

**SKRIPSI**

**PENERAPAN METODE VIOLA-JONES, FACENET, DAN  
DLIB DALAM SISTEM DETEKSI KANTUK DAN  
KLASIFIKASI SISWA SAAT BELAJAR DI KELAS**

**Disusun dan diajukan oleh:**

**MARCELLINO PIRONO  
D121 19 1044**



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

## LEMBAR PENGESAHAN SKRIPSI

### PENERAPAN METODE VIOLA-JONES, FACENET, DAN DLIB DALAM SISTEM DETEKSI KANTUK DAN KLASIFIKASI SISWA SAAT BELAJAR DI KELAS

Disusun dan diajukan oleh


**Marcellino Pirono**  
**D121191044**


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 Agustus 2024  
dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama,

Pembimbing Pendamping,

  
Dr. Ir. Ingrid Nurtanio, M.T.  
NIP 196108131988112001

  
Prof. Dr. Eng. Intan Sari Areni., S.T., M.T.  
NIP 197502032000122002

Ketua Program Studi,

  
Prof. Dr. Ir. Indrabayu, S.T., M.T., M.Bus.Sys., IPM, ASEAN. Eng.  
NIP 197507162002121004



## PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini ;  
Nama : Marcellino Pirono  
NIM : D121191044  
Program Studi : Teknik Informatika  
Jenjang : SI

Menyatakan dengan ini bahwa karya tulisan saya berjudul

Penerapan Metode Viola-Jones, Facenet, dan Dlib dalam Sistem Deteksi Kantuk dan Klasifikasi Siswa Saat Belajar di Kelas

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, 10 Agustus 2024

Yang Menyatakan



Marcellino Pirono

## ABSTRAK

**MARCELLINO PIRONO.** *PENERAPAN METODE VIOLA-JONES, FACENET, DAN DLIB DALAM SISTEM DETEKSI KANTUK DAN KLASIFIKASI SISWA SAAT BELAJAR DI KELAS* (dibimbing oleh Dr. Ir. Ingrid Nurtanio, M.T. dan Prof. Dr.Eng Intan Sari Areni., ST., MT.)

Kantuk saat belajar di kelas merupakan masalah yang umum terjadi di kalangan siswa. Hal ini dapat mengurangi konsentrasi dan mempengaruhi pemahaman materi pelajaran.

Tujuan penelitian ini adalah membangun sistem deteksi kantuk pada siswa saat belajar di kelas dengan menggunakan metode *Viola-Jones*, *FaceNet*, dan *Dlib*. Sistem ini memanfaatkan metode *Viola-Jones* untuk deteksi wajah secara *real time*, *FaceNet* untuk pengenalan wajah sehingga setiap siswa dapat diidentifikasi secara unik, dan *Dlib* untuk mendeteksi estimasi pose kepala serta menghitung rasio mata (*Eye Aspect Ratio*) dan rasio mulut (*Mouth Aspect Ratio*). *Eye Aspect Ratio* digunakan untuk mendeteksi kondisi mata tertutup dan *Mouth Aspect Ratio* untuk mendeteksi aktivitas menguap.

Pada penelitian ini, penerapan metode *Viola-Jones*, *FaceNet*, dan *Dlib* dilakukan melalui beberapa tahapan yaitu input data, preprocessing data, deteksi wajah menggunakan metode *Viola-Jones*, normalisasi citra, rekognisi wajah menggunakan metode *FaceNet*, melakukan perhitungan *Region Eye* dan *Mouth Rasio* menggunakan metode *Dlib*, serta klasifikasi siswa kantuk. Ambang batas EAR ditetapkan sebesar 0,20 dan MAR sebesar 15.

Hasil pengujian deteksi wajah diperoleh akurasi sebesar 83%. Hasil pengujian pengenalan wajah diperoleh akurasi sebesar 92,8%. Untuk pendeteksian kantuk, hasil akurasi yang diperoleh adalah sebesar 99,6%.

Kata Kunci: *Dlib*, *Facenet*, Kantuk, *Machine Learning*, *Viola-Jones*

## ABSTRACT

**MARCELLINO PIRONO.** *APPLICATION OF VIOLA-JONES, FACENET, AND DLIB ALGORITHMS IN A SYSTEM FOR DROWSINESS DETECTION AND CLASSIFICATION OF STUDENTS WHILE STUDYING IN CLASS* (supervised by Dr. Ir. Ingrid Nurtanio, M.T. and Prof. Dr.Eng Intan Sari Areni, ST., MT.)

Drowsiness during class is a common problem among students. This can reduce concentration and affect the understanding of the subject matter.

The purpose of this research is to build a drowsiness detection system for students while studying in class using the Viola-Jones, FaceNet, and Dlib methods. This system utilizes the Viola-Jones method for real time face detection, FaceNet for face recognition so that each student can be uniquely identified, and Dlib to detect head pose estimation and calculate eye ratio (Eye Aspect Ratio) and mouth ratio (Mouth Aspect Ratio). Eye Aspect Ratio is used to detect closed eye conditions and Mouth Aspect Ratio to detect yawning activity.

In this research, the application of the Viola-Jones, FaceNet, and Dlib methods is carried out through several stages, namely data input, data preprocessing, face detection using the Viola-Jones method, image normalization, face recognition using the FaceNet method, calculating the Eye and Mouth Ratio Region using the Dlib method, and classification of drowsy students. The EAR threshold is set at 0.20 and MAR at 15.

Face detection test results obtained an accuracy of 83%. Face recognition test results obtained an accuracy of 92.8%. For drowsiness detection, the accuracy obtained is 99.6%.

Keywords: Dlib, Drowsiness, Facenet, Machine Learning, Viola-Jones

## DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI.....	i
PERNYATAAN KEASLIAN.....	ii
ABSTRAK .....	iii
ABSTRACT.....	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR .....	vii
DAFTAR TABEL.....	ix
DAFTAR SINGKATAN DAN ARTI SIMBOL .....	x
DAFTAR LAMPIRAN.....	xi
KATA PENGANTAR .....	xii
<b>BAB I PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan Penelitian .....	3
1.4 Manfaat Penelitian .....	3
1.5 Ruang Lingkup.....	3
<b>BAB II TINJAUAN PUSTAKA.....</b>	<b>4</b>
2.1 Kantuk.....	4
2.1.1 Parameter Mengantuk .....	4
2.2 Visi Komputer.....	5
2.3 Pengolahan Citra .....	7
2.4 Video Digital.....	8
2.4 Face Detection.....	9
2.5 Face Recognition.....	9
2.6 Viola Jones .....	9
2.6.1 Integral Image .....	10
2.6.2 Haar-like Feature.....	11
2.6.3 Adaboost .....	12
2.6.4 Cascade Classifier .....	14
2.7 Local Binary Pattern .....	14
2.7.1 Local Binary Pattern Histogram.....	16
2.7.2 Parameter LBP .....	16
2.8 Facenet .....	17
2.9 Dlib.....	18
2.10 Precision, Recall, dan F1-Score .....	20
2.11 Python .....	20
2.12 OpenCV.....	21
<b>BAB III METODE PENELITIAN/PERANCANGAN .....</b>	<b>22</b>
3.1 Lokasi Penelitian.....	22

3.2	Alat dan Bahan.....	22
3.3	Tahapan Penelitian.....	23
3.4	Teknik Pengambilan Data.....	24
3.5	Perancangan Implementasi Sistem.....	28
3.5.1	Proses Training.....	32
3.5.2	Proses Testing.....	40
3.6	Pengujian Sistem.....	50
BAB IV HASIL DAN PEMBAHASAN.....		51
4.1	Hasil Penelitian.....	51
4.1.1	Hasil Pengujian Sistem.....	51
4.1.2	Pembahasan.....	56
BAB V KESIMPULAN DAN SARAN.....		63
5.1	Kesimpulan.....	63
5.2	Saran.....	64
DAFTAR PUSTAKA.....		65
LAMPIRAN.....		68

## DAFTAR GAMBAR

Gambar 1 Proses pengenalan pola .....	8
Gambar 2 Nilai pixel pada citra .....	10
Gambar 3 Fitur Wajah.....	11
Gambar 4 Fitur mata, mulut, dan hidung .....	11
Gambar 5 Haar like feature .....	12
Gambar 6 <i>Classifier</i> Lemah .....	12
Gambar 7 Hasil penambahan bobot dari <i>Classifier</i> lemah.....	13
Gambar 8 Hasil kombinasi dari <i>Classifier</i> Lemah.....	13
Gambar 9 Cara kerja metode <i>Local Binary Pattern</i> .....	14
Gambar 10 Facenet Arsitektur .....	17
Gambar 11 <i>Facial landmarks</i> .....	19
Gambar 12 Tahapan Penelitian .....	23
Gambar 13 Ilustrasi proses pengambilan citra sebagai data <i>training</i> .....	25
Gambar 14 Sampel citra latih diambil dari lima sisi: (a) menghadap kedepan, (b) menghadap kiri, (c) menghadap kanan, (d) menghadap atas, (e) menghadap bawah.....	25
Gambar 15 Ilustrasi proses <i>testing</i> .....	26
Gambar 16 Pengujian sistem yang diambil langsung menggunakan kamera <i>smartphone</i> .....	26
Gambar 17 Batas sudut pengujian sistem dari -30 hingga 30 derajat.....	27
Gambar 18 Ilustrasi urutan siswa dalam pengujian sistem .....	27
Gambar 19 Blok diagram perancangan sistem.....	28
Gambar 20 Alur <i>training</i> dan <i>testing</i> model .....	29
Gambar 21 Flowchart Sistem.....	30
Gambar 22 Beberapa data sampel.....	32
Gambar 23 (a) Citra awal berwarna BGR (b) Perubahan warna pada citra dari BGR ke grayscale.....	33
Gambar 24 Nilai <i>pixel</i> pada <i>sliding window</i> .....	34
Gambar 25 Hasil <i>integral image</i> .....	34
Gambar 26. (a) Area perhitungan gelap terang, (b) Angka yang diambil sebagai perhitungan.....	35
Gambar 27 Hasil deteksi wajah.....	37
Gambar 28 <i>Cropping</i> pada area wajah.....	37
Gambar 29 <i>Resize</i> citra wajah yang telah dideteksi ke 200x200 pixel .....	38
Gambar 30 <i>FaceNet</i> mengambil gambar wajah sebagai masukan dan mengeluarkan vektor <i>embedding</i> . .....	39
Gambar 31 Nama dan id siswa disimpan kedalam <i>database</i> MySQL.....	40
Gambar 32 Perubahan warna dari <i>RGB</i> ke <i>Grayscale</i> .....	41
Gambar 33 Proses <i>sliding window</i> .....	41
Gambar 34 Hasil deteksi wajah menggunakan algoritma <i>Viola-Jones</i> .....	42
Gambar 35 Hasil preprocessing pada wajah yang berhasil dideteksi, dengan melakukan <i>crop</i> , <i>resize</i> , dan penambahan <i>batch size</i> .....	42
Gambar 36 vektor <i>embedding</i> mencari jarak dari titik-titik yang sudah ada dalam <i>database</i> .....	43
Gambar 37 Hasil pengenalan wajah menggunakan algoritma <i>Facenet</i> .....	44



Gambar 38 Model fitur wajah 3D .....	45
Gambar 39 Hasil pendeteksian estimasi pose wajah.....	46
Gambar 40 <i>Landmark</i> Wajah.....	46
Gambar 41 Hasil pendeteksian mata dan mulut menggunakan algoritma <i>Dlib</i> ....	47
Gambar 42 <i>Region of eye</i> .....	47
Gambar 43 Hasil Klasifikasi Kantuk .....	49
Gambar 44 Hasil pendeteksian kantuk yang tersimpan kedalam <i>database</i> MySQL.....	50
Gambar 45 Ilustrasi siswa menguap selama 5 detik sehingga terdeteksi sebagai mengantuk.....	59
Gambar 46 Hasil deteksi kantuk pada siswa.....	59
Gambar 47 Ilustrasi kesalahan deteksi sistem yang dimana siswa lagi menghadap kebawah karena sedang menulis, tapi sistem mendeteksi kantuk.....	60
Gambar 48 Hasil kesalahan deteksi oleh sistem yang mengindikasikan bahwa siswa mengantuk .....	60
Gambar 49 Ilustrasi siswa saat menyipitkan mata saat melihat ke papan tulis namun tidak terdeteksi sebagai mengantuk.....	61
Gambar 50 Hasil deteksi sistem yang mengindikasikan siswa tidak mengantuk .....	61
Gambar 51 ilustrasi kesalahan dikarenakan siswa menguap hanya tiga detik dan gerakan yang berlebih saat menguap.....	62
Gambar 52 Hasil bahwa sistem gagal mengidentifikasi siswa sebagai mengantuk .....	62

## DAFTAR TABEL

Tabel 1 Hasil pengukuran Rasio Mata .....	51
Tabel 2 Hasil pengukuran Rasio Mulut .....	52
Tabel 3 Hasil Deteksi Wajah menggunakan algoritma Viola-Jones.....	52
Tabel 4 Hasil Rekognisi wajah dengan algoritma <i>Facenet</i> .....	53
Tabel 5 Hasil Rekognisi wajah dengan algoritma LBPH .....	53
Tabel 6 Hasil deteksi kantuk .....	53
Tabel 7 Hasil bias deteksi kantuk (Data Normal) .....	54
Tabel 8 Hasil bias deteksi kantuk (Data pengujian dengan augmentasi <i>brightness</i> 10 dan <i>contrast</i> 1.1).....	55
Tabel 9 Hasil bias deteksi kantuk (Data pengujian dengan augmentasi <i>brightness</i> -20 dan <i>contrast</i> 1.3).....	55
Tabel 10 Hasil bias deteksi kantuk (Augmentasi pada dataset seperti <i>flip</i> , <i>rotate</i> , <i>brightness</i> , dan <i>contrast</i> ) .....	55

## DAFTAR SINGKATAN DAN ARTI SIMBOL

---

Lambang/Singkatan	Arti dan Keterangan
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
BGR	<i>Blue Green Red</i>
LBP	<i>Local Binary Pattern</i>
LBPH	<i>Local Binary Pattern Histogram</i>
EAR	<i>Eye Aspect Ratio</i>
MAR	<i>Mouth Aspect Ratio</i>
RGB	Red Green Blue

**DAFTAR LAMPIRAN**

Lampiran 1 <i>Consent Form</i> .....	68
Lampiran 2 <i>Preview Dataset</i> .....	69
Lampiran 3 <i>Source code</i> .....	71

## KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Tuhan yang Maha Esa, yang telah memberikan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan tugas akhir ini yang berjudul “Penerapan Metode Viola-Jones, Facenet, dan Dlib Dalam Sistem Deteksi Kantuk dan Klasifikasi Siswa Saat Belajar di Kelas” sebagai salah satu syarat dalam menyelesaikan jenjang Strata-1 di Departemen Teknik Informatika, Fakultas Teknik, Universitas Hasanuddin.

Penulis menyadari bahwa dalam penyusunan dan penulisan tugas akhir ini tidak dapat terselesaikan dengan baik tanpa adanya bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih banyak kepada:

1. Kedua Orang tua penulis, Bapak Benny Pirono dan Ibu Emy Cendhana yang selalu memberikan dukungan, doa, dan memberikan semangat kepada penulis, serta selalu sabar dalam mendidik penulis sejak kecil.
2. Ibu Dr. Ir. Ingrid Nurtanio, M.T. selaku pembimbing I dan Ibu Prof. Dr.Eng Intan Sari Areni., ST., MT. Selaku pembimbing II, yang senantiasa menyediakan waktu, tenaga, pikiran, dan perhatian yang luar biasa dalam mengarahkan penulis untuk menyelesaikan tugas akhir ini.
3. Segenap Dosen dan Staf Departemen Teknik Informatika Fakultas Teknik Universitas Hasanuddin yang telah banyak membantu selama proses perkuliahan hingga penyelesaian tugas akhir ini.
4. Kepala Sekolah dan Guru kelas VI SD Frater Thamrin Makassar yang telah banyak membantu penulis dalam pengambilan data penelitian, mulai dari perizinan pengambilan dan pengujian sistem, uji coba aplikasi, hingga mengarahkan para siswa.
5. Fadhil, Rajab, Arfandy, Dita, Artia, Sila, Brill, Reinhart yang telah memberikan semangat kepada penulis dalam menyelesaikan tugas akhir ini.
6. Teman - Teman Signifier 2019 atas dukungan, bantuan, dan semangat yang telah diberikan selama ini.
7. Serta seluruh pihak yang tidak disebutkan satu per satu, tanpa sadar telah menjadi motivasi penulis dalam menyelesaikan tugas akhir.

Akhirnya dengan segala kerendahan hati, penulis menyadari bahwa tugas akhir ini masih banyak terdapat kesalahan dan kekurangan dalam penyusunan baik dari isi maupun cara penyajiannya. Oleh karena itu penulis mengharapkan adanya saran dan kritik yang bersifat membangun demi kesempurnaan tugas akhir ini. Penulis berharap semoga tugas akhir ini dapat memberikan manfaat bagi pengembangan ilmu selanjutnya.

Gowa, 30 Juni 2024

Penulis  
Marcellino Pirono

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Di era digital saat ini, teknologi telah menjadi bagian tak terpisahkan dari proses pembelajaran, menawarkan berbagai kemudahan dan inovasi dalam menyampaikan materi pendidikan. Penggunaan teknologi dalam pendidikan, seperti penggunaan perangkat digital, aplikasi pembelajaran, dan platform online, telah membuka peluang baru dalam meningkatkan kualitas pengajaran dan interaksi antara siswa dan guru. Namun, di balik berbagai manfaat yang ditawarkan, terdapat pula tantangan yang muncul dari integrasi teknologi ini, salah satunya adalah masalah kantuk saat belajar di kelas.

Dalam konteks pendidikan, keterlibatan siswa dalam proses belajar mengajar menjadi salah satu faktor penting yang menentukan keberhasilan pembelajaran. Dr. Ken Shore, seorang ahli pendidikan, menggarisbawahi sebuah fenomena yang terjadi di lingkungan kelas, terutama pada tingkat pendidikan dasar. Beliau menekankan bahwa siswa pada jenjang ini jarang ditemukan tertidur selama kegiatan belajar berlangsung. Namun, pada kesempatan langka ketika seorang siswa tertidur di kelas, situasi tersebut tidak hanya berpotensi mengganggu proses pembelajaran bagi siswa yang bersangkutan tetapi juga dapat menjadi sumber gangguan bagi guru dan siswa lainnya. Lebih lanjut, Dr. Shore menambahkan bahwa kejadian siswa tertidur di kelas bukan sekadar masalah kecil yang dapat diabaikan. Fenomena ini dapat menjadi indikator adanya masalah yang lebih serius yang memerlukan perhatian dan penanganan. Tidur di kelas mungkin merupakan tanda bahwa siswa tersebut mengalami masalah kesehatan, seperti kurangnya kualitas tidur malam atau kondisi medis tertentu. Alternatif lain, hal ini juga bisa mengindikasikan adanya masalah psikologis atau sosial yang dialami siswa, termasuk stres, kecemasan, atau bahkan depresi (Ken Shore, 2003).

Penelitian yang dilakukan oleh Mohanty dan Hedge yang berjudul *Design of Real-time Drowsiness Detection System using Dlib* menggunakan Dlib model. Metrik kuantitatif yang digunakan dalam algoritma yang diusulkan adalah *Eye Aspect Ratio* (EAR) untuk memantau pola kedipan mata pengemudi dan *Mouth*

*Aspect Ratio* (MAR) untuk menentukan apakah pengemudi menguap dalam video yang diambil secara *real time*. Akurasi pengujian waktu nyata rata-rata yang diperoleh dengan menggunakan Dlib untuk mata dan menguap adalah 82,02% dan 85,44%, serta 93,25% dan 96,71% untuk video yang telah direkam sebelumnya (Mohanty, 2019).

Penelitian terkait dalam mendeteksi kantuk yang pernah dilakukan oleh Siti Khumerah Mufti yang berjudul Sistem Rekognisi Kantuk pada Pengendara Mobil Berbasis Android, menggunakan metode *Haar-Cascade* untuk mendeteksi wajah serta deteksi mulut dan *FindContour* untuk melacak pergerakan dari area gelap di mulut. Pada penelitian tersebut, jumlah frame diukur jika dalam 5-7 detik, sistem mendeteksi setidaknya 35 frame yang mengandung mulut terbuka, maka akan diputuskan sebagai menguap dan notifikasi akan muncul. Hasilnya mencapai akurasi 88,7% (Mufti.K.S, 2018).

Berdasarkan penelitian diatas, penelitian ini akan mengembangkan suatu sistem deteksi kantuk pada siswa saat belajar. Sistem ini akan menggunakan metode Viola-Jones untuk mendeteksi wajah, Facenet untuk mengenali wajah, dan Dlib untuk mendeteksi mata dan bibir siswa yang berguna untuk menentukan kantuk atau tidaknya dari siswa. Pendeteksian kantuk dilakukan di ruangan kelas saat proses belajar mengajar. Kamera diletakkan di samping papan tulis menghadap siswa untuk mengambil video secara *real time*.

Dengan adanya aplikasi ini, diharapkan dapat membantu guru untuk mengetahui siswa yang mengantuk dan memilih metode yang tepat dalam proses belajar mengajar. Selain itu, siswa juga diharapkan dapat lebih fokus belajar dan lebih pandai manajemen waktu. Oleh karena itu, hal ini akan menghasilkan proses akademis yang efisien dan mencetak generasi muda berprestasi untuk kemajuan Indonesia.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang masalah yang telah diuraikan, maka dapat dirumuskan permasalahan pada penelitian ini antara lain:



1. Bagaimana membangun suatu sistem deteksi kantuk dan klasifikasi pada siswa dalam belajar di ruangan kelas menggunakan metode Viola-Jones, Facenet, dan Dlib?
2. Bagaimana unjuk kerja sistem deteksi kantuk dan klasifikasi pada siswa dalam belajar di ruangan kelas menggunakan metode Viola-Jones, Facenet, dan Dlib?

### **1.3 Tujuan Penelitian**

Tujuan yang ingin dicapai pada penelitian ini adalah:

1. Untuk membangun suatu sistem deteksi kantuk dan klasifikasi pada siswa dalam belajar di ruangan kelas menggunakan metode Viola-Jones, Facenet, dan Dlib.
2. Untuk mengetahui kinerja sistem deteksi kantuk dan klasifikasi pada siswa dalam belajar di ruangan kelas menggunakan metode Viola-Jones, Facenet, dan Dlib.

### **1.4 Manfaat Penelitian**

Dengan dilakukannya penelitian ini, manfaat yang diharapkan adalah:

1. Membantu pendidik untuk menentukan metode pembelajaran yang tepat kepada pelajar guna meningkatkan minat dan konsentrasi pelajar.
2. Sebagai referensi dalam pengembangan topik terkait untuk sistem deteksi kantuk pada bidang Visi Komputer.

### **1.5 Ruang Lingkup**

Batasan masalah dari penelitian ini adalah:

1. Maksimal jarak untuk mengukur yaitu hingga 4 meter.
2. Hanya dapat mendeteksi di ruangan dengan cahaya yang cukup.
3. Pengukuran hanya terfokus pada wajah.

## **BAB II TINJAUAN PUSTAKA**

### **2.1 Kantuk**

Kantuk, atau sensasi mengantuk, adalah pengalaman fisik dan psikologis yang umumnya muncul sebagai respons alami terhadap kelelahan atau kurang tidur. Sensasi ini membawa perasaan mengantuk dan mengundang keinginan untuk beristirahat atau tidur. Ada beberapa faktor yang dapat menyebabkan kantuk, seperti kelelahan bekerja dan kurang tidur. Ada berbagai tanda yang menunjukkan bahwa seseorang sedang mengantuk, seperti kelopak mata yang terasa berat, penglihatan yang kabur, dan kepala yang tidak seimbang. Dalam kondisi ini, diperlukan waktu berbaring dan istirahat untuk mengatasi kantuk (Munawaroh dkk, 2021).

#### **2.1.1 Parameter Mengantuk**

##### **a. Kedipan Mata**

Mendeteksi kantuk pada siswa dapat dilakukan dengan menganalisis frekuensi dan durasi kedipan mata mereka. Proses ini melibatkan pengamatan terhadap seberapa sering dan berapa lama mata siswa berkedip selama sesi belajar mengajar. Berdasarkan penelitian yang dilakukan oleh Phillip P. Caffier (2005), tingkatan kantuk dapat diidentifikasi dengan mengelompokkan durasi kedipan mata. Dalam studi tersebut, ditemukan bahwa durasi kedipan mata rata-rata biasanya kurang dari 400 milidetik, sedangkan durasi kedipan minimum adalah sekitar 75 milidetik. Namun, ketika seseorang berada dalam keadaan mengantuk, durasi kedipan mata cenderung meningkat secara signifikan. Hal ini karena refleks kedipan menjadi lebih lambat akibat menurunnya tingkat kewaspadaan dan kelelahan yang dirasakan oleh siswa.

##### **b. Mulut Menguap**

Mulut menguap dapat dijadikan parameter untuk mendeteksi kantuk pada siswa. Ketika menguap, mulut akan terbuka, menarik napas dalam-dalam

melalui mulut dan hidung, dan menghembuskan napas secara perlahan. Terkadang manusia mungkin meregangkan lengan dan memiringkan kepala ke belakang saat menguap. Biasanya, menguap berlangsung antara lima hingga 10 detik. Dikutip dari National Sleep Foundation, Para ilmuwan telah mengidentifikasi sejumlah kemungkinan alasan mengapa manusia menguap sebagai berikut (National Sleep Foundation, 2023):

- Kantuk atau Mengantuk

Manusia biasanya menguap saat mengantuk atau sangat mengantuk. Menguap menandakan bahwa tubuh butuh istirahat. Menguap adalah bagian dari proses tubuh manusia untuk tetap terjaga.

- Bosan

Menguap juga biasanya dikaitkan dengan kebosanan. Jika lingkungan sekitar tidak menstimulasi, maka manusia akan merasa mengantuk. Menguap dapat memberi sinyal kepada orang lain bahwa kita sedang bosan atau lelah.

- Mengatur Suhu Otak

Menguap dapat membantu mendinginkan otak yang kepanasan. Satu penelitian kecil pada manusia menemukan bahwa orang menguap lebih banyak di musim panas

- Melihat orang lain menguap

Beberapa peneliti percaya bahwa menguap adalah bagian dari respons empati. Dalam penelitian pada manusia dan hewan, menguap dari satu hewan telah terbukti memicu menguap pada hewan lain dari spesies yang sama. Namun, penelitian masih terbatas, sehingga hubungan antara menguap dan empati memerlukan penelitian lebih lanjut.

## **2.2 Visi Komputer**

Visi komputer (*computer vision*) adalah bidang ilmu komputer yang bertujuan untuk mengembangkan teknik dan algoritma yang memungkinkan komputer untuk memproses, menganalisis, dan memahami informasi visual. (*“Computer Vision: Algorithms and Applications”*, Richard Szeliski). Berbagai

macam aplikasi dunia nyata menggunakan computer vision meliputi (Szeliski, 2010):

- *Optical Character Recognition (OCR)* : umumnya digunakan untuk mengenali gambar huruf dengan tujuan mengonversinya menjadi teks, seperti mengidentifikasi kode pos yang ditulis tangan pada surat atau melakukan pengenalan otomatis nomor plat (ANPR).
- *Medical Imaging (Pencitraan Medis)* : mencakup pencatatan atau perekaman gambar sebelum dan selama operasi untuk keperluan klinis atau melakukan penelitian jangka panjang pada morfologi bagian tubuh manusia.
- *3D Model Building* : konstruksi secara otomatis dibangun dari foto udara yang kemudian digunakan dalam sistem seperti Bing Maps.
- *Automotive Safety* : melibatkan deteksi hambatan tak terduga, seperti pejalan kaki di jalan raya.
- *Surveillance* : melibatkan pemantauan untuk mendeteksi intrusi dan menganalisis lalu lintas jalan raya.
- *Face Detection* : meningkatkan penekanan pada fokus kamera dan mencari citra yang lebih relevan.

Visi komputer, sebagai cabang ilmu komputer yang berkembang pesat, membahas tentang pengembangan teknologi dan algoritma untuk memungkinkan komputer memahami dan menginterpretasi informasi visual dari dunia nyata. Ini mencakup sejumlah tugas yang melibatkan pengenalan pola, pemrosesan citra, dan analisis visual.

Visi komputer juga memiliki aplikasi luas, seperti dalam bidang kedokteran, dengan teknologi pencitraan medis yang memungkinkan pencatatan dan analisis citra pre-operatif dan intra-operatif. Di bidang keamanan otomotif, visi komputer digunakan untuk mendeteksi hambatan tak terduga di jalan, seperti pejalan kaki. Penerapan visi komputer juga terlihat dalam pemantauan keamanan, di mana sistem digunakan untuk mendeteksi intrusi dan menganalisis lalu lintas.

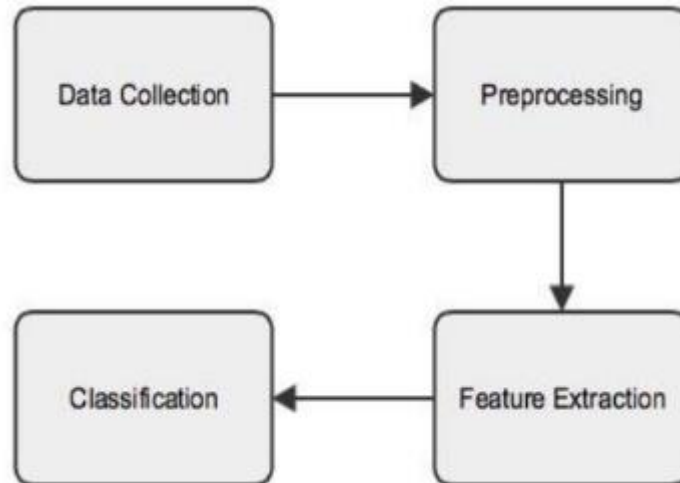
## 2.3 Pengolahan Citra

Citra, dalam konteks pengolahan citra dan visi komputer, merujuk pada representasi visual dari objek atau scene yang direkam atau dihasilkan oleh perangkat seperti kamera atau pemindai. Citra digital terdiri dari kumpulan piksel, yang merupakan elemen dasar yang membentuk gambar. Setiap piksel memiliki nilai numerik yang mewakili warna dan intensitas cahaya pada lokasi tertentu dalam gambar.

Pengolahan citra digital memerlukan berbagai metode dan algoritma untuk memproses dan menganalisis data gambar agar dapat digunakan dalam aplikasi praktis. Proses ini mencakup berbagai tahapan mulai dari peningkatan kualitas gambar hingga ekstraksi informasi yang relevan. Berbagai teknik seperti filterisasi, segmentasi, dan transformasi gambar sering digunakan untuk mengoptimalkan hasil akhir.

Secara sederhana, konsep pemrosesan citra dapat didefinisikan sebagai ranah dalam pemrosesan sinyal yang khusus untuk mengembangkan teknik komputasi guna menganalisis, meningkatkan, mengompres, mengembalikan, dan mengekstraksi informasi dari citra digital. Dengan beragam aplikasinya, pemrosesan gambar telah menjadi topik yang sangat diminati oleh komunitas ilmiah dan industri. Kombinasi minat ini, bersama dengan evolusi teknologi sistem komputer dan kebutuhan akan sistem yang memiliki kinerja yang lebih baik, baik dalam presisi, keandalan, maupun kecepatan pemrosesan, telah memungkinkan perkembangan teknik pemrosesan citra yang luar biasa. Perkembangan ini melibatkan pergeseran dari metode berbasis non-pembelajaran ke penerapan teknik pembelajaran mesin (Valente dkk, 2023).

Pengenalan pola berfungsi untuk mengenali objek dengan menggunakan algoritma yang diterapkan pada gambar guna memperoleh gambar yang lebih baik serta memahami gambar tersebut. Penelitian dan teknologi dalam bidang ini bertujuan untuk memperbaiki dan mengembangkan algoritma untuk mengumpulkan dan memproses informasi, serta mengekstraksi fitur dari gambar digital, frasa musik, atau suku kata yang mewakili sebuah kata atau bahkan teks komputer, seperti yang terlihat pada gambar 1 (Reem M. Hussein dkk, 2021).



Gambar 1 Proses pengenalan pola

## 2.4 Video Digital

Video merupakan aplikasi teknologi yang digunakan untuk menangkap, merekam, memproses, menyimpan, dan merekonstruksi sekuen gambar yang berurutan. Seperti yang dijelaskan oleh Alan C. Bovik dalam "Handbook of Image and Video Processing," video digital adalah konversi dan kuantisasi dari video analog. Pada dasarnya, proses sampling dan kuantisasi antara citra digital dan video digital tidak berbeda (Tekalp, 2015).

Kualitas video sangat dipengaruhi oleh nilai karakteristik tertentu dari video digital, yang secara langsung mempengaruhi sensitivitas visual manusia terhadap video tersebut. Beberapa karakteristik penting dari video digital meliputi:

1. Resolusi, yang merupakan ukuran dari setiap frame dalam video, biasanya diukur dalam *pixel x pixel*. Resolusi yang lebih tinggi menyediakan kualitas gambar yang lebih baik, tetapi juga mengharuskan penggunaan lebih banyak bit, sehingga meningkatkan ukuran file.
2. Kedalaman Bit, yang menunjukkan jumlah bit yang digunakan untuk merepresentasikan setiap piksel dalam frame. Tingkat kedalaman bit yang lebih tinggi membutuhkan lebih banyak bit.
3. Laju Frame, yang mengacu pada jumlah frame yang ditampilkan per detik, sering kali disebut sebagai frame per second (fps). Laju frame ini mempengaruhi seberapa mulus gerakan objek dalam video tersebut. Standar umum untuk fps adalah 30 fps dan 25 fps.

## 2.4 Face Detection

Deteksi wajah telah digunakan untuk berbagai tujuan, seperti pengawasan, identifikasi orang, lingkungan cerdas dan robotika. Dengan demikian, akurasi deteksi waktu nyata dan tinggi sangat penting faktor. Deteksi wajah adalah langkah dasar untuk identifikasi pribadi, sistem pemantauan, hukum pidana, dan interaksi manusia dan komputer. Era yang terus berkembang menuntut pengembangan yang lebih akurat teknologi. Lebih khusus lagi, banyak masalah di bidang teknologi dan keamanan kriminal yang membutuhkan identifikasi klasifikasi wajah dalam menyelesaikan masalah (Leo M. dkk, 2017).

## 2.5 Face Recognition

Pengenalan Wajah adalah bagian dari Computer Vision yang digunakan untuk mendapatkan koordinat wajah dari sebuah gambar. Banyak algoritma yang telah dikembangkan untuk mendukung pendeteksian wajah seperti *Cascade Face Detection* yang dikembangkan oleh Viola dan Jones (Yang dkk, 2014)

Beberapa faktor yang mempengaruhi akurasi ataupun presisi dari sistem *face recognition* adalah *noise, occlusion, illumination, pose variation, low resolution, aging, expressions, plastic surgery, and database*. Elemen-elemen ini dapat dikelompokkan ke dalam dua kelas: faktor alami dan faktor luar. Bagian karakteristik menggabungkan kondisi normal wajah manusia seperti tampilan, penuaan, operasi plastik, dan sebagainya, yang memengaruhi kerangka, sedangkan elemen luar mencakup hal-hal yang memengaruhi wajah dari luar, seperti kebisingan, oklusi, pencahayaan, resolusi rendah, dan variasi pose. Namun demikian, basis data juga merupakan faktor utama dalam menentukan akurasi pengenalan wajah (Meena dkk, 2022).

## 2.6 Viola Jones

Algoritma pendeteksian wajah Viola-Jones memindai gambar dengan jendela untuk mencari fitur wajah manusia. Jika fitur-fitur ini ditemukan dan memiliki nilai tertentu sebagai wajah, maka jendela tertentu dari gambar tertentu dari gambar tersebut diperkirakan sebagai wajah. Untuk menyelesaikan kasus dengan ukuran wajah yang berbeda, jendela diskalakan dengan dengan proses

yang berulang untuk setiap gambar. Mengurangi jumlah fitur yang harus diperiksa setiap jendela, jendela melewati beberapa tahap yang berbeda (Rahmad, 2020).

### 2.6.1 Integral Image

*Integral image* digunakan untuk mempercepat komputasi sum area yang efisien di dalam citra. Citra ini adalah representasi di mana setiap titik pada citra integral memiliki nilai yang merupakan total dari semua piksel di atas dan ke kiri titik ini, termasuk titik itu sendiri.



Gambar 2 Nilai pixel pada citra

Nilai pixel pada citra tersebut kemudian akan dihitung nilai dari *integral image* menggunakan persamaan berikut.

$$I(x, y) = G(x, y) + I(x - 1, y) + I(x, y - 1) - I(x - 1, y - 1) \quad (1)$$

Dimana:

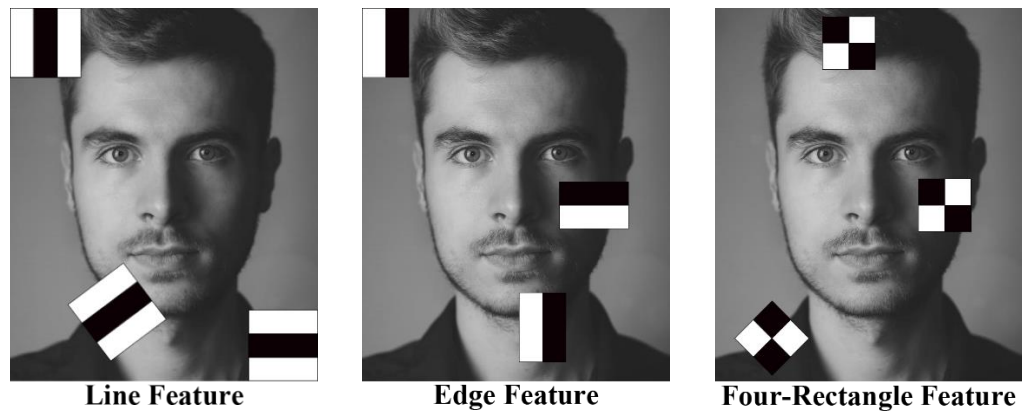
- $I(x, y)$  adalah nilai pada posisi  $(x, y)$  dalam integral image.
- $G(x, y)$  adalah intensitas piksel pada posisi  $(x, y)$  dalam gambar asli.
- $I(x - 1, y)$  adalah nilai integral image sebelumnya pada baris yang sama.
- $I(x, y - 1)$  adalah nilai integral image sebelumnya pada kolom yang sama.
- $I(x - 1, y - 1)$  adalah nilai integral image pada posisi diagonal kiri atas dari posisi saat ini.

Tujuan dari penggunaan citra integral adalah untuk memungkinkan perhitungan cepat dari fitur-fitur Haar pada langkah selanjutnya, karena dengan citra integral, area sum dapat dihitung dalam waktu konstan (Wang, 2014).

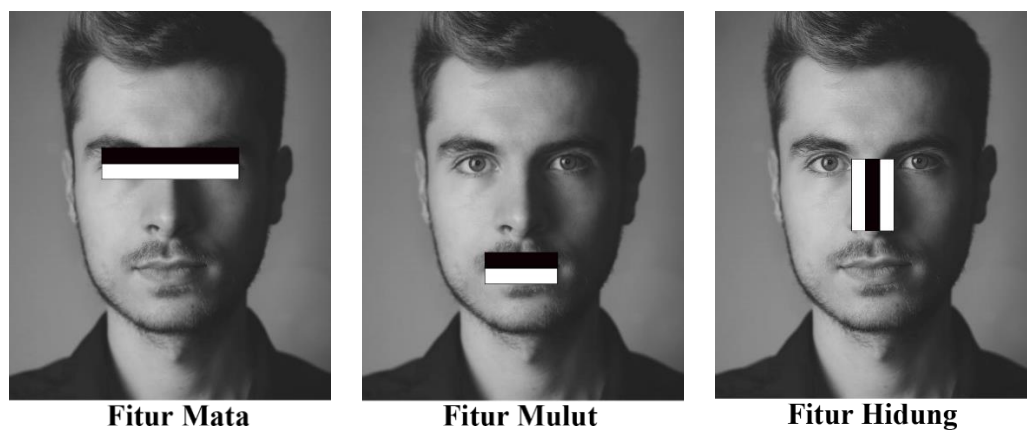


### 2.6.2 Haar-like Feature

Fitur *Haar* mengukur perbedaan intensitas piksel antara daerah terang dan gelap. Fitur ini bentuknya sederhana, umumnya berupa persegi panjang, dan disusun dalam berbagai konfigurasi seperti dua persegi vertikal atau horizontal, tiga persegi, atau empat persegi dalam pola catur.



Gambar 3 Fitur Wajah

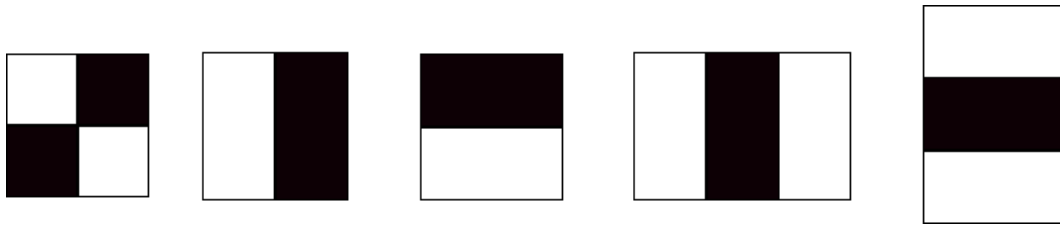


Gambar 4 Fitur mata, mulut, dan hidung

Berikut ini merupakan persamaan untuk mendapatkan nilai dari fitur *Haar* dengan:

$$Feature\ Value = \Sigma(\text{white region}) - \Sigma(\text{black region}) \quad (2)$$

Perhitungan ini efisien karena menggunakan citra integral, yang memungkinkan pengurangan cepat area piksel.

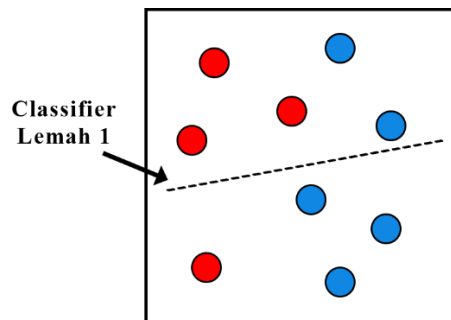


Gambar 5 Haar like feature

### 2.6.3 Adaboost

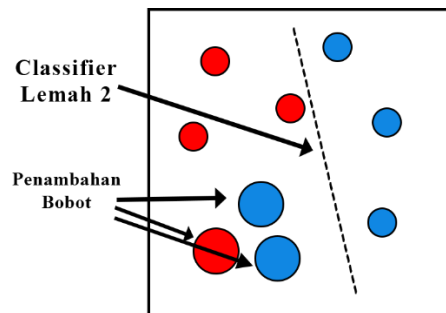
Adaboost digunakan untuk memilih fitur yang paling efektif dari ribuan atau lebih fitur Haar. Algoritma ini menggabungkan banyak pengklasifikasi lemah (berdasarkan fitur Haar individu) untuk membentuk pengklasifikasi yang kuat. Adaboost memberikan bobot kepada setiap pengklasifikasi berdasarkan keakuratannya, memungkinkan model untuk fokus pada fitur yang paling informatif.

Proses Adaboost dimulai dengan pemberian bobot yang sama kepada setiap sampel dalam dataset pelatihan. Tujuan awal adalah untuk memastikan bahwa setiap contoh memiliki peluang yang sama untuk dipertimbangkan dalam pembentukan klasifikasi lemah pertama. Klasifikasi lemah yang dimaksud umumnya merupakan fungsi sederhana yang dibuat berdasarkan satu fitur Haar tertentu, dengan tujuan meminimalkan kesalahan berbobot pada data pelatihan.

Gambar 6 *Classifier* Lemah

Dalam setiap siklus iterasi, algoritma Adaboost memilih klasifikasi lemah yang berhasil meminimalkan jumlah kesalahan berbobot terhadap sampel yang diklasifikasikan. Tingkat kesalahan ini diukur berdasarkan bobot yang telah ditentukan, dimana bobot ini akan disesuaikan setelah setiap iterasi. Klasifikasi yang membuat kesalahan akan meningkatkan bobot sampel yang salah

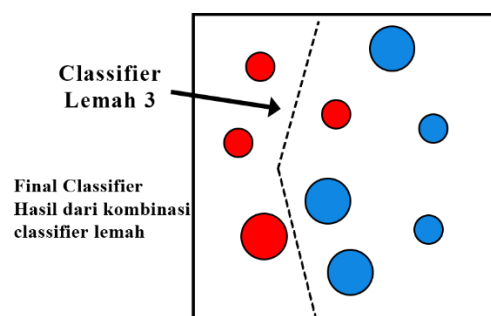
diklasifikasikan, sehingga di iterasi berikutnya, sampel tersebut mendapat perhatian lebih dalam proses pelatihan.



Gambar 7 Hasil penambahan bobot dari *Classifier* lemah

Sebagai hasil dari kesalahan klasifikasi, algoritma menghitung nilai bobot untuk setiap klasifikasi lemah, yang menunjukkan seberapa besar pengaruhnya terhadap model akhir. Nilai ini, yang disebut sebagai  $\alpha$ , dihitung berdasarkan formula yang menggambarkan proporsi kesalahan: semakin rendah tingkat kesalahan, semakin besar pengaruh klasifikasi dalam model final. Dengan demikian, klasifikasi yang paling akurat menjadi yang paling dominan dalam keputusan model gabungan.

Setelah serangkaian iterasi, semua klasifikasi lemah—yang masing-masing sekarang memiliki bobot berdasarkan efektivitasnya—dikombinasikan menjadi sebuah model klasifikasi yang kuat. Model ini kemudian dapat digunakan untuk mengklasifikasikan data baru. Dalam kasus algoritma Viola-Jones, model ini digunakan untuk memutuskan apakah suatu area dari citra mengandung wajah atau tidak (Yulina, 2021).



Gambar 8 Hasil kombinasi dari *Classifier* Lemah

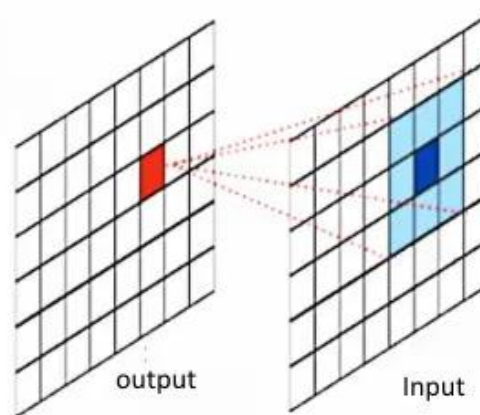
### 2.6.4 Cascade Classifier

Algoritma ini menggunakan *Cascade Classifier* untuk mengurangi area pencarian secara bertahap. Setiap tahap dari *cascade* memiliki kemampuan untuk menolak jendela yang tidak mengandung wajah dengan penghitungan yang minimal. Jendela yang lolos dari satu tahap akan diproses di tahap selanjutnya yang lebih kompleks. Ini memastikan bahwa hanya jendela yang paling mungkin mengandung wajah yang dianalisis lebih lanjut, meningkatkan efisiensi deteksi.

Untuk mendeteksi wajah dengan berbagai ukuran, algoritma menerapkan *Cascade Classifier* pada berbagai skala. Ini dilakukan dengan mengubah ukuran jendela deteksi dan menggesernya di seluruh citra, memungkinkan algoritma untuk mendeteksi wajah pada berbagai ukuran dan posisi (Rani & Chakkaravarthy, 2022).

### 2.7 Local Binary Pattern

LBP menyajikan konsep yang sederhana namun kuat. Algoritma ini bekerja pada tingkat piksel, di mana setiap piksel dalam citra diubah menjadi pola biner berdasarkan perbandingan intensitasnya dengan piksel tetangga. Pendekatan ini memberikan representasi pola tekstur yang khas untuk setiap piksel, menciptakan citra LBP yang mencerminkan tekstur lokal di sekitar setiap titik citra. Ilustrasi cara kerja dari algoritma *Local Binary Pattern* dapat dilihat pada gambar 9.



Gambar 9 Cara kerja metode *Local Binary Pattern*

Seperti namanya, Local Binary Pattern (disingkat LBP) adalah fitur representasi lokal dari suatu citra. Fitur ini terdiri atas nilai relatif dengan

membandingkan setiap piksel dengan piksel tetangganya (Harmouch, 2020). Langkah-langkah metode LBP adalah sebagai berikut.

- Konversikan citra ke dalam ruang skala abu-abu.
- Untuk setiap piksel ( $gp$ ) pada citra, pilih  $P$  neighborhoods yang mengelilingi piksel pusat. koordinat  $gp$  diberikan oleh

$$(gc_x - R_{\sin} \frac{2\pi p}{p}, gc_y + R_{\cos} \frac{2\pi p}{p}) \quad (3)$$

- Ambil piksel tengah ( $gc$ ) dan tetapkan sebagai ambang batas untuk  $P$  neighbors.
- Tetapkan ke 1 jika nilai piksel yang berdekatan lebih besar atau sama dengan nilai piksel tengah, 0 jika tidak.
- Sekarang, hitung nilai LBP: Secara berurutan berlawanan arah jarum jam, tuliskan angka biner yang terdiri atas angka-angka yang berdekatan dengan piksel tengah. Angka biner ini (atau ekuivalen desimalnya) disebut kode piksel tengah LBP dan, selanjutnya, digunakan sebagai karakteristik tekstur lokal yang dipilih.

$$LBP(gp_x, gp_y) \sum_{p=0}^{P-1} S(gp - gc) \times 2^p \quad (4)$$

Yang dimana:

- $gc$  merupakan nilai intensitas *pixel* pusat
- $gp$  merupakan nilai intensitas *pixel* tetangga (*neighbors*) dengan indeks  $p$

fungsi  $S$  dapat dinyatakan pada persamaan 5 berikut.

$$S(x) = \begin{cases} 1 & \text{jika } x \geq 0 \\ 0 & \text{jika } x < 0 \end{cases} \quad (5)$$

Nilai keabuan dari tetangga yang tidak berada tepat di tengah-tengah piksel (blok) diestimasi dengan interpolasi (Harmouch, 2020).

### 2.7.1 *Local Binary Pattern Histogram*

*Local Binary Pattern Histogram* (LBPH) adalah sebuah karakteristik untuk klasifikasi yang menggabungkan teknik histogram dan merupakan inovasi terbaru dari pendekatan *Local Binary Pattern* (LBP) untuk meningkatkan kinerja dalam pengenalan wajah. Secara umum, LBP dirancang untuk mengenali pola tekstur dalam citra. Metode LBPH menjadi pilihan yang optimal untuk implementasi pengenalan wajah pada perangkat bergerak Android karena mengadopsi perhitungan yang simpel, sesuai dengan keterbatasan sumber daya pada platform Android (T. Filareti, 2011).

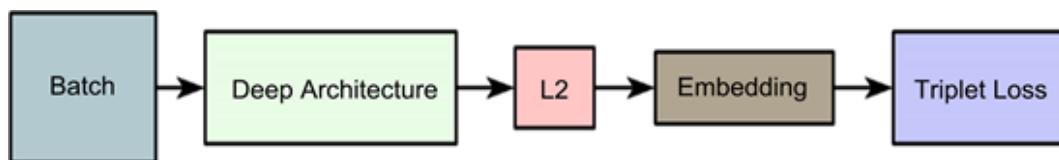
### 2.7.2 **Parameter LBP**

Algoritma *Local Binary Pattern* (LBP) memiliki beberapa parameter yang digunakan untuk menyesuaikan dan mengonfigurasi operasinya. Empat parameter utama pada algoritma LBP adalah:

- Radius ( $R$ ), parameter ini menentukan jarak dari piksel pusat ke piksel tetangga yang akan dianalisis. Radius ini mengukur seberapa jauh piksel tetangga diambil dalam perhitungan pola biner lokal. Nilai radius yang berbeda dapat menghasilkan hasil yang berbeda dan memengaruhi ukuran area lokal yang dianalisis.
- *Neighbors* ( $P$ ), menentukan jumlah tetangga yang akan diambil dalam perhitungan pola biner lokal. Jumlah tetangga ini menentukan seberapa banyak perbandingan intensitas dilakukan untuk membentuk pola biner pada setiap piksel.
- Grid X, mengindikasikan jumlah sel dalam arah mendatar. Semakin banyak sel yang digunakan, semakin halus gridnya, dan semakin besar dimensi vektor yang dihasilkan.
- Grid Y, pada sisi lain, mencerminkan jumlah sel dalam arah vertikal. Semakin banyak sel yang digunakan, semakin halus gridnya, dan semakin besar dimensi dari vektor yang dihasilkan.

## 2.8 Facenet

*FaceNet* adalah sistem Pengenalan Wajah yang dikembangkan pada tahun 2015 yang mencapai hasil yang canggih pada dataset tolok ukur pengenalan wajah. Sistem *FaceNet* dapat digunakan untuk mengekstrak fitur berkualitas tinggi dari wajah yang disediakan dalam gambar, yang disebut penyematan wajah. Penyematan wajah kemudian digunakan untuk melatih sistem identifikasi wajah. Selain itu, *FaceNet* secara langsung mempelajari pemetaan dari gambar wajah ke ruang Euclidian yang ringkas, di mana jarak secara langsung sesuai dengan ukuran kemiripan wajah. Perbedaan antara *FaceNet* dan metode lainnya adalah bahwa *FaceNet* mempelajari peta ping dari gambar atau wajah dan membuat *embedding* daripada menggunakan botol apa pun lapisan leher untuk tugas pengenalan atau verifikasi. Setelah penyematan dibuat, semua tugas-tugas lain seperti verifikasi dan pengenalan dapat dilakukan dengan menggunakan standar teknik standar. *FaceNet* menggunakan *triplet loss*, yang secara langsung akan mencerminkan apa yang kita inginkan yang ingin kita capai dalam verifikasi, - pengenalan, dan pengelompokan wajah. Alur kerja dari *facenet* dapat dilihat pada gambar 10.



Gambar 10 Facenet Arsitektur

Dalam ruang ini, panjang jarak spasial berbanding lurus dengan tingkat kemiripan antar wajah. Setelah ruang tersebut terbentuk, teknik standar dengan penyematan *FaceNet* sebagai vektor fitur dapat dengan mudah digunakan untuk melakukan tugas-tugas seperti pengenalan wajah, verifikasi, dan pengelompokan. Kelebihan model ini adalah efisiensi dalam mengolah gambar yang digunakan sebagai input. Selain itu, model ini memiliki akurasi yang sangat tinggi pada berbagai kumpulan data dan aplikatif untuk pengenalan wajah pada telepon seluler.

Struktur jaringannya ditunjukkan pada Gambar 12.

Jaringan *FaceNet* terdiri dari lapisan input *batch*, jaringan konvolusi yang mendalam, dilanjutkan dengan normalisasi L2 yang menghasilkan penyematan

wajah, dan akhirnya, penghitungan *triplet loss* yang meminimalisir jarak antara objek serupa dan memaksimalkan jarak antara objek yang berbeda. Model ini menggunakan jaringan saraf konvolusional yang mendalam untuk mempelajari metode penyematan Euclidean dari setiap gambar, mengoptimalkan jaringan sehingga jarak L2 kuadrat dalam ruang penyematan sesuai dengan kemiripan wajah. FaceNet menggunakan fungsi *Loss* dari LMNN berbasis *Triples* (Klasifikasi Tetangga Terdekat Batas Maksimum) untuk melatih jaringan, dengan output dari jaringan adalah ruang vektor 128 dimensi. *Triplet* yang dipilih mengandung dua gambar mini wajah yang serupa dan satu gambar mini wajah yang berbeda, dengan tujuan fungsi *Loss* untuk membedakan antara kelas positif dan negatif berdasarkan batas jarak.

## 2.9 Dlib

Dlib merupakan sebuah *library machine learning* yang ditulis dalam bahasa pemrograman C++ dan ditunjukkan untuk menyelesaikan masalah kehidupan sehari-hari. Dlib adalah *library open source* yang biasa dipakai untuk pengembangan software machine learning. Inti dari dlib adalah sebuah aljabar linier dengan *Basic Linear Algebra Subprograms* (BLAS). BLAS biasa digunakan untuk implementasi bayesian network dan algoritma kernel-based untuk klasifikasi, *clustering*, deteksi anomali, regresi dan *feature ranking*. Dlib mempunyai dua komponen penting yaitu aljabar linier dan *machine learning tools*.

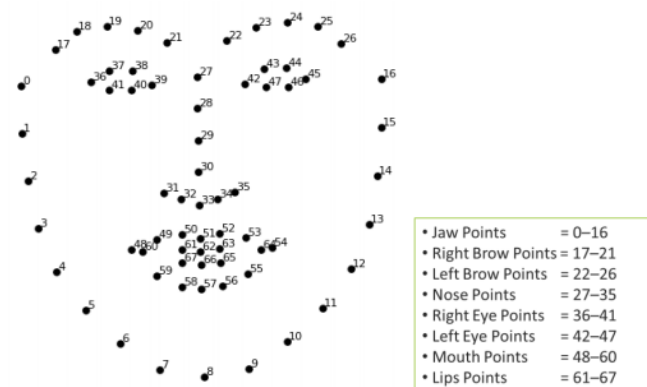
Komponen aljabar linier berbasis pada teknik template expression yang ada pada Blitz++. Blitz++ merupakan sebuah *numerical software* yang dikembangkan oleh Veldhuizen dan Ponnambalam (1996). Dlib menggunakan BLAS agar mendapatkan performa terbaik dan menambah kecepatan sebagai sebuah library. Dlib dapat menunjukkan semua transformasi pada semua ekspresi dengan melibatkan BLAS yang dapat digunakan. Lalu BLAS dapat membuat pengguna melakukan penulisan atau penjumlahan dalam bentuk intuisi dan diberikan pada *library*.

Tujuan utama pada *machine learning tools* adalah untuk menyajikan sebuah arsitektur modular dan mudah digunakan. Dlib dapat diimplementasikan pada



gambar, kolom vektor atau struktur bentuk data. Implementasi dari algoritma akan berbeda pada data dengan apa yang akan dioperasikan nantinya. Karena *dlib* sangat fleksibel maka dapat dioperasikan secara langsung pada objek apapun dimana itu membuat kernel menjadi custom dimana biasanya kernel hanya bisa pada objek yang sudah jelas ukurannya. *Facial Landmark* digunakan oleh *Dlib* untuk mendeteksi wajah manusia. *Dlib* menggunakan *training* set yang sudah ditandai (Wahyudiana, 2019).

*Facial landmark* merupakan teknik yang digunakan untuk mengidentifikasi dan melokalisasi titik-titik penting pada wajah, yang sering disebut sebagai titik jangkar. Teknik ini sangat penting untuk menganalisis dan menentukan struktur biologis wajah manusia. Dalam proses ekstraksi titik-titik ini, dibutuhkan sebuah alat atau metode, salah satunya adalah *Dlib Regression Tree*. Metode ini efektif dalam melokalisasi 68 titik penting yang telah dilatih sebelumnya seperti pada gambar 11. Teknik *landmark* wajah ini tidak hanya mampu mengalokasikan titik-titik pada wajah untuk pengenalan, tetapi juga dapat mendeteksi bagian-bagian wajah seperti mata, alis, mulut, dan hidung, yang mendukung berbagai aplikasi seperti pengenalan wajah, rekonstruksi wajah, dan analisis atribut wajah seperti usia, jenis kelamin, ekspresi senyum, serta penggunaan kacamata. Teknologi ini memiliki aplikasi luas dalam berbagai bidang termasuk keamanan, interaksi manusia dan komputer, serta kesehatan untuk menganalisis kondisi medis atau psikologis melalui ekspresi wajah.



Gambar 11 *Facial landmarks*

## 2.10 Precision, Recall, dan F1-Score

Gambaran yang lengkap tentang kinerja suatu sistem bisa dilihat dari kurva *precision-recall*, yang sering dirangkum dalam satu indikator dengan rata-rata precision pada berbagai tingkat recall standar atau jumlah dokumen. *Precision* (Presisi) adalah ukuran akurasi dari prediksi positif yang dibuat oleh model. *Precision* menunjukkan berapa banyak dari semua prediksi positif yang benar-benar relevan. Dalam konteks deteksi wajah atau klasifikasi lainnya, precision dapat didefinisikan sebagai rasio antara jumlah true positive (TP) dengan jumlah keseluruhan prediksi positif (true positive + false positive). Perhitungan presisi dapat dilihat pada persamaan 6 berikut (Goutte dan Gaussier, 2005).

$$Presisi = \frac{TP}{TP+FP} \quad (6)$$

*Recall* adalah ukuran sensitivitas dari model. *Recall* menunjukkan berapa banyak dari semua kasus positif yang sebenarnya berhasil terdeteksi oleh model. Dalam konteks deteksi wajah atau klasifikasi lainnya, recall dapat didefinisikan sebagai rasio antara jumlah *true positive* (TP) dengan jumlah keseluruhan kasus positif sebenarnya (true positive + false negative).

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

*F1 Score* adalah rata-rata harmonis dari precision dan recall. *F1 Score* memberikan keseimbangan antara *precision* dan *recall* dan berguna ketika ada ketidakseimbangan antara jumlah kasus positif dan negatif. *F1-Score* memberikan gambaran yang lebih baik tentang kinerja model secara keseluruhan dalam situasi di mana keseimbangan antara precision dan recall penting.

$$F1 - Score = 2 \times \left( \frac{Presisi \times Recall}{Presisi + Recall} \right) \quad (8)$$

## 2.11 Python

Python merupakan salah satu bahasa pemrograman yang banyak digunakan oleh perusahaan besar maupun para developer untuk mengembangkan berbagai

macam aplikasi berbasis desktop, web dan *mobile* (Romzi dkk., 2020). Python adalah bahasa pemrograman *open-source*, yang artinya dapat dikembangkan oleh siapapun. Dikenal sebagai bahasa tingkat tinggi, Python memudahkan pemrograman berorientasi objek melalui pendekatannya yang sederhana. Python sangat relevan bagi ilmuwan data dan sering digunakan dalam berbagai aplikasi ilmu data. Bahasa ini juga menawarkan fungsionalitas yang sangat baik untuk mengelola matematika, statistik, dan fungsi ilmiah (Saabith dkk., 2019).

## **2.12 OpenCV**

Open Computer Vision (OpenCV) adalah perpustakaan sumber terbuka yang difokuskan pada pengolahan citra. Fokus utamanya adalah memberikan kemampuan pada komputer untuk melakukan pengolahan visual dengan cara yang serupa dengan manusia. OpenCV telah menghadirkan berbagai algoritma dasar dalam bidang visi komputer. Selain itu, OpenCV juga menyediakan modul deteksi objek yang menggunakan metode visi komputer (Muchtar dan Apriadi, 2019).