BAB II

TINJAUAN LITERATUR

2.1 Penelitian Terkait

Bagian ini memaparkan kajian teoretis dan empiris yang menjadi landasan dalam penelitian klasifikasi kecantikan wajah wanita berbasis ekstraksi fitur geometris. Kajian pustaka disusun secara sistematis dengan menelaah penelitian-penelitian terdahulu yang relevan sebagai acuan perbandingan sekaligus justifikasi atas pendekatan yang digunakan dalam penelitian ini.

Penelitian yang dilakukan oleh (Fansyuri and Yunita, 2023) berfokus pada implementasi algoritma K-Nearest Neighbor (KNN) untuk klasifikasi jenis kelamin berdasarkan analisis citra wajah. Penelitian ini bertujuan untuk mengatasi permasalahan ketidakakuratan deteksi wajah, terutama dalam membedakan wajah laki-laki dan perempuan, yang disebabkan oleh faktor seperti tumpang tindih objek, halangan pada citra, atau ketidakjelasan gambar. Metode yang digunakan terdiri dari dua fase utama, yaitu fase pelatihan dan fase pengujian. Dataset yang digunakan berjumlah 200 citra wajah dari 10 orang, yang dibagi menjadi 150 data latih dan 50 data uji. Proses ekstraksi fitur dilakukan dengan menggabungkan fitur warna (nilai rata-rata HSV dan RGB) dan fitur bentuk (area, eccentricity, dan metric) setelah melalui segmentasi menggunakan K-Means Clustering dan operasi morfologi untuk menghilangkan noise. Hasil pengujian menunjukkan bahwa metode KNN dengan nilai K=2 menghasilkan akurasi tertinggi sebesar 96% berdasarkan *confusion matrix,* dengan nilai AUC sebesar 0,996 yang mengindikasikan kinerja klasifikasi yang sangat baik. Penelitian ini menyimpulkan

bahwa KNN merupakan algoritma yang sangat baik untuk klasifikasi citra wajah berdasarkan warna dan bentuk, dan menyarankan penambahan objek lain seperti tinggi badan atau postur tubuh untuk hasil yang lebih optimal di masa mendatang.

Penelitian yang dilakukan oleh (Agnes and Yamasari, 2024) berfokus pada analisis perbandingan antara dua metode pengenalan wajah, yaitu K-Nearest Neighbor (KNN) dan Local Binary Pattern Histogram (LBPH), dalam implementasi sistem absensi online. Proses penelitian meliputi akuisisi gambar, preprocessing (konversi grayscale, deteksi wajah dengan Haar Cascade, resize ke 100x100 piksel), ekstraksi fitur, dan klasifikasi. Pada metode KNN, dataset dibagi menjadi 80% data latih dan 20% data uji, dengan pengujian nilai K dari 1 hingga 9 menggunakan perhitungan jarak Euclidean. Hasil pengujian menunjukkan bahwa KNN dengan K=1 menghasilkan akurasi tertinggi sebesar 98,52%, sedangkan K=3, 5, dan 9 menghasilkan akurasi 95,57%. Sementara itu, metode LBPH diuji pada 8 individu dengan total 808 citra dan menghasilkan tingkat kecocokan tertinggi sebesar 81% per individu. Penelitian ini menyimpulkan bahwa KNN lebih unggul dalam hal akurasi klasifikasi dan lebih sesuai untuk sistem absensi, sedangkan LBPH lebih cocok untuk identifikasi individu berdasarkan fitur tekstur wajah. Penulis menyarankan untuk penelitian selanjutnya agar menggabungkan kedua metode atau mengintegrasikannya dengan metode lain untuk meningkatkan performa sistem.

Penelitian yang dilakukan oleh (Singh, Sharma and Singh, 2022) berfokus pada analisis komparatif kemampuan berbagai fitur tekstur dan geometri dalam klasifikasi massa payudara pada citra mammogram menggunakan algoritma k-

nearest neighbor (k-NN). Dataset yang digunakan adalah INbreast yang terdiri dari 106 citra mammogram digital dengan total 115 lesi massa payudara (52 jinak dan 63 ganas). Metode yang digunakan meliputi ekstraksi 125 fitur tekstur dan geometri dari berbagai model, kemudian seleksi fitur menggunakan empat algoritma filter-based (Relief-F, Pearson correlation coefficient, neighborhood component analysis, dan term variance) untuk memilih 20 fitur paling diskriminatif. Hasil menunjukkan bahwa Relief-F merupakan algoritma seleksi fitur terbaik dengan akurasi 85,2%. Eksperimen lebih lanjut menghasilkan 9 fitur paling diskriminatif, yaitu FD26 (Fourier Descriptor), Euler number, solidity, mean, FD14, FD13, periodicity, skewness, dan contrast. Penelitian ini menyimpulkan bahwa fitur geometri (terutama Fourier Descriptor) lebih diskriminatif dibandingkan fitur tekstur, namun kombinasi keduanya memberikan hasil terbaik. Seleksi fitur berkontribusi signifikan dalam meningkatkan akurasi dari 76,0% (tanpa seleksi) menjadi 90,4% (dengan 9 fitur terpilih). Penelitian ini menunjukkan bahwa pendekatan machine learning konvensional dengan seleksi fitur yang tepat masih mampu memberikan performa yang kompetitif.

Penelitian yang dilakukan oleh (Marinatul, Luluk and Siti, 2024) berfokus pada implementasi algoritma K-Nearest Neighbor (K-NN) dengan ekstraksi ciri Gray Level Co-occurrence Matrix (GLCM) untuk mengklasifikasikan enam jenis ekspresi wajah manusia, yaitu bahagia, sedih, marah, takut, terkejut, dan jijik. Dataset yang digunakan terdiri dari citra ekspresi wajah mahasiswi yang diambil menggunakan kamera DSLR dengan jarak 1 meter, latar hijau, dan pencahayaan blitz. Metode yang digunakan meliputi konversi citra RGB ke grayscale, deteksi

wajah, ekstraksi fitur GLCM (Contrast, Correlation, Energy, dan Homogeneity), serta klasifikasi menggunakan K-NN dengan nilai k = 1, 3, dan 5. Pengujian dilakukan pada dua skenario: data pelatihan 390 dan data uji 102 (jumlah lebih banyak), serta data pelatihan 138 dan data uji 30 (jumlah lebih sedikit). Hasil penelitian menunjukkan bahwa akurasi tertinggi dicapai pada skenario data lebih banyak dengan k = 1, yaitu 100% pada proses pelatihan dan pengujian, sedangkan penggunaan k = 3 dan k = 5 menghasilkan akurasi yang jauh lebih rendah (40-46% untuk pelatihan dan 33-40% untuk pengujian). Penelitian ini menyimpulkan bahwa algoritma K-NN dapat mencapai akurasi terbaik dalam klasifikasi ekspresi wajah ketika menggunakan jumlah data yang lebih banyak dan nilai k = 1. Penelitian ini juga menekankan bahwa penentuan nilai k yang tepat sangat krusial karena penggunaan k yang terlalu besar dapat menurunkan performa klasifikasi secara signifikan.

Penelitian yang dilakukan oleh (Nisa, Muslem and Almuslim, 2026) berfokus pada pengembangan model pengenalan wajah menggunakan algoritma K-Nearest Neighbors (KNN) dengan ekstraksi fitur Histogram of Oriented Gradients (HOG). Penelitian ini dilatarbelakangi oleh kebutuhan akan sistem pengenalan wajah yang ringan secara komputasi namun tetap mampu memberikan performa stabil pada kondisi nyata, sebagai alternatif dari metode deep learning yang membutuhkan sumber daya besar. Hasil eksperimen menunjukkan bahwa nilai k optimal adalah 4 dengan akurasi validasi 75,37%. Pengujian lebih lanjut pada data uji menghasilkan akurasi 81,37% dengan rata-rata F1-score 0,81. Analisis confusion matrix menunjukkan bahwa sebagian besar prediksi berada pada

diagonal utama, meskipun masih terdapat kesalahan klasifikasi pada kelas dengan kemiripan fitur wajah, seperti kelas Rifa (recall 0,23) dan Aceng (precision 0,45). Penelitian ini menyimpulkan bahwa kombinasi HOG dan KNN efektif dan efisien untuk pengenalan wajah pada dataset skala kecil hingga menengah dengan keunggulan waktu pelatihan yang sangat cepat, namun memiliki keterbatasan pada tahap prediksi yang semakin lambat seiring bertambahnya data. Penulis menyarankan penelitian selanjutnya menggunakan dataset lebih besar atau mengkombinasikan metode yang lebih kompleks untuk meningkatkan performa dan ketahanan sistem.

2.2 Pengolahan Citra Digital

Pengolahan citra digital (digital image processing) merupakan serangkaian teknik komputasi yang digunakan untuk memanipulasi dan menganalisis citra dalam format digital guna menghasilkan informasi yang lebih bermakna. Secara fundamental, citra digital direpresentasikan sebagai matriks dua dimensi yang terdiri dari elemen-elemen piksel (picture element), di mana setiap piksel memiliki nilai intensitas yang merepresentasikan kecerahan atau warna pada posisi spasial tertentu. Pada citra berwarna, ruang warna RGB (Red, Green, Blue) merepresentasikan setiap piksel dengan tiga nilai kanal warna yang masing-masing berkisar antara 0 hingga 255 (Hardini *et al.*, 2023). Rangkaian operasi dalam pengolahan citra biasanya mengalir secara berurutan, dimulai dari akuisisi, deteksi objek, ekstraksi ciri, hingga interpretasi atau klasifikasi. Pada penelitian ini, citra wajah wanita diakuisisi dalam format RGB, kemudian diproses langsung oleh modul deteksi landmark tanpa memerlukan tahap konversi grayscale maupun resize

eksplisit, karena pustaka MediaPipe Face Mesh yang digunakan telah dirancang untuk bekerja pada berbagai resolusi dan kondisi pencahayaan citra masukan.

2.3 Kecantikan Wajah Dan Proposrsi Klasik

Kajian tentang kecantikan wajah telah menjadi objek penelitian lintas disiplin sejak zaman Yunani kuno, yang melahirkan berbagai kanon proporsi wajah ideal, salah satunya golden ratio atau rasio emas dengan nilai mendekati 1,618. Golden ratio dipercaya merepresentasikan proporsi yang secara estetis dianggap paling harmonis oleh persepsi visual manusia, dan sering diaplikasikan pada rasio tinggi terhadap lebar wajah maupun rasio antar-segmen wajah (Sadiq and Abdulazeez, 2024). Selain golden ratio, kanon proporsi neoklasik (neoclassical canons) membagi wajah secara vertikal menjadi tiga segmen yang idealnya memiliki panjang yang sama, yaitu dari batas rambut ke alis, dari alis ke dasar hidung, dan dari dasar hidung ke dagu.

Selain proporsi, simetri wajah (facial symmetry) juga merupakan salah satu indikator estetika yang banyak dikaji, karena wajah manusia secara alami memiliki sedikit asimetri antara sisi kiri dan kanan, dan tingkat asimetri yang lebih rendah kerap diasosiasikan dengan persepsi kecantikan yang lebih tinggi pada berbagai studi psikologi persepsi (Iyer et al., 2021). Meskipun prinsip-prinsip proporsi klasik tersebut telah digunakan secara luas sebagai proksi kuantitatif kecantikan pada berbagai penelitian facial beauty prediction, penting untuk ditegaskan bahwa prinsip-prinsip ini merupakan konvensi estetika yang berasal dari tradisi budaya tertentu (Yunani dan Tiongkok kuno), dan tidak dapat diklaim sebagai definisi kecantikan yang berlaku universal lintas budaya. Penelitian ini menggunakan

prinsip-prinsip tersebut murni sebagai kerangka pengukuran geometris yang terukur dan dapat direproduksi, bukan sebagai klaim ilmiah atas standar kecantikan yang mutlak.

2.4 MediaPipe Face Mesh

MediaPipe Face Mesh merupakan solusi machine learning yang dikembangkan oleh Google untuk mengestimasi topologi tiga dimensi wajah manusia secara real-time hanya dengan menggunakan citra atau video sebagai masukan, tanpa memerlukan perangkat keras khusus seperti sensor kedalaman (depth sensor). Pustaka ini mampu mendeteksi lebih dari 468 titik landmark tiga dimensi pada wajah, yang mencakup kontur wajah, alis, mata, hidung, bibir, serta area sekitar mulut (Effendi and Widyaningrum, 2026). Setiap titik landmark direpresentasikan dalam koordinat (x, y, z) ternormalisasi terhadap ukuran citra, sehingga posisi setiap landmark dapat dikonversi menjadi koordinat piksel aktual dengan mengalikan nilai x dan y terhadap lebar dan tinggi citra. Pada penelitian ini, MediaPipe Face Mesh digunakan dengan mode refine_landmarks aktif untuk meningkatkan presisi deteksi pada area mata dan bibir, serta ambang kepercayaan deteksi (minimum detection confidence) sebesar 0,3 agar sistem tetap mampu mendeteksi wajah pada kondisi pencahayaan maupun sudut pengambilan citra yang bervariasi.

2.5 Ekstraksi Fitur Geometris Wajah

Setelah landmark wajah berhasil dideteksi, sistem mengekstraksi 12 fitur geometris yang dikelompokkan ke dalam empat kategori, yaitu simetri wajah, proporsi wajah/golden ratio, proporsi hidung, dan proporsi bibir. Seluruh fitur

dihitung berdasarkan jarak Euclidean antar titik landmark tertentu pada bidang dua dimensi.

2.5.1 Simetri Wajah

Simetri wajah dihitung dengan membandingkan jarak titik landmark kiri dan kanan terhadap garis tengah wajah (mid-x), yang diperoleh dari rata-rata koordinat horizontal pipi kiri dan kanan. Untuk sepasang titik landmark kiri dan kanan, rasio simetri dihitung menggunakan Persamaan :

$$S = \frac{min(dL, dR)}{max(dL, dR)}$$

dengan dL dan dR masing-masing merupakan jarak absolut titik landmark kiri dan kanan terhadap garis tengah wajah. Nilai S mendekati 1 menunjukkan simetri yang sempurna antara sisi kiri dan kanan wajah. Penelitian ini menghitung empat nilai simetri, yaitu simetri mata, simetri pipi, simetri rahang, dan simetri alis, yang masing-masing menjadi satu fitur pada vektor fitur akhir.

2.5.2 Golden Ratio dan Proporsi Wajah

Rasio golden ratio wajah dihitung sebagai perbandingan antara tinggi wajah (jarak dahi ke dagu) terhadap lebar wajah (jarak pipi kiri ke pipi kanan), sebagaimana dirumuskan pada Persamaan:

$$GR = \frac{Htinggi\ wajah}{Wlebar\ wajah}$$

dengan nilai GR ideal mendekati 1,618 sesuai dengan nilai golden ratio klasik. Selain itu, sistem juga menghitung skor proporsi tiga segmen vertikal wajah

(dahi–mata, mata–hidung, hidung–dagu) berdasarkan koefisien variasi (coefficient of variation) ketiga segmen tersebut, sebagaimana dirumuskan pada Persamaan:

$$CV = \frac{abs(s1 - sbar) + abs(s2 - sbar) + abs(s3 - sbar)}{3\ sbar}$$

$$Skorproporsi = \frac{1}{1 + CV}$$

dengan s1, s2, dan s3 masing-masing merupakan panjang ketiga segmen vertikal wajah, dan sbar merupakan rata-rata ketiganya. Semakin kecil nilai CV (semakin seragam ketiga segmen), semakin tinggi skor proporsi yang dihasilkan, sesuai dengan prinsip kanon neoklasik yang mengasumsikan proporsi ideal ketika ketiga segmen memiliki panjang yang sama.

2.5.3 Proporsi Hidung dan Bibir

Proporsi hidung dihitung melalui rasio aspek (aspect ratio) antara tinggi dan lebar hidung, rasio lebar hidung terhadap lebar wajah, serta simetri sayap hidung menggunakan rumus simetri yang sama dengan Persamaan. Rasio aspek hidung dirumuskan pada Persamaan:

$$ARhidung = \frac{Hhidung}{Whidung}$$

Sementara itu, proporsi bibir dihitung melalui rasio lebar bibir terhadap lebar wajah dan rasio lebar bibir terhadap lebar hidung, sebagaimana dirumuskan pada Persamaan:

$$Rbibir = \frac{Wbibir}{Wwajah}$$

Kedua belas fitur di atas (empat fitur simetri, tiga fitur golden ratio/proporsi wajah, tiga fitur proporsi hidung, dan dua fitur proporsi bibir) kemudian digabungkan menjadi satu vektor fitur berdimensi 12 untuk setiap citra wajah, yang menjadi masukan bagi tahap normalisasi dan klasifikasi.

2.6 Normalisasi Fitur dengan StandardScaler

Karena kedua belas fitur geometris memiliki rentang nilai dan satuan yang berbeda-beda (misalnya rasio simetri berkisar 0–1, sedangkan golden ratio berkisar di sekitar 1–2), diperlukan proses normalisasi agar setiap fitur memiliki kontribusi yang seimbang pada perhitungan jarak KNN. Penelitian ini menggunakan StandardScaler, yang mengubah setiap fitur menjadi distribusi dengan rata-rata 0 dan simpangan baku 1, sebagaimana dirumuskan pada Persamaan:

$$z = \frac{x - \mu}{\sigma}$$

dengan x merupakan nilai fitur asli, μ merupakan rata-rata fitur pada data pelatihan, dan σ merupakan simpangan baku fitur pada data pelatihan. Proses fit standardisasi hanya dilakukan pada data pelatihan untuk menghindari kebocoran data (data leakage), kemudian parameter yang sama digunakan untuk mentransformasi data pengujian.

2.7 Algoritma K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) merupakan algoritma klasifikasi non-parametrik yang mengklasifikasikan suatu data baru berdasarkan mayoritas kelas dari K tetangga terdekatnya pada ruang fitur. Kedekatan antar data pada penelitian

ini diukur menggunakan jarak Euclidean, sebagaimana dirumuskan pada Persamaan:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(xi - yi)^2}$$

dengan x dan y merupakan dua vektor fitur (masing-masing berdimensi n=12) yang dibandingkan, dan xi, yi merupakan nilai fitur ke-i dari masing-masing vektor. Berbeda dengan KNN standar yang memberi bobot yang sama pada seluruh K tetangga terdekat, penelitian ini menggunakan skema KNN berbobot jarak (distance-weighted), di mana tetangga yang lebih dekat memberikan kontribusi suara yang lebih besar terhadap keputusan klasifikasi. Bobot suara tiap tetangga dirumuskan pada Persamaan :

$$wi = \frac{1}{di}$$

dengan di merupakan jarak Euclidean antara data uji dan tetangga ke-i. Skema pembobotan ini dipilih karena secara empiris memberikan hasil yang lebih stabil dibandingkan pembobotan seragam (uniform), khususnya ketika distribusi data pada ruang fitur tidak sepenuhnya merata. Nilai K pada penelitian ini dapat diatur melalui antarmuka pengguna (dengan nilai default K=5), dan dioptimalkan melalui skema validasi silang yang dijelaskan pada subbab berikutnya.

2.8 Evaluasi Kinerja Sistem dengan Confusion Matrix

Evaluasi kinerja model klasifikasi multi-kelas pada penelitian ini dilakukan menggunakan confusion matrix, yaitu tabel kontingensi yang merangkum jumlah

prediksi benar dan salah untuk setiap kelas dibandingkan dengan label aktualnya. Tabel 2.1 menyajikan struktur umum confusion matrix untuk kasus tiga kelas yang digunakan pada penelitian ini.

Tabel 2. 1 Struktur Confusion Matrix Tiga Kelas

	Prediksi: Sangat Cantik	**Prediksi: Cantik**	**Prediksi: Cukup Cantik**
Aktual: Sangat Cantik	Jumlah benar	Salah klasifikasi	Salah klasifikasi
Aktual: Cantik	Salah klasifikasi	Jumlah benar	Salah klasifikasi
Aktual: Cukup Cantik	Salah klasifikasi	Salah klasifikasi	Jumlah benar

Berdasarkan confusion matrix, untuk setiap kelas i dapat dihitung nilai True Positive (TPi), False Positive (FPi), dan False Negative (FNi), yang selanjutnya digunakan untuk menghitung empat metrik evaluasi utama sebagai berikut (Hidayah and Dodiman, 2024).

1. Akurasi (Accuracy): mengukur proporsi prediksi yang benar dari keseluruhan data uji, dirumuskan pada Persamaan :

$$Accuracy = \frac{Jumlah\ prediksi\ benar}{Jumlah\ seluruh\ data}$$

2. Presisi (Precision): mengukur proporsi prediksi suatu kelas yang benar-benar tepat dari seluruh prediksi kelas tersebut yang dihasilkan sistem, dirumuskan pada Persamaan :

$$Precision_i = \frac{TP_i}{TPi + FPi}$$

3. Recall (Sensitivity): mengukur proporsi data suatu kelas yang berhasil dideteksi dengan benar oleh sistem, dirumuskan pada Persamaan :

$$Recall_i = \frac{TP_i}{TPi + FNi}$$

4. F1-Score: rata-rata harmonik dari presisi dan recall, yang memberikan gambaran keseimbangan antara kemampuan mendeteksi suatu kelas dan menghindari kesalahan klasifikasi, dirumuskan pada Persamaan :

$$F1_i = 2 \times \frac{Precisioni \times Recalli}{Precisioni + Recalli}$$

Keempat metrik di atas dihitung untuk setiap kelas secara individual, kemudian dirata-ratakan (macro-average) untuk memperoleh gambaran kinerja model secara keseluruhan. Penggunaan confusion matrix sebagai alat evaluasi dipilih karena mampu memberikan gambaran performa sistem yang komprehensif, tidak hanya dari sisi ketepatan prediksi secara keseluruhan (akurasi), tetapi juga dari kemampuan sistem dalam membedakan ketiga kategori kecantikan secara individual, termasuk mengidentifikasi kelas mana yang paling sering tertukar dengan kelas lainnya.