ini diimplementasikan algoritma *Support Vector Machine* sebagai metode klasifikasi.

1.2 Rumusan Masalah

1. Bagaimana implementasi algoritma SVM dalam mengklasifikasikan kualitas biji kakao berdasarkan warna dan tekstur menggunakan *video processing*?
2. Bagaimana kinerja klasifikasi pada biji kakao berdasarkan warna dan teksturnya menggunakan *video processing*?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengimplementasikan algoritma SVM dalam mengklasifikasikan kualitas biji kakao berdasarkan warna dan teksturnya menggunakan *video processing*.
2. Mengetahui kinerja klasifikasi biji kakao berdasarkan warna dan teksturnya menggunakan *video processing*.

1.4 Manfaat Penelitian

Dengan dilakukannya penelitian ini, diharapkan dapat memberikan manfaat sebagai berikut:

1. Diharapkan dapat menjadi referensi bagi akademisi maupun peneliti berikutnya dalam melakukan penelitian yang terkait dengan tema yang penulis teliti.
2. Sebagai inovasi yang dapat meminimalisir terjadinya kesalahan saat penentuan kualitas biji kakao.
3. Membantu meningkatkan kualitas kontrol pada industri pengolahan kakao.
4. Dapat meningkatkan penggunaan teknologi *video processing* pada bidang perkebunan, khususnya pada pengolahan kakao.

1.5 Ruang Lingkup

1. Biji kakao yang digunakan adalah biji yang sudah difermentasi.
2. Parameter yang digunakan adalah warna dan tekstur biji kakao.

3. Biji kakao yang diklasifikasi dibagi menjadi 2 kategori yaitu biji berkualitas baik dan berkualitas buruk.
4. Data yang digunakan yaitu data citra berupa video dan gambar yang diambil dari satu sisi biji kakao.
5. Kamera yang digunakan saat pengujian adalah kamera statis Logitech Brio Ultra HD beresolusi 1080 piksel, bukan auto fokus.
6. Algoritma yang digunakan untuk klasifikasi adalah *Support Vector Machine* (SVM) kernel *Radial Basis Function* (RBF), dan Algoritma *Gray Level Co-Occurrence Matrix* untuk ekstraksi fitur tekstur.
7. Penelitian ini hanya mengimplementasikan algoritma SVM, tidak sampai menerapkan tahap pemisahan otomatis.

BAB II TINJAUAN PUSTAKA

2.1 Kakao

Kakao (*Theobroma cacao L.*) merupakan salah satu komoditas perkebunan yang sesuai untuk perkebunan rakyat, karena tanaman ini dapat berbunga dan berbuah sepanjang tahun, sehingga dapat menjadi sumber pendapatan harian atau mingguan bagi pekebun. Tanaman kakao berasal dari daerah hutan hujan tropis di Amerika Selatan. Di daerah asalnya, kakao merupakan tanaman kecil di bagian bawah hutan hujan tropis dan tumbuh terlindung pohon-pohon yang besar (Radifan, 2022). Berikut adalah taksonomi tanaman kakao:

- Kingdom : *Plantae* (Tanaman)
- Subkingdom : *Tracheobionta* (Tanaman berpembuluh)
- Superdivisi : *Spermatophyta* (Menghasilkan biji)
- Divisi : *Magnoliophyta* (Tumbuhan berbunga)
- Kelas : *Magnoliopsida* (Dikotil)
- Subkelas : *Dilleniidae*
- Ordo : *Malvales*
- Famili : *Malvaceae* (Famili kakao)
- Genus : *Theobroma*
- Spesies : *Theobroma cacao L.*

Tanaman Kakao secara umum mulai berproduksi pada umur 2,5 - 3 tahun setelah tanam. Buah kakao dihasilkan melalui proses penyerbukan antara bunga jantan dan bunga betina yang tumbuh menempel pada semua bagian batang tanaman sehingga menghasilkan buah kakao (Radifan, 2022) . Buah kakao terdiri atas kulit buah (*pod*), arilus (*pulp*), dan biji. Pada setiap buah kakao bisa menghasilkan sebanyak 30 – 50 biji kakao tergantung varietas tanaman dan kondisi pertumbuhan tanaman kakao (Hilyatunnisa, 2013).

Biji kakao merupakan bahan yang kaya akan flavonoid, yang mempunyai kapasitas antioksidan dalam menangkal radikal bebas. Biji kakao yang telah diproses menghasilkan beberapa produk utama, seperti *Cocoa Mass*, *Cocoa Butter*, *Cocoa Powder*, *Chocolate*, *Cocoa Nibs*, dll (Hilyatunnisa, 2013). Proses pembuatan

produk-produk ini melibatkan beberapa langkah utama mulai dari buah kakao yang telah dipanen dibuka untuk mengeluarkan biji kakao yang ditutupi *pulp* putih yang berserat. Biji-biji tersebut lalu ditumpuk di atas tikar, daun pisang atau dalam kotak dan ditutup. Kemudian dilakukan proses fermentasi, bakteri dan khamir (*yeast*) di udara menyebabkan terjadinya perubahan pada komponen kimiawi dalam *pulp* sehingga menjadi alkohol dan asam asetat dalam proses fermentasi. Ini adalah salah satu tahapan yang paling penting dalam tahap pengolahan biji kakao, karena di sinilah awal pembentukan citarasa dari kakao. Secara berkala, petani akan membolak-balik tumpukan biji secara perlahan agar proses terjadi secara merata. Proses fermentasi dapat memakan waktu selama 5-8 hari. Selanjutnya yaitu proses pengeringan, Biji kakao disebar di atas nampan dan dikeringkan oleh sinar matahari selama kurang lebih seminggu. Setelah dikeringkan, kandungan air pada biji kakao berkurang menjadi 6-7 persen, dan beratnya berkurang hampir setengahnya. Biji kemudian disortasi berdasarkan ukuran dan kualitas, dan dikemas untuk dikirim ke pabrik. Proses sortasi atau penentuan mutu biji ini tentu sangat penting karena membantu mengurangi kontaminasi dan memastikan hanya biji berkualitas tinggi yang diolah, proses pemilahan juga memungkinkan untuk memisahkan biji yang rusak, yang dapat mempengaruhi rasa dan kualitas akhir produk kakao. Dan juga proses produksi kakao dapat menjadi lebih efisien dan menghasilkan produk akhir yang lebih konsisten. Setelah melalui serangkaian proses tersebut, biji kakao siap untuk dikirim ke pabrik untuk kemudian diolah lebih lanjut menjadi produk coklat. (PT. Freyabadi, 2022).

Berdasarkan Standar Nasional Indonesia (SNI), berikut adalah perbedaan ciri-ciri biji kakao berkualitas baik dan berkualitas buruk yang ditunjukkan pada Tabel 1.

Tabel 1. Mutu Biji Kakao Berdasarkan SNI

	Biji Baik	Biji Buruk
Warna	Warna yang merah kecoklatan	Warna gelap kehitaman
Cangkang	Berongga, tidak kopong, memiliki kepala yang keras, padat, dan tidak mudah hancur	tekstur yang pejal, gepeng, terdapat cacat pecah atau saling melekat

Kadar Air	Biji baik memiliki kadar air yang rendah, seperti 6-7% untuk Grade AA, 7-8% untuk Grade A, dan 7,5% untuk Grade B	memiliki kadar air yang tinggi, seperti 10% untuk Grade C
Kadar Kotoran	0% untuk Grade AA, 1-2% untuk Grade A, dan 2,5% untuk Grade B	4% untuk Grade C
Kadar Keabuan	Kadar keabuan rendah	Kadar keabuan tinggi (slaty)
Kadar Kafein	kadar kafein tinggi	Kadar kafein rendah

Sumber: (Badan Standarisasi Nasional, 2008)

2.2 Visi Komputer

Visi komputer dapat dijelaskan sebagai proses pemrosesan gambar yang terlibat dalam berbagai tahapan mulai dari pengambilan gambar, pemrosesan, kategorisasi, hingga analisis dan pengambilan keputusan berdasarkan identifikasi gambar tersebut. Tujuan utama dari teknologi visi komputer adalah untuk meniru kapasitas visual manusia dalam perangkat elektronik, memungkinkan perangkat tersebut untuk memahami dan menafsirkan makna dari gambar yang diberikan (Prabowo & Abdullah, 2018).

Sebagai bidang akademik, visi komputer melibatkan teori yang mendukung sistem buatan untuk mengekstrak informasi dari gambar. Data gambar ini bisa berasal dari berbagai sumber, termasuk urutan video, pandangan dari beberapa kamera, atau data multidimensi dari pemindai medis. Dalam konteks teknologi, visi komputer berupaya menerapkan teori dan model ini dalam pengembangan sistem visi komputer. Beberapa aplikasi dari visi komputer termasuk, namun tidak terbatas pada, sistem untuk kontrol proses (seperti robot industri atau kendaraan otonom), deteksi kejadian (misalnya untuk pengawasan atau menghitung orang), pengorganisasian informasi (seperti pengindeksan database gambar), pemodelan objek atau lingkungan (seperti dalam inspeksi industri atau analisis citra medis), dan interaksi (seperti input untuk interaksi manusia-komputer) (Prabowo & Abdullah, 2018).

2.3 Citra

Citra adalah representasi grafis pada permukaan datar yang menunjukkan variasi dalam tingkat kegelapan, kecerahan, atau beragam warna, yang diperoleh menggunakan sistem visual yang memiliki kemiripan dengan suatu objek (Sulistiyanti et al., 2016). Secara umum citra terbagi menjadi dua sifat yaitu citra analog dan citra digital:

2.3.1 Citra Analog

Citra analog adalah citra yang bersifat continue, seperti gambar pada monitor televisi, foto sinar X, foto yang tercetak di kertas foto, lukisan, pemandangan alam, hasil CT scan, gambar-gambar yang terekam pada pita kaset, dan lain sebagainya. Citra analog tidak dapat direpresentasikan dalam komputer, sehingga tidak bisa diproses di komputer secara langsung. Oleh sebab itu, agar ini dapat diproses di komputer, proses konversi analog ke digital harus dilakukan terlebih dahulu. Citra analog dihasilkan dari alat-alat analog, seperti video kamera analog, kamera foto analog, cam, CT scan, sensor rontgen untuk foto thorax, sensor gelombang pendek pada sistem radar, sensor ultrasound pada sistem USG, dan lain-lain (Yanu F. et al., 2022).

2.3.2 Citra Digital

Citra digital merupakan representatif dari citra yang diambil oleh mesin dengan bentuk pendekatan berdasarkan sampling dan kuantisasi. Sampling menyatakan besarnya kotak-kotak yang disusun dalam baris dan kolom. Dengan kata lain, sampling pada citra menyatakan besar kecilnya ukuran pixel (titik) pada citra, dan kuantisasi menyatakan besarnya nilai tingkat kecerahan yang dinyatakan dalam nilai tingkat keabuan (*grayscale*) sesuai dengan jumlah bit biner yang digunakan oleh mesin, dengan kata lain kuantisasi pada citra menyatakan jumlah warna yang ada pada citra (Yanu F. et al., 2022).

2.3.2.1 Citra Biner

Citra biner adalah metode representasi gambar digital yang hanya menggunakan dua warna: hitam dan putih. Setiap warna ini

direpresentasikan oleh angka biner, yaitu 0 dan 1. Dalam citra biner, terdapat dua model representasi warna yang berbeda. Pertama, dalam model citra cahaya, angka 1 menunjukkan warna putih (simbolisasi adanya cahaya), sedangkan angka 0 menunjukkan warna hitam (simbolisasi ketiadaan cahaya). Kedua, dalam model citra tinta atau cat, angka 1 menunjukkan warna hitam (menandakan keberadaan cat), dan angka 0 menunjukkan warna putih (menandakan ketiadaan cat) (Yanu F. et al., 2022).

Citra biner sering sekali muncul sebagai hasil dari proses pengolahan, seperti segmentasi, morfologi ataupun *dithering*. Fungsi dari binerisasi sendiri adalah untuk mempermudah proses pengenalan pola, karena pola akan lebih mudah terdeteksi pada citra yang mengandung lebih sedikit warna (C. Gonzalez & E. Woods, 2008).

2.3.2.2 Citra Grayscale

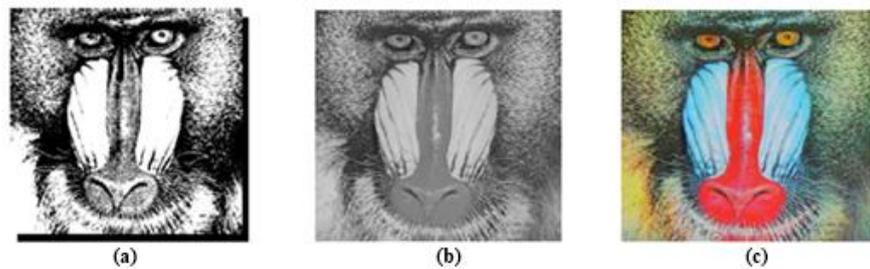
Citra *grayscale* merupakan suatu cara dalam merepresentasikan citra digital dengan menggunakan skala derajat keabuan, dimana derajat keabuan yang ada merupakan hasil pemangkatan nilai bit yang ada terhadap angka 2 (2^n). Misalkan skala keabuan 4bit memiliki rentang skala keabuan sebanyak 24 warna = 16 warna, yang diwakili dengan angka 0 hingga 15. (angka 0 / minimal mewakili warna hitam, dan angka 15 / maksimal mewakili warna putih). Adapun angka diantara 0 hingga 15 merepresentasikan warna abu dalam skala kecerahan yang berbeda (Yanu F. et al., 2022).

2.3.2.3 Citra Berwarna

Citra warna merupakan metode dalam merepresentasikan suatu citra secara digital, dimana metode ini menggunakan kombinasi dari tiga warna primer (merah, hijau dan biru = RGB) untuk membentuk suatu citra. Adapun setiap titik pada citra mewakili kombinasi dari ketiga warna ini (Yanu F. et al., 2022). Setiap warna dasar memiliki intensitas tersendiri dengan nilai minimum nol (0) dan nilai maksimum 255 (8 bit). RGB didasarkan pada teori bahwa mata manusia peka terhadap

panjang gelombang 630nm (merah), 530 nm(hijau), dan 4nm (biru) (C. Gonzalez & E. Woods, 2008).

Berikut adalah contoh dari citra biner, citra *grayscale* dan citra berwarna:



Gambar 1 Citra Biner (a), Citra *Grayscale* (b), Citra Berwarna (c)
Sumber: (Henry A., 2014)

2.4 Pengolahan Citra Digital

Pengolahan citra digital adalah teknik yang digunakan untuk memanipulasi dan mengubah piksel-piksel dalam gambar untuk mencapai tujuan tertentu. Awalnya, teknik ini dikembangkan untuk meningkatkan kualitas gambar. Namun, seiring dengan perkembangan teknologi komputer, yang ditandai dengan peningkatan kapasitas dan kecepatan pemrosesan, serta kemajuan dalam bidang ilmu komputasi, pengolahan citra sekarang juga memungkinkan pengambilan informasi yang berguna dari gambar tersebut (Yanu F. et al., 2022).

Secara garis besar, dalam melakukan pengolahan citra digital perlu dilakukan tiga hal berikut:

a. *Pre-processing*

Citra yang akan diproses, terlebih dahulu dilakukan pra pemrosesan. Pre-processing adalah proses operasi citra pada tingkat abstraksi yang terendah dan merupakan tahapan awal yang dilakukan dalam pengolahan citra. Tujuan dari pre-processing ini adalah untuk meningkatkan kualitas citra sehingga dapat dilakukan pemrosesan lebih lanjut. Pre-processing dapat menghilangkan noise, memperjelas fitur yang akan diekstraksi pada proses selanjutnya. Contoh proses pre-processing adalah augmentasi, *resize*, *sharpening*, *cropping*, dan lain sebagainya. Selain perbaikan kualitas, pre-processing juga digunakan untuk mengubah tipe data dan ukuran data

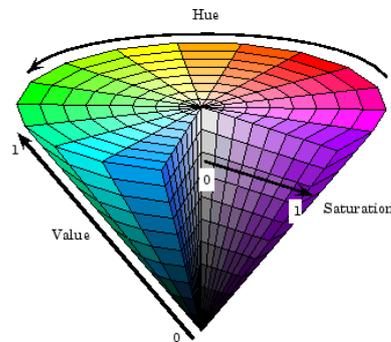
gambar. Hal ini disesuaikan dengan kebutuhan jenis data yang akan diproses pada tahap selanjutnya (Indrabulan, 2016).

b. Segmentasi

Segmentasi adalah proses yang bertujuan untuk memisahkan antara objek atau *foreground* dengan latar belakang atau *background*. Pada umumnya keluaran hasil segmentasi citra adalah berupa citra biner di mana objek (*foreground*) yang dikehendaki berwarna putih (1), sedangkan *background* nya yang ingin dihilangkan berwarna hitam (0). Segmentasi juga merupakan proses membagi citra menjadi beberapa bagian. Pembagian ini biasanya berdasarkan karakteristik dari piksel pada citra. Beberapa metode segmentasi citra yang biasa digunakan adalah *thresholding*, *active contour*, segmentasi warna, deteksi tepi, *watershed* dan transformasi *Hough* (Indrabulan, 2016). Pada penelitian ini, metode segmentasi yang digunakan adalah segmentasi warna *Hue*, *Saturation* dan *Value* (HSV). Segmentasi HSV dilakukan dengan menganalisis nilai warna tiap piksel citra sesuai fitur yang diinginkan dengan nilai toleransi pada setiap dimensi warna HSV.

HSV merupakan warna yang memiliki model ruang warna yang berbeda dengan RGB. *Hue* digunakan untuk menyatakan warna sesungguhnya seperti merah, jingga, kuning dll. *Hue* digunakan untuk membedakan tingkat warna kemerahan, kehijauan, kekuningan, kebiruan, dll berdasarkan gelombang cahaya. Nilai ini menyatakan bahwa model hue ini berbentuk lingkaran. Sehingga apabila kita telusuri terus dari warna merah maka kita akan kembali lagi ke warna merah. *Saturation* digunakan untuk menyatakan tingkat keputihan dari setiap warna pada *hue*. Sehingga apabila bila warna merah memiliki tingkat *saturation* yang tinggi atau bernilai 1 maka tidak ada kandungan warna putih pada warna *hue*. Sedangkan apabila *saturation* memiliki nilai 0 maka kandungan warna tersebut tidak memiliki *hue* atau sepenuhnya berwarna putih. *Value* pada model HSV merupakan nilai kadar warna yang berada pada warna tersebut. Mirip seperti *saturation*, *value* ini memiliki rentang dari 0 hingga 1. Apabila nilai *value* ini sebesar 0, maka warna tersebut akan menjadi hitam. Sedangkan apabila nilai *value* bernilai 1

maka tingkat kandungan warna hitam pada warna tersebut menjadi hilang (Annas T.S., 2019).



Gambar 2 Ruang Warna HSV
Sumber: (kitainformatika.com, 2020)

Untuk lebih jelasnya bisa dilihat pada gambar diatas. Dapat dilihat bahwa puncak nilai value merupakan seluruhnya warna yang tidak mengandung warna hitam. Sehingga yang mempengaruhi warna adalah hue dan saturation. Sangat berbeda dengan HSL yang apabila L bernilai 1 apapun nilai hue dan nilai saturation tidak dapat mengubah warna. Model warna HSV juga disebut sebagai model warna *user oriented*, karena warna yang dihasilkan merupakan warna sebenarnya dari citra (Annas T.S., 2019).

c. Operasi Morfologi

Proses morfologi menunjukkan operasi tertentu yang dikenakan pada sebuah objek dengan elemen terstruktur. Operasi morfologi adalah teknik pengolahan citra yang didasarkan pada bentuk segmen atau region dalam citra. Operasi ini antara lain meliputi pencarian batas/kontur, dilasi, erosi, penutupan (*closing*), pembukaan (*opening*), dan pengisian (*filling*) (K. Jain, 1989).

2.5 Pengolahan Video Digital

Pengolahan video digital adalah proses manipulasi dan analisis data video dalam domain digital. Ini melibatkan berbagai teknik untuk memanipulasi, menganalisis, dan mengolah data video, seperti kompresi video, deteksi gerakan, pelacakan objek, pengenalan wajah, segmentasi, dan banyak lagi. Tujuan dari pengolahan video digital dapat bervariasi, termasuk untuk keperluan hiburan,

keamanan, medis, penelitian ilmiah, dan industri lainnya. Perbedaan utama antara pengolahan video digital dan pengolahan citra digital adalah dimensi waktu. Pengolahan citra digital umumnya hanya melibatkan manipulasi gambar dalam satu *frame* tunggal, sedangkan pengolahan video digital memperhitungkan urutan waktu dari berbagai *frame* dalam video. Oleh karena itu, teknik dan algoritma yang digunakan dalam pengolahan video digital harus mempertimbangkan aspek temporal dari data, seperti deteksi pergerakan, analisis perubahan antar-*frame*, dan kompresi video. (C. Gonzalez & E. Woods, 2008)

2.6 Deteksi Objek

Deteksi objek adalah sebuah teknik dalam pengolahan citra komputer yang bertujuan untuk mengidentifikasi dan menemukan lokasi serta jenis objek yang ada dalam sebuah gambar atau video. Tujuan utamanya adalah untuk mengidentifikasi keberadaan objek tertentu dalam gambar dan membatasi atau mengelilingi objek dengan kotak pembatas (*bounding box*) yang tepat. Teknik ini sering digunakan dalam berbagai aplikasi seperti pengawasan keamanan, pengenalan wajah, kendaraan otonom, dan lain sebagainya (Ren et al., 2017).

Pada penelitian ini, deteksi objek dilakukan menggunakan teknik deteksi kontur. Kontur dalam pengolahan citra adalah representasi garis luar atau batas dari objek dalam gambar. Deteksi kontur bertujuan untuk mengidentifikasi dan mengekstrak tepi objek yang berbeda dari latar belakangnya. Teknik ini berfokus pada penemuan tepi objek melalui analisis bentuk dan intensitas pada gambar, di mana hasil dari proses ini adalah kumpulan titik yang membentuk kontur objek. Metode deteksi kontur sering digunakan untuk memisahkan objek dari latar belakang, serta untuk menganalisis fitur-fitur penting seperti bentuk, area, dan perimeter dari objek yang terdeteksi (Suzuki & be, 1985). Dalam proses deteksi kontur, beberapa parameter seperti *thresholding*, binarisasi, dan pemilihan metode aproksimasi kontur harus diperhatikan agar kontur yang dihasilkan optimal. Proses dimulai dengan mengidentifikasi tepi objek menggunakan metode seperti *Canny Edge Detection* atau *thresholding*, kemudian kontur yang terbentuk dianalisis untuk menentukan karakteristik objek tersebut, seperti bentuk atau ukuran, yang akan menjadi fokus deteksi. Salah satu metode *thresholding* yang umum digunakan

adalah segmentasi berbasis HSV yang memanfaatkan rentang nilai *upper* dan *lower* ditentukan untuk memilih warna tertentu, yang memisahkan objek dari latar belakangnya.

2.7 Ekstraksi Fitur

Fitur merupakan karakteristik unik dari suatu objek. Fitur dibedakan menjadi dua yaitu fitur alami dan buatan. Fitur alami adalah bagian dari gambar, misalnya kecerahan dan tepi objek. Sedangkan fitur buatan adalah fitur yang diperoleh dengan operasi tertentu pada gambar, misalnya histogram tingkat keabuan. Sehingga ekstraksi fitur adalah proses untuk mendapatkan ciri-ciri pembeda yang membedakan suatu objek dari objek yang lain (Hidayat et al., 2018).

Menurut (Amirullah, 2018) Ekstraksi fitur merupakan suatu proses pengambilan ciri dari gambar digital dengan menghitung jumlah piksel dalam suatu gambar dengan menggunakan metode tertentu. Pada penelitian ini, dua buah metode ekstraksi fitur digunakan, yaitu Ekstraksi fitur warna menggunakan nilai rata-rata HSV dan ekstraksi fitur tekstur menggunakan *Gray Level Co-Occurance Matrix* (GLCM).

2.7.1 Fitur Warna (HSV)

Fitur warna adalah fitur penting dalam citra sebab warna dapat dilihat oleh mata manusia. Untuk membedakan suatu objek dengan warna tertentu dapat menggunakan nilai *hue* yang merupakan representasi dari cahaya tampak (merah, jingga, kuning, hijau, biru, ungu). Nilai *hue* dapat dikombinasikan dengan nilai *saturation* dan *value* yang merupakan tingkat kecerahan suatu warna. Untuk mendapatkan ketiga nilai tersebut, perlu dilakukan konversi ruang warna citra yang semula RGB menjadi HSV.

2.7.2 Fitur Tekstur (GLCM)

Ciri tekstur merupakan ciri penting dalam sebuah gambar yang merupakan informasi berupa susunan struktur permukaan suatu gambar. Teknik statistik yang terkenal untuk ekstraksi fitur adalah matriks *Gray Level Co-occurrence* (GLCM). GLCM merupakan suatu matriks yang menggambarkan frekuensi munculnya pasangan dua piksel dengan intensitas tertentu dalam jarak dan arah tertentu dalam citra (Praseptiyana et al., 2019).

Koordinat pasangan piksel memiliki jarak d dan orientasi sudut θ . Jarak direpresentasikan dalam derajat. Orientasi sudut terbentuk berdasarkan empat arah sudut yaitu 0° , 45° , 90° , dan 135° , dan jarak antar piksel sebesar 1 piksel. Tahapan yang dilakukan pada perhitungan GLCM adalah sebagai berikut:

1. Pembentukan matriks awal GLCM dari pasangan dua piksel yang berjajar sesuai dengan arah 0° , 45° , 90° , dan 135° .
2. Membentuk matriks yang simetris dengan menjumlahkan matriks awal GLCM dengan nilai transposnya.
3. Menormalisasikan matriks GLCM dengan membagi setiap elemen matriks dengan jumlah pasangan piksel.
4. Ekstraksi ciri yaitu:
 - a. Kontras

Kontras adalah perhitungan perbedaan intensitas antara piksel satu dan piksel yang berdekatan diseluruh gambar. Kontras bernilai nol untuk gambar yang konstan (Praseptiyana et al., 2019). Persamaan kontras:

$$kontras = \sum_i^{Ng} \sum_j^{Ng} (i - j)^2 p(i, j) \quad (1)$$

- b. Homogenitas

Homogenitas atau *homogeneity* menunjukkan kehomogenan variasi intensitas dalam citra. Persamaan *homogeneity* sebagai berikut (Praseptiyana et al., 2019) :

$$homogenitas = \sum_i^{Ng} \sum_j^{Ng} \frac{p(i, j)}{1 + (i - j)^2} \quad (2)$$

- c. ASM (Angular Second Moment)

ASM bisa juga disebut *uniformity* atau *energy*. ASM adalah penjumlahan pangkat dari elemen matriks GLCM (Praseptiyana et al., 2019). Adapun persamaan ASM sebagai berikut:

$$ASM = \sum_i^{Ng} \sum_j^{Ng} (p(i,j))^2 \quad (3)$$

d. Energi

Energi merupakan fitur GLCM yang digunakan untuk mengukur konsentrasi pasangan intensitas pada matriks GLCM. Jika semakin tinggi nilai energi, maka tingkat homogenitas tekstur tinggi dan variasi intensitas dalam citra mengecil (Praseptiyana et al., 2019), dan didefinisikan dengan persamaan sebagai berikut:

$$energi = \sqrt{ASM} \quad (4)$$

e. Korelasi

Korelasi (Correlation) menyatakan ukuran hubungan linier dari nilai gray level piksel ketetanggaan. Nilai korelasi dapat dicari menggunakan persamaan (Septiarini et al., 2015):

$$korelasi = \sum_i \sum_j (\sigma_x \sigma_y (i - \mu_x)(j - \mu_y) p(i,j)) \quad (5)$$

f. Disimilaritas

Disimilaritas mengukur ketidakmiripan suatu tekstur, yang akan bernilai besar bila acak dan sebaliknya akan bernilai kecil bila seragam (Yunus, 2020). Berikut persamaannya:

$$dissimilarity = \sum_i \sum_j |i - j| p(i,j) \quad (6)$$

Keterangan persamaan-persamaan diatas:

p : probabilitas (0-1) yaitu elemen matriks

(i,j) : pada baris ke-i serta kolom ke-j

Ng : frekuensi tingkat keabuan citra

$\mu_x \mu_y$: mean nilai keabuan piksel

$\sigma_x \sigma_y$: deviasi standar dari nilai keabuan piksel

2.8 Pearson Correlation

Korelasi adalah ukuran dari asosiasi monotonik antara 2 variabel. Hubungan monotonik antara 2 variabel adalah hubungan dimana saat nilai 1 variabel meningkat, demikian pula nilai variabel lainnya atau saat nilai 1 variabel meningkat, nilai variabel lainnya menurun. Dalam data yang berkorelasi, perubahan dalam magnitudo 1 variabel berhubungan dengan perubahan dalam magnitudo variabel lainnya, baik dalam arah yang sama atau berlawanan (Schober et al., 2018).

Pearson Product Moment Correlation Coefficient adalah suatu metode yang menggambarkan hubungan antara 2 variabel. Titik-titik data 2 variabel menentukan arah atau tanda nilai koefisien korelasi (r) (Zinzenoff Okwonu et al., 2020). Korelasi yang sempurna adalah -1 atau $+1$ artinya semua titik data berada tepat pada garis lurus (Schober et al., 2018b). Korelasi dapat positif, yang artinya searah jika variabel pertama besar, maka variabel kedua semakin besar juga. Korelasi negatif, yang artinya berlawanan arah jika variabel pertama besar, maka variabel kedua semakin mengecil (Jonathan Sarwono, 2018). Arti nilai koefisien korelasi ditunjukkan pada Tabel 2.

Tabel 2. Interpretasi nilai korelasi pearson

Nilai Koefisien Korelasi (r)	Kekuatan Korelasi
$< 0,25$	Hubungan dapat dianggap tidak ada
$0,25 - 0,50$	Hubungan rendah
$0,50 - 0,80$	Hubungan cukup
$> 0,80 - 1,00$	Hubungan tinggi

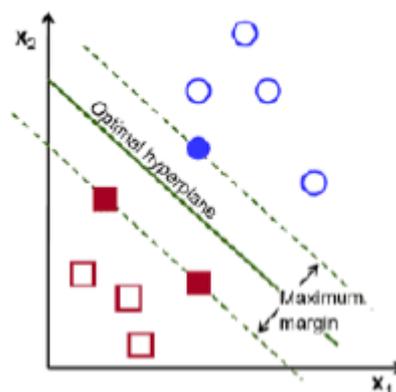
Sumber: (Jonathan Sarwono, 2018)

2.9 Klasifikasi dengan *Support Vector Machine* (SVM)

Klasifikasi adalah proses pengelompokan yang artinya mengumpulkan benda atau entitas yang sama dan memisahkan objek atau entitas yang tidak sama. Menurut (Sumarlin, 2015) Klasifikasi merupakan proses penemuan model (fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar

bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui.

Salah satu metode statistik yang dapat digunakan untuk klasifikasi adalah *Support Vector Machine* (SVM). SVM merupakan salah satu teknik dalam proses klasifikasi yang relatif baru, teknik ini sudah biasa dipakai para peneliti untuk menuntaskan masalah yang berkaitan dengan komputasi. Konsep dasar dari SVM adalah melalui pembentukan *hyperplane* (maximal margin hyperplane) (Prawira et al., 2021). SVM merupakan pembelajaran yang mengarah ke pemrograman kuadratik dengan kendala linear. Berdasarkan minimalisasi risiko prinsip terstruktur, SVM berusaha untuk meminimalkan batas atas kesalahan generalisasi bukan kesalahan empiris, sehingga model prediksi baru efektif menghindari *over-pas* masalah. Selain itu, model SVM bekerja di ruang fitur berdimensi tinggi yang dibentuk oleh pemetaan non-linear dari N-dimensi vektor input x ke dalam ruang fitur K-dimensi ($K > N$) melalui penggunaan fungsi ϕ non-linear (x) (Li et al., 2015).



Gambar 3 Penentuan Hyperplane Terbaik

Sumber: (Prasetyo, 2012)

Hyperplane (batas keputusan) pemisah terbaik antara kedua kelas dapat ditemukan dengan mengukur margin hyperplane tersebut dan mencari titik maksimalnya. Margin adalah jarak antara hyperplane tersebut dengan data terdekat dari masing-masing kelas. Garis solid pada Gambar 3 di atas menunjukkan *hyperplane* yang terbaik, yaitu yang terletak tepat pada tengah-tengah kedua kelas, sedangkan data lingkaran dan bujur sangkar yang dilewati garis batas margin (garis

putus-putus) adalah *support vector*. (Prasetyo, 2012). Rumus perhitungan *hyperplane* algoritma SVM (Putri, 2023):

$$f : w \cdot x + b = 0 \quad (7)$$

Keterangan:

f = fungsi *hyperplane*

w = parameter *hyperplane* (garis tegak lurus antara garis *hyperplane* dan titik *support vector*)

x = data input SVM (x_1 = index kata, x_2 = bobot kata)

b = parameter *hyperplane* nilai bias

Dalam SVM terdapat fungsi kernel trik yang terbagi dua yaitu kernel linear dan kernel non-linear. Kernel linear digunakan ketika data yang akan diklasifikasi dapat terpisah dengan sebuah garis (*hyperplane*). Sedangkan kernel non-linear digunakan ketika data hanya dapat dipisahkan dengan garis lengkung atau sebuah bidang pada ruang dimensi tinggi. Contoh kernel non-linear adalah *Polynomial*, *Radial Basis Function* (RBF), *Sigmoid*, *Additive* (Rachmat, 2018).

Fungsi kernel yang umum digunakan pada metode SVM adalah:

1. Kernel Linear

$$K(x, y) = x \cdot y \quad (8)$$

2. Kernel *Polynomial*

$$K(x, y) = (\gamma(x \cdot y) + r)^d \quad (9)$$

3. Kernel *Radial Basis Function* (RBF)

$$K(x, y) = \exp(-\gamma \|x - y\|^2) \quad (10)$$

4. Kernel *Sigmoid*

$$K(x, y) = \tanh(\gamma(x \cdot y) + r) \quad (11)$$

Keterangan:

- (x, y) adalah vektor input
- r adalah oefisien bebas yang digunakan pada kernel polinomial dan sigmoid untuk memperluas fleksibilitas pemetaan.
- d adalah derajat (degree) dari kernel polinomial, yang menentukan kompleksitas dari polinomial.

- *Exp* adalah fungsi eksponensial, digunakan dalam kernel RBF untuk menentukan seberapa cepat fungsi meluruh.
- *Tanh* pada kernel sigmoid adalah fungsi tangens hiperbolik, digunakan dalam kernel sigmoid untuk transformasi non-linear.
- γ (gamma): Parameter ini menentukan seberapa besar pengaruh setiap contoh data tunggal. Nilai γ yang lebih tinggi menyebabkan model lebih fokus pada contoh data yang dekat dengan margin.
- C merupakan parameter regulasi dalam SVM yang mengontrol *trade-off* antara margin yang lebih besar dan *misclassification error* (kesalahan klasifikasi). Semakin besar nilai C , semakin sedikit penalti untuk kesalahan, yang menyebabkan model cenderung *overfitting*. C menentukan kompleksitas *decision surface*, nilai kecil akan memperhalus *decision surface*, sedangkan nilai tinggi akan menargetkan semua data latih diklasifikasikan dengan benar.

Pemilihan fungsi kernel yang tepat merupakan hal yang sangat penting karena akan menentukan *feature space* dimana fungsi *classifier* akan dicari. Menurut (Schölkopf & Smola, 2001) fungsi kernel RBF memiliki kelebihan yaitu secara otomatis menentukan nilai, lokasi dari center serta nilai pembobot dan bisa mencakup nilai rentang tak terhingga. RBF juga efektif menghindari *overfitting* dengan memilih nilai yang tepat untuk parameter C dan γ dan RBF baik digunakan ketika tidak ada pengetahuan terdahulu. Fungsi kernel yang direkomendasikan adalah fungsi kernel RBF karena dapat memetakan hubungan tidak linear, RBF lebih *robust* terhadap *outlier* karena fungsi kernel RBF berada antara selang 0 – 1, sedangkan fungsi kernel yang lain memiliki rentang antara -1 sampai dengan 1. (Chih-Wei Hsu, 2003).

2.10 Confusion Matrix

Confusion matrix adalah tabel yang digunakan untuk mengukur kinerja dari algoritma klasifikasi. *Confusion matrix* terdiri dari empat kondisi yang digunakan untuk menentukan metrik pengukuran pengklasifikasi. Keempat kondisi tersebut adalah (P. Singh et al., 2021):

1. TP (*True Positive*): TP mewakili data positif yang terklasifikasi dengan benar sebagai data positif.
2. TN (*True Negative*): TN mewakili data negatif yang terklasifikasi dengan benar sebagai data negatif.
3. FP (*False Positive*): FP mewakili data negatif yang terklasifikasi sebagai data positif.
4. FN (*False Negative*): FN mewakili data positif yang terklasifikasi sebagai data negatif.

Metrik kinerja dari sebuah algoritma adalah *accuracy*, *precision*, *recall*, dan *F1 score*, yang dihitung berdasarkan nilai dari TP, TN, FP, dan FN (P. Singh et al., 2021).

1. Accuracy

Accuracy dari suatu algoritma direpresentasikan sebagai rasio data yang diklasifikasikan dengan benar dengan jumlah total data. Berikut persamaan *akurasi*:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (12)$$

2. Precision

Precision direpresentasikan sebagai rasio data positif yang terklasifikasi dengan benar (TP) dibagi total data yang terklasifikasi sebagai data positif. Berikut persamaan *precision*:

$$Precision = \frac{(TP)}{(TP + FP)} \quad (13)$$

3. Recall

Recall didefinisikan sebagai rasio data positif yang terklasifikasi dengan benar dibagi dengan total data positif. Berikut persamaan *recall*:

$$Recall = \frac{(TP)}{(TP + FN)} \quad (14)$$

4. F1 Score

F1 *score* dikenal sebagai F *Measure*. F1 *score* menyatakan keseimbangan antara *precision* dan *recall*. Berikut persamaan F1 *Score*:

$$F1\ Score = \frac{2 \times precision \times recall}{precision + recall} \quad (15)$$