BAB II
TINJAUAN PUSTAKA

2.1 Penelitian Sebelumnya

Pada beberapa penelitian terkait sebelumnya telah memberikan kemudahan dan digunakan sebagai referensi untuk memperkaya bahan penelitian. Berikut adalah penelitian-penelitian yang memiliki kesinambungan dengan penelitian yang akan dilakukan dan menjadi bahan perbandingan dan pertimbangan penulis.

Penelitian pertama dari Irianto, dkk. pada tahun 2019 mengenai sistem keamanan pintu. Teknologi keamanan pada pintu yang biasa digunakan diantaranya yaitu menggunakan kunci, *id card* dan *password*. Cara tersebut dianggap masih memiliki kerentanan dalam keamanannya seperti kehilangan hingga lupa dengan *password*. Sehingga diperlukan keamanan menggunakan teknologi biometrik berupa *face recognition* untuk proses verifikasinya. Penelitian ini menggunakan metode *Haar Cascade Classifier* untuk melakukan deteksi wajah dan untuk proses pengenalan wajah menggunakan metode *Local Binary Pattern Histogram* (LBPH). Menggunakan *dataset* tiga macam citra wajah dengan masing-masing sebanyak 100 gambar. Pengujian dilakukan terhadap 5 orang (3 terdaftar dan 2 tidak terdaftar) dengan masing-masing dilakukan sebanyak lima kali. Hasilnya menunjukkan nilai *confident* lebih dari 70% untuk pengguna terdaftar dan kurang dari 67% untuk pengguna tidak terdaftar [17].

Penelitian kedua dilakukan oleh Santoso dan Kristianto pada tahun 2020 mengenai sistem presensi mahasiswa. Sistem presensi pada perguruan tinggi masih banyak dijumpai menggunakan presensi secara manual dan menggunakan *barcode* pada Kartu Tanda Mahasiswa (KTM). Kedua metode tersebut dinilai masih memiliki kelemahan, seperti fenomena titip absen pada presensi manual dan risiko kehilangan KTM. Diperlukan solusi inovasi baru pada teknologi presensi yang dapat mencegah kecurangan mahasiswa yaitu

menggunakan teknologi *face recognition* yang merupakan teknologi biometrik. Penelitian dilakukan dengan menggunakan metode *Haar Cascade Classifier* dan LBPH. *Dataset* wajah mahasiswa diambil sebanyak 30 kali dengan sudut dan pose berbeda. Pengujian dilakukan dan hasilnya sistem mampu mendeteksi dan mengenali beberapa wajah dalam satu *frame* dengan jarak maksimal 150 cm. Jika wajah terhalang benda atau jarak lebih dari 150 cm sistem tidak dapat mendeteksi dan mengenali wajah [18].

Penelitian ketiga yaitu mengenai sistem presensi yang dilakukan oleh Buana pada tahun 2021. Pada masa pandemi COVID-19 dilakukan pembatasan aktivitas seperti menjaga jarak. Hal ini berdampak pada sistem presensi yang mengharuskan mahasiswa untuk tidak menyentuh tempat publik dan menjaga jarak. Sehingga diperlukan solusi dengan menerapkan teknologi *face recognition* yang mana cara kerjanya tidak perlu menyentuh apapun untuk memverifikasi presensi. Metode Viola-Jones (*Haar Cascade*) digunakan pada penelitian ini untuk melakukan deteksi wajah dan metode LBPH untuk mengenali wajah. *Dataset* yang digunakan sebanyak 100 gambar wajah. Jarak maksimal adalah 100 cm dengan jarak minimal 40 cm. Dilakukan pengujian dan diperoleh hasil untuk jarak kurang dari 40 cm dan lebih dari 100 cm wajah tidak dapat dikenali. Kemudian pada sudut kurang maupun lebih 30^{o} ke kanan maupun ke atas juga tidak dapat dikenali [16].

Penelitian keempat pada tahun 2021 oleh Rehman, dkk. mengenai pelatihan robotika *Artificial Intelligence* (AI) menggunakan mini komputer untuk melakukan deteksi wajah (*facial detection*). Mini komputer yang digunakan adalah NVIDIA Jetson Nano. Dipilih karena sudah terintegrasi dengan GPU dan CPU dengan performansi yang sangat baik. Metode deteksi wajah yang digunakan dalam penelitian ini adalah SSD (*Single Shot Detector*). Menggunakan *dataset* sebanyak 139 citra wajah dari 3 orang dengan 29 untuk validasi dan 110 untuk pelatihan. Hasilnya diperoleh akurasi model mencapai 97%, sangat efektif dalam mendeteksi wajah [19].

Penelitian kelima oleh Dang pada tahun 2023 mengenai sistem presensi cerdas. Hampir kebanyakan sistem identifikasi menggunakan metode

konvensional seperti *id card*, *password* dan *fingerprint*. Namun kehilangan *id card* hingga lupa dengan *password* kerap menjadi masalah umum. Kemudian hadir teknologi identifikasi menggunakan QR Code dan RFID. Akan tetapi keduanya masih memiliki masalah dimana QR Code tidak dapat dipindai menggunakan *barcode scanner* dan memerlukan *scanner* khusus QR Code, sementara RFID kerap bermasalah dengan frekuensinya yang dapat terhalang sehingga tidak dapat berfungsi dengan baik. Sehingga diperlukan solusi menggunakan teknologi biometrik pengenalan wajah. Penelitian ini menggunakan model FaceNet dengan MobileNetV2 sebagai *backbone* atau struktur dasarnya untuk melakukan pengenalan wajah. Kemudian metode SSD digunakan untuk proses deteksi wajah. *Dataset* yang digunakan sebanyak 13.000 citra wajah dari 50 orang. Dilakukan pengujian dan hasilnya diperoleh akurasi mencapai 99% dengan kecepatan proses berkisar 20-23 FPS, menunjukkan sistem yang sangat efektif dan efisien [20].

Penelitian keenam menggunakan metode *Haar Cascade Classifier* dan LBPH yang dilakukan oleh Taib, dkk. pada tahun 2021. Pada sistem keamanan dan pengawasan gedung masih banyak dilakukan secara manual. Hal ini menyebabkan orang yang tidak berkepentingan dapat dengan mudah masuk, menunjukkan sistem keamanan yang rentan dan belum efektif. Berdasarkan hal ini diperlukan suatu teknologi untuk memberikan keamanan ekstra yaitu dengan menggunakan pengenalan wajah (*face recognition*). Metode *Haar Cascade* digunakan untuk mendeteksi wajah dan metode LBPH digunakan untuk klasifikasi wajah. Penelitian ini menggunakan dua data wajah dengan dilakukan pengujian sebanyak lima kali untuk setiap wajah. Parameter pengujiannya adalah jarak dekat dan jauh, pencahayaan terang dan gelap, dan posisi wajah menghadap atas, depan dan bawah. Hasilnya diperoleh untuk parameter jarak mencapai 70% (dekat) dan 73% (jauh), parameter pencahayaan mencapai 36% (terang) dan 57% (gelap), dan parameter posisi wajah mencapai 57% (atas), 71% (depan) dan 69% (bawah) [13].

Penelitian ketujuh oleh Sukusvieri (2020) menggunakan metode SSD untuk pengenalan wajah. Penggunaan teknologi terbaru pada fasilitas umum

banyak digunakan seperti kamera pengawas berbasis *IoT*, alat pendeteksi logam hingga sistem identifikasi manusia. Salah satu pengimplementasian sistem identifikasi manusia adalah pengenalan wajah, yang menggunakan ciri wajah sebagai fokus identifikasinya. Berdasarkan hal tersebut, penelitian dilakukan dengan membuat sistem pengenalan wajah menggunakan metode SSD. Metode SSD digunakan untuk proses deteksi dan juga untuk proses pengenalannya. Penelitian menggunakan 5 orang responden (1 orang terdaftar/dikenal dan 4 tidak terdaftar/dikenal) dan dilakukan pengujian menggunakan parameter sudut wajah (depan, atas, bawah, kanan, kiri). Hasilnya dari kelima responden diperoleh rata-rata akurasi deteksi wajah sebesar 100%, sementara untuk akurasi pengenalan wajah sebesar 88% dan presisinya 63% [8].

Penelitian kedelapan melakukan perbandingan antar metode deteksi wajah oleh Farokhah pada tahun 2021. Di dalam pengenalan wajah, proses deteksi wajah sangat berperan penting. Perkembangan teknologi deteksi wajah sangat pesat, beberapa diantaranya yang cukup terkenal yaitu metode *Haar Cascade Classifier (Viola Jones)* yang merupakan metode paling tua yang paling banyak digunakan dan masih relevan hingga saat ini. Kemudian metode Dlib CNN yang diklaim cukup akurat dan mampu menghasilkan deteksi wajah yang paling baik. Metode lainnya adalah SSD, yang dikembangkan untuk mengatasi kekurangan *Haar Cascade*. SSD sudah diterapkan dalam mendeteksi wajah yang terhalang *obstacle* seperti kacamata, masker, dsb. dan hasilnya menunjukkan performansi yang baik. Penelitian ini bertujuan untuk membandingkan performansi ketiga metode deteksi wajah dengan pendekatan *wild condition* atau kondisi liar yang lebih realistis dengan keadaan nyata. Data uji diambil secara acak dari Google dengan lima kondisi yaitu variasi pose dan sudut pandang, halangan benda pada wajah, pencahayaan, gangguan *background* dan perubahan ekspresi wajah. Hasilnya diperoleh untuk metode *Haar Cascade* dari 5 pengujian hanya 1 yang sesuai sehingga memiliki performansi sebesar 20% dibandingkan Dlib dan SSD. Hasil performansi Dlib dan SSD memiliki persentase yang sama yaitu 80% dimana dari 5 pengujian

hanya 1 yang tidak sesuai. Metode Dlib dan SSD memiliki performansi yang sangat baik dalam mendeteksi wajah [15].

Penelitian kesembilan oleh Detila, dkk. pada tahun 2019 yang membandingkan metode *Eigenface*, *Fisherface* dan LBPH. Teknologi dalam penerapan sistem pengenalan wajah sudah banyak diciptakan, beberapa diantaranya adalah metode *Eigenface*, *Fisherface* dan LBPH. *Eigenface* adalah metode identifikasi wajah yang memanfaatkan algoritma *Principal Component Analysis* (PCA) untuk mengurangi dimensi citra dari dimensi aslinya menjadi dimensi fitur. PCA merupakan teknik statistik multivarian yang banyak digunakan. Berikutnya metode *Fisherface*, merupakan turunan dari *Fisher's Linear Discriminant* (FLD). Dikembangkan untuk mengatasi kekurangan *Eigenface*. Prinsip dasar dari metode *Fisherface* adalah melakukan reduksi dimensi yang pada saat yang sama meningkatkan rasio jarak antara kelas (*between-class scatter*) dan jarak intra kelas (*within-class scatter*). Dan metode LBPH, yang dasarnya dari metode *Local Binary Patterns* (LBP), merupakan jenis *visual descriptor* yang digunakan dalam bidang visi komputer untuk klasifikasi. Operator LBP merupakan metode unggulan dalam pengenalan tekstur. Untuk meningkatkan kinerja pengenalan wajah, LBP telah dimodifikasi dengan penambahan histogram, menjadi metode yang dikenal sebagai *Local Binary Patterns Histograms* (LBPH). Berdasarkan hal ini, penelitian dilakukan untuk membandingkan kinerja dari setiap metode. Penelitian ini menggunakan *dataset* dengan 6 wajah untuk dilakukan pengenalan wajah dan dilatih menggunakan ketiga metode. Pengujian dilakukan dengan 4 parameter: tingkat kecerahan rendah, kecerahan normal, wajah ekspresi datar, dan wajah ekspresi senyum. Hasilnya diperoleh rata-rata akurasi *Eigenface* sebesar 46%, *Fisherface* sebesar 54% dan LBPH sebesar 83%, menunjukkan bahwa metode LBPH memiliki akurasi yang paling baik [14].

Tabel 2.1 Penelitian Sebelumnya

No.	Penulis (Tahun)	Masalah	Metode	*Dataset*	Hasil
1	R. Irianto, S. Prabowo, dan R. Yasirandi (2019) [17]	Teknologi keamanan pada pintu yang biasa digunakan diantaranya yaitu menggunakan kunci, *id card* dan *password*. Cara tersebut dianggap masih memiliki kerentanan dalam keamanannya seperti kehilangan hingga lupa dengan *password*.	*Haar Cascade Classifier* dan *Local Binary Pattern Histogram* (LBPH)	Menggunakan *dataset* tiga macam citra wajah dengan masing-masing sebanyak 100 gambar.	Hasilnya menunjukkan nilai *confident* lebih dari 70% untuk pengguna terdaftar dan kurang dari 67% untuk pengguna tidak terdaftar.
2	B. Santoso dan R. P. Kristianto (2020) [18]	Sistem presensi pada perguruan tinggi masih banyak dijumpai menggunakan presensi secara manual dan menggunakan *barcode* pada Kartu Tanda Mahasiswa (KTM). Kedua metode tersebut dinilai masih memiliki kelemahan, seperti fenomena titip absen pada presensi manual dan risiko kehilangan KTM.	*Haar Cascade Classifier* dan LBPH	*Dataset* wajah mahasiswa diambil sebanyak 30 kali dengan sudut dan pose berbeda.	Hasilnya sistem mampu mendeteksi dan mengenali beberapa wajah dalam satu *frame* dengan jarak maksimal 150 cm. Jika wajah terhalang benda atau jarak lebih dari 150 cm sistem tidak dapat mendeteksi dan mengenali wajah.

No.	Penulis (Tahun)	Masalah	Metode	*Dataset*	Hasil
3	I. K. S. Buana (2021) [16]	Pada masa pandemi COVID-19 dilakukan pembatasan aktivitas seperti menjaga jarak. Hal ini berdampak pada sistem presensi yang mengharuskan mahasiswa untuk tidak menyentuh tempat publik dan menjaga jarak.	Viola-Jones (*Haar Cascade*) dan LBPH	*Dataset* yang digunakan sebanyak 100 gambar wajah. Jarak maksimal adalah 100 cm dengan jarak minimal 40 cm.	Diperoleh hasil untuk jarak kurang dari 40 cm dan lebih dari 100 cm wajah tidak dapat dikenali. Kemudian pada sudut kurang maupun lebih 30^{o} ke kanan maupun ke atas juga tidak dapat dikenali.
4	S. U. Rehman, M. R. Razzaq, dan M. H. Hussian (2021) [19]	Pelatihan robotika *Artificial Intelligence* (AI) menggunakan mini komputer untuk melakukan deteksi wajah (*facial detection*). Mini komputer yang digunakan adalah NVIDIA Jetson Nano. Dipilih karena sudah terintegrasi dengan GPU dan CPU dengan performansi yang sangat baik.	SSD (*Single Shot Detector*)	Menggunakan *dataset* sebanyak 139 citra wajah dari 3 orang dengan 29 untuk validasi dan 110 untuk pelatihan.	Hasilnya diperoleh akurasi model mencapai 97%, sangat efektif dalam mendeteksi wajah.

No.	Penulis (Tahun)	Masalah	Metode	*Dataset*	Hasil
5	T. V. Dang (2023) [20]	Sebagian besar sistem identifikasi saat ini mengandalkan metode konvensional seperti kartu identitas, kata sandi, dan sidik jari. Namun, kerugian kartu identitas dan lupa kata sandi seringkali menimbulkan masalah umum. Meskipun teknologi identifikasi baru seperti QR Code dan RFID telah muncul, keduanya masih memiliki kendala. QR Code sulit dipindai menggunakan *scanner barcode* dan memerlukan perangkat khusus, sementara RFID seringkali terpengaruh oleh gangguan frekuensi yang menghambat kinerjanya.	Model FaceNet dengan MobileNetV2 dan SSD	*Dataset* yang digunakan sebanyak 13.000 citra wajah dari 50 orang.	Hasilnya diperoleh akurasi mencapai 99% dengan kecepatan proses berkisar 20-23 FPS, menunjukkan sistem yang sangat efektif dan efisien.
6	Taib S, Sudin S, Muhammad A EJD (2021) [13]	Sistem keamanan dan pengawasan gedung saat ini sering kali masih dilakukan secara manual, menyebabkan	*Haar Cascade Classifier* dan LBPH	Menggunakan dua data wajah dengan dilakukan pengujian sebanyak	Hasilnya diperoleh untuk parameter jarak mencapai 70% (dekat) dan 73% (jauh), parameter

No.	Penulis (Tahun)	Masalah	Metode	*Dataset*	Hasil
		kerentanan terhadap akses yang tidak diinginkan. Oleh karena itu, dibutuhkan teknologi pengenalan wajah untuk meningkatkan keamanan secara efektif.		lima kali untuk setiap wajah	pencahayaan mencapai 36% (terang) dan 57% (gelap), dan parameter posisi wajah mencapai 57% (atas), 71% (depan) dan 69% (bawah)
7	Sukusvieri A (2020) [8]	Penggunaan teknologi terbaru pada fasilitas umum banyak digunakan seperti kamera pengawas berbasis *IoT*, alat pendeteksi logam hingga sistem identifikasi manusia. Salah satu pengimplementasian sistem identifikasi manusia adalah pengenalan wajah, yang menggunakan ciri wajah sebagai fokus identifikasinya.	SSD	Penelitian menggunakan 5 orang responden (1 orang terdaftar/dikenal dan 4 tidak terdaftar/dikenal)	Hasilnya dari kelima responden diperoleh rata-rata akurasi deteksi wajah sebesar 100%, sementara untuk akurasi pengenalan wajah sebesar 88% dan presisinya 63%
8	Lia Farokhah (2021) [15]	Di dalam pengenalan wajah, proses deteksi wajah sangat berperan penting. Perkembangan teknologi deteksi wajah sangat pesat,	*Haar Cascade Classifier (Viola Jones)*, Dlib CNN dan SSD	Data uji diambil secara acak dari Google dengan lima kondisi yaitu variasi pose dan sudut pandang,	Hasilnya diperoleh untuk metode *Haar Cascade* dari 5 pengujian hanya 1 yang sesuai sehingga memiliki performansi sebesar 20%

No.	Penulis (Tahun)	Masalah	Metode	*Dataset*	Hasil
		beberapa diantaranya yang cukup terkenal yaitu metode *Haar Cascade Classifier (Viola Jones)*, Dlib CNN dan SSD. Diperlukan pengujian ketiganya untuk mengetahui performansi dari masing-masing metode dalam mendeteksi wajah.		halangan benda pada wajah, pencahayaan, gangguan *background* dan perubahan ekspresi wajah.	dibandingkan Dlib dan SSD. Hasil performansi Dlib dan SSD memiliki persentase yang sama yaitu 80% dimana dari 5 pengujian hanya 1 yang tidak sesuai. Metode Dlib dan SSD memiliki performansi yang sangat baik dalam mendeteksi wajah.
9	Detila Q, Eri D, Wibowo P (2019) [14]	Teknologi dalam penerapan sistem pengenalan wajah sudah banyak diciptakan, beberapa diantaranya adalah metode Eigenface, Fisherface dan LBPH. Ketiganya memiliki karakteristiknya masing-masing, untuk mengetahui metode mana yang paling baik, diperlukan pengujian untuk membandingkan mana yang paling baik dari ketiganya.	*Eigenface*, *Fisherface* dan LBPH	Menggunakan *dataset* dengan 6 wajah	Hasilnya diperoleh rata-rata akurasi *Eigenface* sebesar 46%, *Fisherface* sebesar 54% dan LBPH sebesar 83%, menunjukkan bahwa metode LBPH memiliki akurasi yang paling baik

Penelitian-penelitian di atas dijadikan sebagai referensi dalam penyusunan penelitian ini. Penelitian di atas menunjukkan bahwa penggunaan metode SSD dan LBPH cukup baik di dalam melakukan deteksi dan pengenalan wajah. Pada penelitian yang dijadikan sebagai referensi utama masih menggunakan *dataset* wajah yang terbatas dan untuk penerapannya hanya sebatas sistem pengenalan wajah saja, sehingga penelitian ini akan memperbaiki penelitian sebelumnya dengan menambah jumlah *dataset* dan wajah yang bervariasi. Kemudian sistem pengenalan wajah akan diimplementasikan dalam bentuk sistem presensi berbasis *face recognition*.

2.2 Landasan Teori

2.2.1 Presensi

Presensi adalah aktivitas perekaman kehadiran seseorang dalam suatu acara atau kegiatan, misalnya belajar mengajar, seminar, pertemuan dan lain sebagainya. Pada lingkup pendidikan tinggi dapat diartikan sebagai pencatatan kehadiran mahasiswa dalam bentuk berkas atau dokumen. Dokumen ini dapat berupa daftar hadir manual atau daftar hadir digital yang otomatis tercatat oleh sistem pencatat kehadiran. Data kehadiran ini nantinya dapat dipergunakan sebagai bahan evaluasi akademik mahasiswa [12].

2.2.2 Biometri

Sistem biometri merupakan sebuah pendekatan komputerisasi yang memanfaatkan karakteristik biologis, terutama fitur-fitur yang unik dan spesifik yang dimiliki oleh individu manusia. Fitur-fitur fisiologis yang khas tersebut meliputi sidik jari, wajah, tangan, iris, retina, suara, tanda tangan, dan cara pengetikan [2].

Karakteristik biometri yang khas pada setiap individu dapat digunakan sebagai metode identifikasi pemiliknya. Keunikan tersebut dapat diukur melalui penggunaan sensor dan data yang dihasilkan dapat diambil untuk proses pengenalan dan identifikasi. Data yang berhasil

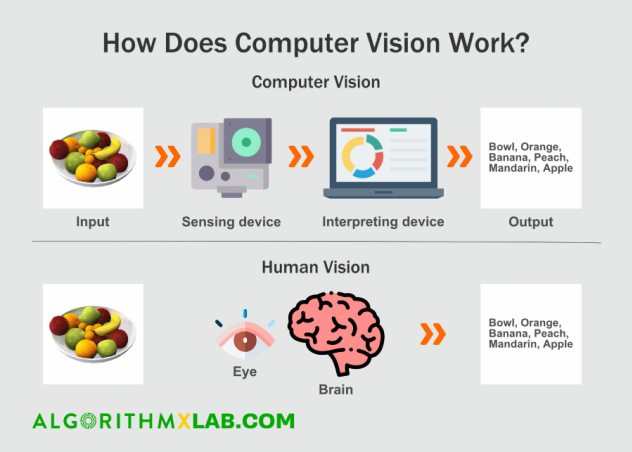
diperoleh kemudian diubah menjadi format digital, sehingga dapat digunakan sebagai data untuk mengidentifikasi individu tersebut [12].

2.2.3 Computer Vision

Computer vision atau visi komputer adalah bidang ilmu kecerdasan buatan yang fokus pada pengembangan metode untuk mengambil informasi numerik atau simbolik dari citra. Dalam *computer vision*, terdapat beberapa proses yang meliputi pengambilan citra, pemrosesan citra, segmentasi, ekstraksi fitur, dan pengklasifikasian.

Tujuan dari *computer vision* adalah untuk mengenali dan memahami konten dari gambar atau citra digital. Tujuan ini sering dicapai dengan menggunakan metode pengembangan yang berusaha mereplikasi kemampuan penglihatan manusia (*human vision*) melalui pemberian data pelatihan yang sesuai. Memahami isi citra digital dapat dengan serta merta melakukan eksplorasi informasi yang terkandung dalam deskripsi citra.

Pada dasarnya, *computer vision* berupaya meniru kemampuan visual manusia. Proses penglihatan manusia sangat kompleks, di mana manusia melihat objek melalui indra penglihatan, lalu informasi visual tersebut diteruskan ke otak untuk diterjemahkan sehingga manusia dapat memahami objek yang terlihat. Hasilnya kemudian dapat digunakan untuk proses pengambilan keputusan. Sama halnya dengan mata dan otak, *computer vision* merupakan sistem yang memiliki kemampuan untuk menganalisis objek secara visual setelah objek tersebut diwakili dalam bentuk citra [21].



Gambar 2.1 Prinsip kerja *computer vision* [22]

2.2.4 Face Recognition

Face recognition atau pengenalan wajah merupakan aplikasi dari teknologi *computer vision* yang menggunakan wajah sebagai objek visualnya. Pengenalan wajah merupakan implementasi dari sistem biometri yang bertujuan untuk identifikasi dan verifikasi seseorang berdasarkan wajahnya. Wajah adalah bagian dari tubuh manusia yang bersifat unik atau memiliki ciri khas tersendiri pada setiap individu. Dengan demikian wajah dapat dijadikan sebagai objek untuk mengidentifikasi atau memverifikasi seseorang [23].

Proses pada sistem pengenalan wajah dilakukan dengan cara membandingkan citra suatu wajah dengan wajah lainnya yang sudah dilatih dan disimpan di dalam *database* hingga memperoleh kecocokan pada wajah yang sedang dikenali [3][23].

Mekanisme proses pengenalan wajah terbagi menjadi tiga tahapan sebagai berikut:

a) Deteksi wajah (*face detection*)

Deteksi wajah adalah proses pengenalan bentuk wajah pada manusia dengan membandingkan pola tekstur dan kontur wajah pada citra digital [23]. Beberapa metode untuk deteksi wajah diantaranya

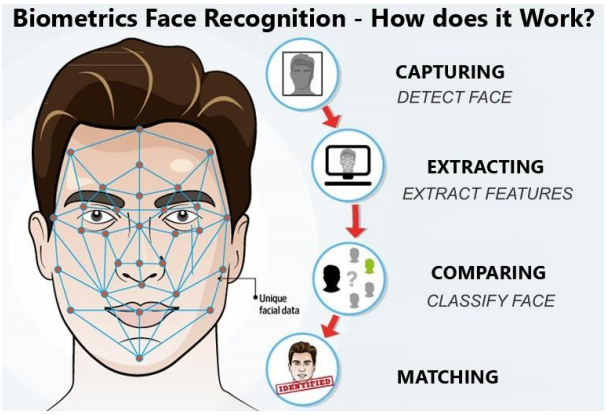
adalah *Haar Cascade Classifier* (Viola-Jones), *Histogram of Oriented Gradients* (HOG), pengembangan arsitektur CNN seperti R-CNN (*Region-based Convolution Neural Network*), YOLO (*You Only Look Once*) dan SSD (*Single Shot Detector*).

b) Ekstraksi fitur (*feature extraction*)

Ekstraksi fitur adalah tahap di mana ciri-ciri atau informasi penting diambil dari objek (wajah) yang ada pada citra. Fitur, pada gilirannya, merujuk pada karakteristik unik yang dimiliki oleh objek tersebut [2]. Beberapa metode yang dapat digunakan untuk ekstraksi fitur adalah *Principal Component Analysis* (PCA) dan pengembangannya seperti *Eigenface* dan *Fisherface*, *Local Binary Pattern* (LBP) dan pengembangannya seperti *Local Binary Pattern Histogram* (LBPH), CNN dan pengembangannya seperti *Deep Convolutional Neural Networks* (DCNN).

c) Klasifikasi wajah (*face classification*)

Klasifikasi wajah adalah proses pengategorian wajah ke dalam kelas-kelas yang tepat berdasarkan nilai kecocokan pada ciri-ciri wajah [2]. Beberapa metode yang dapat digunakan untuk klasifikasi wajah adalah LBPH, *Support Vector Machines* (SVM), *k-Nearest Neighbors* (k-NN), dan juga CNN serta pengembangannya seperti DNN.



Gambar 2.2 Proses pengenalan wajah [24]

2.2.5 OpenCV

OpenCV (*Open-Source Computer Vision*) adalah pustaka atau *library* perangkat lunak pembelajaran mesin dengan lisensi sumber terbuka (*open source*). OpenCV menyediakan 2500 lebih algoritma yang telah dioptimalkan untuk digunakan dalam bidang *computer vision* dan *machine learning*. Algoritma-algoritma yang tersedia dapat digunakan untuk mendeteksi wajah, mengenali wajah, mengidentifikasi objek, dan masih banyak lagi [25].

Beberapa algoritma pada OpenCV yang digunakan untuk *face recognition* diantaranya adalah *Haar Cascade Classifier*, LBPH (*Local Binary Pattern Histogram*), *Eigenface*, *Fisherface* dan DNN (*Deep Neural Network*) seperti SSD (*Single Shot Multibox Detector*) [26].

2.2.6 *Single Shot Detector* (SSD)

Single Shot Multibox Detector atau disebut juga *Single Shot Detector* merupakan metode yang digunakan untuk mendeteksi atau mengenali objek pada citra menggunakan jaringan saraf tiruan tunggal (*single deep neural network*). Metode ini merupakan salah satu algoritma deteksi objek yang sangat populer karena kemudahan implementasinya dan menghasilkan akurasi yang baik dengan kebutuhan komputasi yang

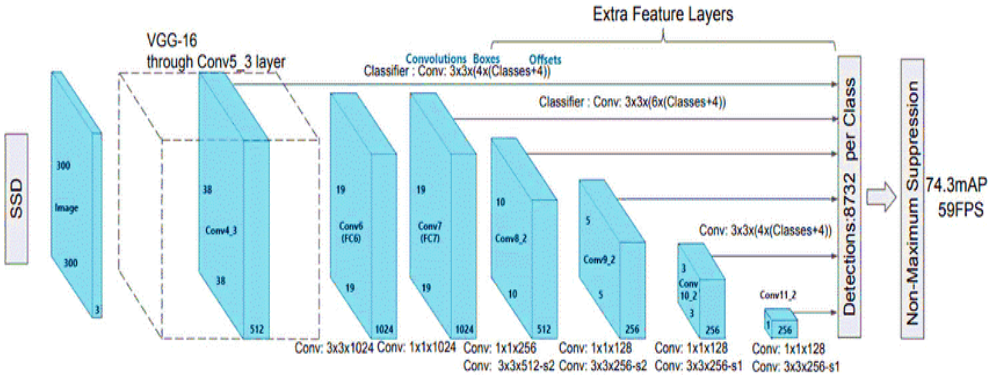
relatif rendah, dimana SSD hanya memerlukan satu pemrosesan tunggal untuk mendeteksi *multiple* objek dalam sebuah gambar [8].

SSD termasuk dalam kategori kerangka kerja (*framework*) One-Step, juga dikenal sebagai *Regression* atau *Classification Based Framework*, seperti halnya YOLO (*You Only Look Once*) atau RetinaNet. Dalam *framework* ini, terdapat pemetaan eksplisit antara nilai piksel, koordinat kotak pembatas (*bounding box*), dan probabilitas kelas, berbeda dengan *framework* berbasis proposal wilayah (*Region Proposal-Based*) seperti Faster R-CNN. Oleh karena itu, dibandingkan dengan Faster R-CNN dan arsitektur serupa, SSD memiliki waktu inferensi yang lebih cepat untuk mencapai kinerja *real-time* [27]. SSD tergolong sebagai algoritma yang sangat kuat untuk melakukan deteksi objek [28].

Gambar 2.3 *Real-time detection* menggunakan SSD [28]

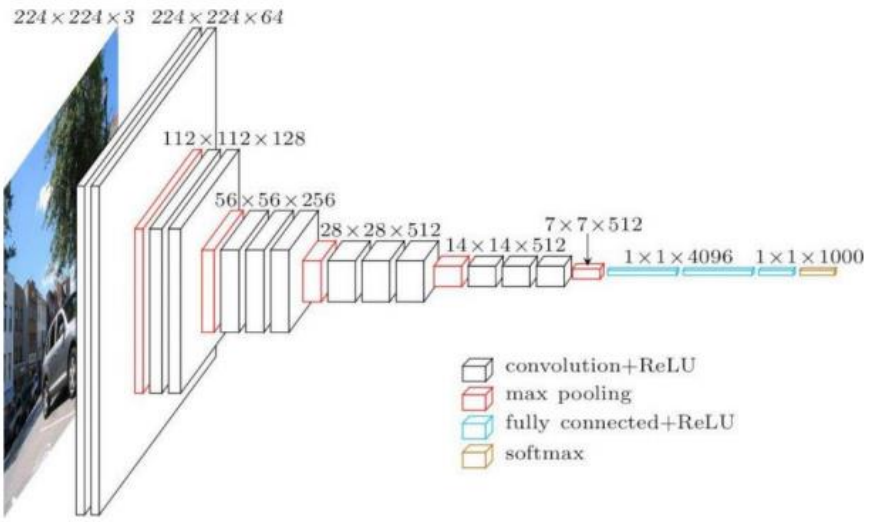
Arsitektur SSD, yang terlihat pada Gambar 2.4, terdiri dari dua komponen utama: ekstraksi fitur dan deteksi objek. Bagian pertama menggunakan model klasifikasi yang canggih seperti jaringan VGG16 (seperti yang ditampilkan pada Gambar 2.5), namun juga memungkinkan penggunaan model lain seperti ResNet atau MobileNet. Bagian ini disebut sebagai *backbone* atau struktur jaringan dasar yang tujuannya adalah menghasilkan peta fitur (*feature map*) tingkat tinggi dari gambar

input. Selain itu, SSD menambahkan enam *feature map* tambahan dengan dimensi spasial yang lebih rendah; seperti yang ditunjukkan pada Gambar 2.5 [27].



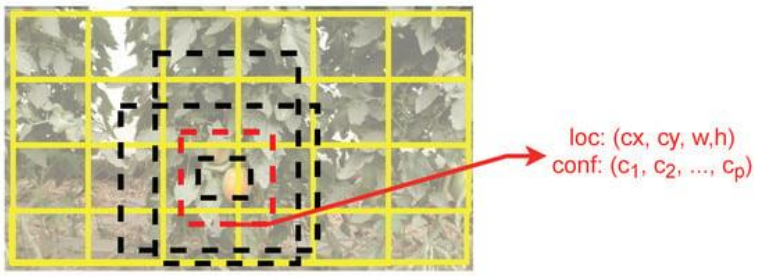
Gambar 2.4 Arsitektur SSD menggunakan *backbone* VGG16 [8]

Alasan mengapa VGG16 digunakan sebagai *backbone* atau jaringan dasar adalah karena kinerjanya yang kuat dalam tugas klasifikasi citra yang berkualitas tinggi serta popularitasnya dalam masalah-masalah di mana *transfer learning* dapat membantu meningkatkan hasil, yang dalam hal ini adalah *pretrained model*. Dengan memodifikasi VGG asli, *layer-layer* konvolusional tambahan (mulai dari conv6 dan seterusnya) ditambahkan, yang memungkinkan ekstraksi fitur pada berbagai skala dan secara bertahap mengurangi ukuran input ke setiap *layer* berikutnya. Gambar 2.5 menunjukkan contoh arsitektur dari VGG16 [8].



Gambar 2.5 Arsitektur VGG16

SSD menggunakan serangkaian *bounding box* atau kotak pembatas *default* (juga dikenal sebagai *anchor box*) dengan rasio aspek yang berbeda dan skala yang beragam (lihat Gambar 2.6). Hal ini membantu mengurangi jumlah bentuk yang mungkin diwakili oleh kotak pembatas. *Layer* konvolusi bertanggung jawab untuk memprediksi *offset* lokasi dan nilai *confidence* masing-masing kelas untuk setiap *anchor box* dalam setiap operasi konvolusi pada *feature map* tambahan. Penerapan lapisan konvolusi ini juga dilakukan pada keluaran Conv4_3 (Gambar 2.4) dalam kasus arsitektur VGG16. Dengan menggabungkan prediksi dari *feature map* dengan resolusi yang berbeda, SSD dapat mendeteksi objek dengan ukuran yang berbeda. Dalam Gambar 2.4, penggunaan *feature map* ke arah kanan akan menghasilkan deteksi objek yang lebih besar, dan sebaliknya [27].



Gambar 2.6 *Anchor box* digunakan pada arsitektur SSD [27]

Akan ada banyak prediksi deteksi yang dihasilkan, oleh karena itu, *layer Non-Maximum Suppression* (NMS) (Gambar 2.4) digunakan untuk mempertahankan kotak pembatas dengan skor tertinggi. Dalam proses pelatihan, bobot yang seimbang antara *localization loss* (misalnya, *smooth L1*) dan *confidence loss* (misalnya, *softmax*) digunakan [27].

2.2.7 *Local Binary Pattern Histogram* (LBPH)

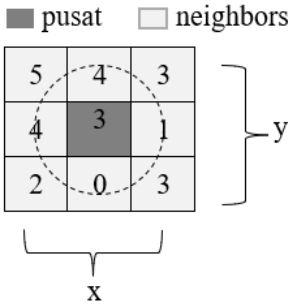
LBPH adalah salah satu metode klasifikasi wajah yang sangat populer. LBPH menyempurnakan algoritma LBP (*Local Binary Pattern*) dengan melakukan pemrosesan pada nilai histogram [29]. LBP adalah sebuah metode untuk membedakan objek dari latar belakang dalam proses pengenalan objek. Metode LBP diperkenalkan oleh Timo Ojala dan David Hardwood. Operasi LBP menggunakan perbandingan nilai *grayscale* dari tiap piksel dengan tetangganya. Salah satu keunggulan metode LBP adalah kemudahan implementasinya serta kemampuan untuk mengekstraksi fitur dengan kecepatan tinggi dan komputasi yang rendah [30].

Tahapan-tahapan dasar pada metode LBPH sama dengan metode LBP, yaitu menggunakan operasi LBP. Berikut langkah-langkah atau tahapan algoritma LBPH [25]:

a) Variabel Parameter

LBP memiliki 4 variabel parameter yaitu radius, *neighbors* (tetangga), *grid* x, dan *grid* y. Variabel radius digunakan untuk membentuk LBP *circular* dengan menggambarkan area lingkaran di

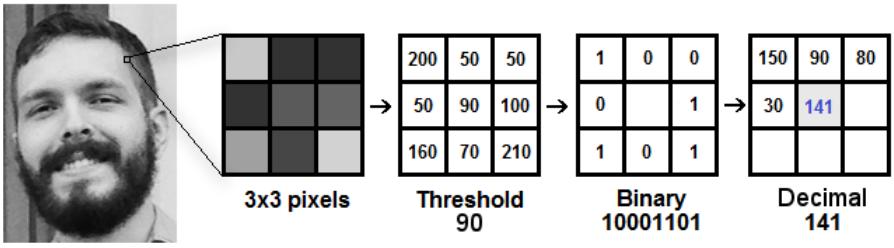
sekitar piksel pusat dan nilai *default* radius ini adalah 1 (1 piksel). Variabel *neighbors* menentukan jumlah titik sampel piksel yang digunakan dalam pembentukan LBP *circular*, dengan nilai *default* biasanya adalah 8. Variabel *grid* x dan *grid* y digunakan untuk membagi gambar menjadi sel-sel piksel secara horizontal dan vertikal.



Gambar 2.7 Parameter Operator LBP pada piksel 3x3

b) Menerapkan Operasi LBP

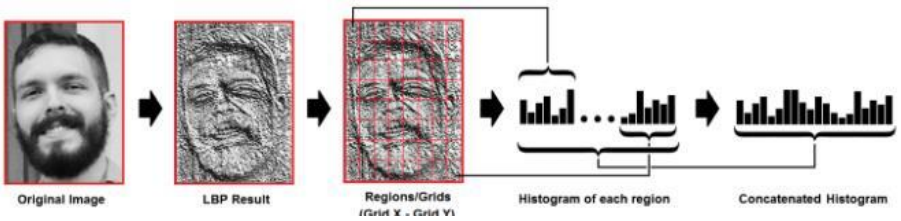
Operasi ini dilakukan untuk memperoleh nilai LBP dari setiap piksel berukuran 3x3 pada citra yang telah diubah menjadi *grayscale*. Nilai yang terdapat pada setiap piksel mencerminkan intensitas cahaya dengan rentang antara 0 hingga 255. Kemudian, nilai pada piksel pusat digunakan sebagai ambang batas (*threshold*). Nilai *threshold* akan menentukan nilai piksel tetangga sebanyak 8 piksel. Jika nilai piksel tetangga lebih besar dari nilai *threshold*, nilainya akan diubah menjadi 1, sedangkan jika lebih kecil dari nilai *threshold*, nilainya akan diubah menjadi 0. (Lihat Gambar 2.8) Setelah proses ini selesai, semua nilai dalam piksel 3x3 tersebut akan menjadi nilai biner. Nilai biner ini akan diatur searah jarum jam dan dikonversikan menjadi nilai desimal. Nilai desimal ini akan menjadi *threshold* untuk membentuk citra baru. Proses ini akan diulang sampai seluruh piksel pada citra asli dikonversi dan membentuk citra baru yang mewakili karakteristik wajah dengan lebih baik daripada citra asli.



Gambar 2.8 Operasi LBP pada piksel 3x3 pada citra *grayscale* [25]

c) Ekstraksi Histogram

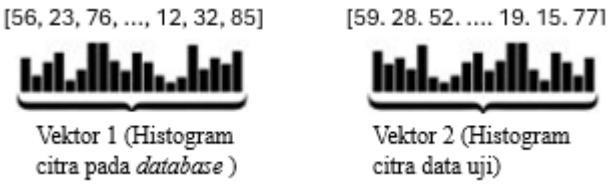
Pada langkah ini, citra yang telah dihasilkan akan dibagi menjadi beberapa *grid*, yaitu *grid* x dan *grid* y. Proses transformasi dari gambar *grayscale* menjadi histogram dilakukan dengan menggunakan metode LBP, seperti yang ditunjukkan dalam Gambar 2.9. Karena citra dalam bentuk *grayscale*, histogram yang dihasilkan hanya memiliki 256 posisi (0-255). Setiap histogram dari *dataset* akan digabungkan untuk membentuk pola (*pattern*) histogram yang lebih besar, yang akan merepresentasikan karakteristik keseluruhan dari citra atau gambar aslinya.



Gambar 2.9 Proses ekstraksi nilai LBP ke nilai histogram [25]

d) Pencocokan Nilai Histogram

Tahap berikutnya adalah pencocokan nilai histogram dalam proses pengenalan wajah. Pada pengujian input citra wajah hasil deteksi akan diekstrak nilai histogramnya sama seperti langkah-langkah sebelumnya, kemudian setelah diperoleh nilai histogramnya akan dibandingkan dengan nilai histogram citra wajah pada *database* yang telah di-*training*. Perbandingan berdasarkan perbedaan jarak *Euclidean*, yaitu jarak antar vektor histogram. Semakin kecil nilai jarak *Euclidean*, maka semakin mirip kedua histogram tersebut [31].



Gambar 2.10 Pencocokan Nilai 2 Histogram

Jika nilai histogram citra wajah uji mendekati nilai histogram citra wajah pada *database*, maka akan menghasilkan nilai *confidence* yang menunjukkan wajah dikenali [25]. Perhitungan jarak *Euclidean ED* untuk dua vektor $(p_1, p_2, \dots, p_n)$ dan $(q_1, q_2, \dots, q_n)$ dihitung menggunakan rumus berikut:

$$ED = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}$$

Keterangan:

$ED =$ Jarak Euclidean

$p_i =$ Frekuensi distribusi pola biner ke-i citra *database* model

$q_i =$ Frekuensi distribusi pola biner ke-i citra data uji (input)

2.2.8 Caffe Model

Caffe (*Convolutional Architecture for Fast Feature Embedding*) adalah sebuah *framework deep learning open-source* yang banyak digunakan untuk tugas-tugas *computer vision* dan *machine learning*. Model yang dilatih menggunakan *Caffe framework* disebut sebagai *Caffe* model. *Caffe Model* dapat digunakan untuk berbagai tugas *deep learning*, seperti [32]:

1) Klasifikasi gambar
2) Deteksi objek
3) Deteksi wajah

4) *Natural language processing*
5) *Machine translation*

Caffe Model memiliki beberapa keunggulan, yaitu:

1) Kecepatan komputasi yang tinggi
2) Akurasi yang baik
3) Fleksibilitas dalam penggunaan

2.2.9 Confusion Matrix

Confusion Matrix adalah tabel yang digunakan untuk menggambarkan kinerja suatu model klasifikasi pada suatu set data, dengan membandingkan hasil prediksi model dengan kelas sebenarnya dari data tersebut. *Confusion matrix* mencakup empat elemen [33]:

a) *True Positive* (TP): Jumlah kasus di mana model dengan benar memprediksi suatu kelas sebagai positif.
b) *True Negative* (TN): Jumlah kasus di mana model dengan benar memprediksi suatu kelas sebagai negatif.
c) *False Positive* (FP): Jumlah kasus di mana model keliru memprediksi suatu kelas sebagai positif (kesalahan *Type* I).
d) *False Negative* (FN): Jumlah kasus di mana model keliru memprediksi suatu kelas sebagai negatif (kesalahan *Type* II).

Contoh *Confusion matrix* digambarkan pada Tabel 2.2 berikut:

Tabel 2.2 *Confusion Matrix*

	Prediction	
Actual	Terdeteksi	Tidak Terdeteksi
Terdeteksi	TP	FN
Tidak Terdeteksi	FP	TN

Confusion matrix dapat digunakan untuk mengukur nilai akurasi dan presisi:

$$Presisi\ (Precision) = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP)\ +\ False\ Positive\ (FP)}$$

$$Akurasi\ (Accuracy) = \frac{True\ Positive\ (TP)\ +\ True\ Negative\ (TN)}{Total\ Data}$$

2.2.10 Python

Python merupakan bahasa pemrograman tingkat tinggi (*high-level*) yang paling mendekati bahasa alami manusia. Python menjadi bahasa pemrograman yang paling populer terutama di dalam pembelajaran mesin (*machine learning*) karena kemampuannya yang andal dan *powerful*, baik secara sintaksis maupun performa [8][31].

Python memiliki lisensi sumber terbuka (*open source*) yang menyediakan banyak pustaka (*library*) *multi-platform*. Salah satu *library* yang menggunakan Python sebagai salah satu bahasa pemrogramannya dan sering digunakan untuk *computer vision* adalah OpenCV [34].

2.2.11 Flask

Flask adalah kerangka kerja (*framework*) web mikro yang menggunakan bahasa pemrograman Python. Flask adalah *framework* yang ringan dan sederhana yang dirancang untuk membangun aplikasi web dengan cepat dan mudah tanpa perlu membuatnya dari awal. Flask memiliki fokus yang kuat pada kesederhanaan, fleksibilitas, dan kebebasan dalam pengembangan. Walaupun Flask terkenal sebagai sebuah *framework* dengan inti yang sederhana, ini tidak berarti bahwa Flask memiliki keterbatasan dalam hal fungsionalitas. Flask memiliki inti yang sederhana namun tetap memungkinkan untuk menambahkan ekstensi yang diperlukan sesuai dengan kebutuhan. Dengan demikian, Flask dapat dianggap sebagai sebuah *framework* yang memiliki fleksibilitas dan skalabilitas yang tinggi jika dibandingkan dengan *framework* lainnya [35].