

ABSTRAK

KLASIFIKASI BENTUK WAJAH WANITA CANTIK MENGUNAKAN METODE K-NEAREST NEIGHBOR BERBASIS EKTRAKSI FITUR GEOMETRIS DENGAN PYTHON

Nama : Abillaza Abdullah

NPM : 2055201140

Pembimbing : Yovi Apridiansyah, S.Kom., M.Kom

Penelitian ini mengimplementasikan sistem klasifikasi kecantikan wajah wanita berbasis computer vision menggunakan metode K-Nearest Neighbor (KNN) yang berbasis ekstraksi fitur geometris. Sistem memanfaatkan pustaka MediaPipe Face Mesh untuk mendeteksi landmark wajah, kemudian mengekstraksi 12 fitur geometris yang meliputi empat fitur simetri wajah, tiga fitur golden ratio/proporsi wajah, tiga fitur proporsi hidung, dan dua fitur proporsi bibir. Karena tidak tersedia ground truth kecantikan yang independen, pelabelan dataset primer sebanyak 100 citra wajah wanita dilakukan secara semi-otomatis berdasarkan skor komposit dari fitur geometris yang sama, kemudian ditinjau ulang secara manual, menghasilkan tiga kategori kelas yang relatif seimbang, yaitu Sangat Cantik (33 citra), Cantik (34 citra), dan Cukup Cantik (33 citra). Fitur yang telah diekstraksi dinormalisasi menggunakan StandardScaler, kemudian diklasifikasikan menggunakan algoritma KNN berbobot jarak (distance-weighted) dengan pengukuran jarak Euclidean. Evaluasi kinerja sistem dilakukan menggunakan skema 10-fold stratified cross-validation. Hasil pengujian terhadap 100 citra menunjukkan bahwa sistem mencapai akurasi sebesar 75,00%, dengan nilai presisi rata-rata 0,7669, recall rata-rata 0,7504, dan F1-score rata-rata 0,7484. Kelas Sangat Cantik memperoleh performa terbaik (F1-score 0,8732), sedangkan kelas Cantik yang berada di posisi tengah memperoleh performa terendah (F1-score 0,6575) karena lebih rentan tertukar dengan kedua kelas yang berbatasan langsung dengannya. Penelitian ini menegaskan bahwa kecantikan merupakan konstruksi persepsi yang subjektif, sehingga hasil klasifikasi yang diperoleh merefleksikan konsistensi model terhadap skema pelabelan berbasis proporsi geometris yang digunakan, bukan penilaian kecantikan oleh manusia secara independen.

Kata Kunci: Klasifikasi Kecantikan Wajah, K-Nearest Neighbor, Ekstraksi Fitur Geometris, MediaPipe Face Mesh, Golden Ratio, Confusion Matrix, Python.

ABSTRACT

CLASSIFICATION OF BEAUTIFUL FEMALE FACE SHAPES USING K-NEAREST NEIGHBOR METHOD BASED ON GEOMETRIC FEATURE EXTRACTION IN PYTHON

Name : Abillaza Abdullah

NPM : 2055201140

Advisor : Yovi Apridiansyah, S.Kom., M.Kom

This study implements a women's facial beauty classification system based on computer vision using the K-Nearest Neighbor (KNN) method based on geometric feature extraction. The system utilizes the MediaPipe Face Mesh library to detect facial landmarks, then extracts 12 geometric features covering four facial symmetry features, three golden ratio/facial proportion features, three nose proportion features, and two lip proportion features. Since no independent beauty ground truth was available, the primary dataset of 100 women's facial images was labeled semi-automatically based on a composite score derived from the same geometric features, then manually reviewed, resulting in three relatively balanced classes: Sangat Cantik (33 images), Cantik (34 images), and Cukup Cantik (33 images). The extracted features were normalized using StandardScaler and classified using a distance-weighted KNN algorithm with Euclidean distance. System performance was evaluated using a 10-fold stratified cross-validation scheme. Testing on 100 images shows that the system achieves an accuracy of 75.00%, with an average precision of 0.7669, average recall of 0.7504, and average F1-score of 0.7484. The Sangat Cantik class achieved the best performance (F1-score 0.8732), while the middle Cantik class achieved the lowest performance (F1-score 0.6575) as it is more prone to being confused with its two neighboring classes. This study emphasizes that beauty is a subjective perceptual construct, so the classification results reflect the model's consistency with the geometric-proportion-based labeling scheme used, not an independent human judgment of beauty.

Keywords: Facial Beauty Classification, K-Nearest Neighbor, Geometric Feature Extraction, MediaPipe Face Mesh, Golden Ratio, Confusion Matrix, Python.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi kecerdasan buatan (Artificial Intelligence) dan pengolahan citra digital telah membuka berbagai peluang baru dalam berbagai bidang kehidupan, termasuk dalam dunia kecantikan dan estetika wajah. Salah satu aspek yang menarik perhatian adalah kemampuan sistem komputer untuk mengenali dan mengevaluasi karakteristik wajah manusia secara otomatis (Maulana, Fadlil and Yudhana, 2024). Penampilan wajah, termasuk persepsi kecantikan, memiliki peranan penting secara sosial dan psikologis dalam interaksi antarindividu, dan telah lama menjadi objek kajian lintas disiplin, mulai dari seni, antropologi, hingga psikologi kognitif.

Penilaian kecantikan wajah secara konvensional banyak mengacu pada prinsip-prinsip proporsi klasik, seperti golden ratio (rasio 1,618), simetri wajah, dan proporsi neoklasik yang telah dikenal sejak zaman Yunani dan Tiongkok kuno sebagai acuan estetika wajah (Sadiq and Abdulazeez, 2024). Namun, penilaian tersebut ketika dilakukan secara manual oleh manusia seringkali bersifat subjektif dan tidak konsisten antarpemilai (Marinescu, 2021). Ketidakkonsistenan ini membuka peluang untuk mengembangkan pendekatan komputasional yang mampu mengukur kesesuaian geometris wajah terhadap prinsip-prinsip proporsi klasik tersebut secara sistematis dan dapat direproduksi. Penilaian kecantikan wajah pada penelitian ini diukur melalui empat aspek kriteria geometris yang diadaptasi dari prinsip-prinsip proporsi klasik dan penelitian facial beauty prediction sebelumnya,

yaitu simetri wajah, proporsi wajah (golden ratio), proporsi hidung, dan proporsi bibir. Simetri wajah dihitung dari empat titik pembandingan kiri-kanan (mata, pipi, rahang, dan alis), karena tingkat asimetri yang rendah antara sisi kiri dan kanan wajah telah lama diasosiasikan dengan persepsi kecantikan yang lebih tinggi pada berbagai studi psikologi persepsi (Iyer *et al.*, 2021).

Proporsi hidung diukur melalui rasio aspek (tinggi terhadap lebar), rasio lebar hidung terhadap lebar wajah, serta simetri sayap hidung, sedangkan proporsi bibir diukur melalui rasio lebar bibir terhadap lebar wajah dan terhadap lebar hidung. Keempat kelompok kriteria tersebut merepresentasikan karakteristik psikofisik persepsi kecantikan seperti keserasian (averageness) dan simetri, yang secara implisit menjadi dasar berbagai pendekatan facial beauty prediction berbasis fitur geometris. Keempat belas fitur turunan dari keempat aspek ini kemudian digabungkan menjadi satu vektor fitur untuk diklasifikasikan menggunakan algoritma K-Nearest Neighbor.

Ekstraksi fitur geometris merupakan salah satu pendekatan yang umum digunakan dalam penilaian estetika wajah berbasis komputer, karena secara implisit merepresentasikan karakteristik psikofisik persepsi kecantikan seperti keserasian (averageness) dan simetri (Wei *et al.*, 2022). Metode ini mengambil informasi geometris dari wajah seperti rasio simetri antara sisi kiri dan kanan wajah (mata, pipi, rahang, alis), rasio tinggi terhadap lebar wajah sebagai representasi golden ratio, proporsi tiga segmen vertikal wajah, serta proporsi hidung dan bibir terhadap lebar wajah. Fitur-fitur geometris ini relatif tidak terpengaruh oleh perubahan pencahayaan dan ekspresi wajah, sehingga lebih robust digunakan sebagai

representasi kuantitatif dari prinsip estetika wajah. Dari sisi algoritma klasifikasi, metode K-Nearest Neighbor (KNN) dipilih pada penelitian ini karena kesederhanaannya, sifatnya yang non-parametrik, serta telah terbukti memberikan performa yang kompetitif pada berbagai penelitian facial beauty prediction sebelumnya.

Berdasarkan uraian tersebut, penelitian ini mengusulkan pengembangan sistem klasifikasi kecantikan wajah wanita ke dalam tiga tingkatan kategori, yaitu Sangat Cantik, Cantik, dan Cukup Cantik, menggunakan metode K-Nearest Neighbor berbasis ekstraksi 12 fitur geometris yang diimplementasikan dengan bahasa pemrograman Python. Sistem memanfaatkan pustaka MediaPipe Face Mesh untuk mendeteksi landmark wajah tanpa memerlukan perangkat keras khusus, kemudian mengekstraksi fitur simetri, golden ratio, proporsi hidung, dan proporsi bibir dari setiap citra. Fitur-fitur tersebut dinormalisasi menggunakan StandardScaler, lalu diklasifikasikan menggunakan KNN berbobot jarak (distance-weighted) dengan pengukuran jarak Euclidean, di mana nilai K dioptimalkan melalui skema 10-fold stratified cross-validation.

Penelitian ini juga perlu menegaskan sejak awal bahwa kecantikan (beauty) adalah konstruksi persepsi yang bersifat subjektif dan dipengaruhi oleh faktor budaya, sehingga tidak terdapat satu definisi kecantikan yang berlaku universal. Sistem yang dikembangkan pada penelitian ini tidak dimaksudkan untuk mengklaim kebenaran mutlak mengenai kecantikan seseorang, melainkan untuk mengukur sejauh mana fitur geometris suatu wajah sesuai dengan prinsip proporsi klasik tertentu yang telah banyak digunakan sebagai proksi kuantitatif dalam

penelitian facial beauty prediction. Label kategori kecantikan pada dataset penelitian ini ditetapkan melalui skema semi-otomatis berbasis skor komposit dari fitur geometris yang sama, kemudian ditinjau ulang secara manual, mengingat keterbatasan sumber daya untuk memperoleh penilaian juri manusia independen dalam jumlah besar. Pendekatan ini merupakan keterbatasan metodologis yang diakui secara eksplisit: performa klasifikasi yang dilaporkan pada penelitian ini merefleksikan konsistensi model terhadap skema pelabelan berbasis proporsi geometris tersebut, bukan terhadap penilaian kecantikan oleh manusia secara independen.

Penelitian ini juga mempertimbangkan preseden kegagalan etis pada sistem penilaian kecantikan berbasis kecerdasan buatan, seperti kasus kontes Beauty.AI pada tahun 2016 yang algoritmenya terbukti bias terhadap warna kulit akibat data latih yang tidak representatif. Sebagai langkah mitigasi, sistem pada penelitian ini dirancang murni untuk kepentingan akademis dalam mendemonstrasikan ekstraksi fitur geometris dan klasifikasi KNN, dan tidak dimaksudkan untuk digunakan dalam pengambilan keputusan nyata yang berdampak pada individu, seperti seleksi, penilaian sosial, atau aplikasi kencan daring. Dataset yang digunakan juga memiliki keterbatasan keberagaman demografis subjek, sehingga generalisasi hasil penelitian ini perlu dilakukan secara hati-hati.

Dengan adanya penelitian ini, sistem yang dibangun diharapkan dapat memberikan pengukuran kuantitatif dan konsisten terhadap kesesuaian wajah dengan prinsip-prinsip proporsi klasik, sebagai alat bantu analisis dan pembelajaran di bidang computer vision, bukan sebagai vonis mutlak atas kecantikan seseorang.

1.2 Pertanyaan Penelitian

1. Bagaimana pustaka MediaPipe Face Mesh dapat dimanfaatkan untuk mendeteksi landmark wajah dan mengekstraksi fitur geometris yang merepresentasikan simetri wajah serta proporsi klasik (golden ratio, proporsi hidung, dan proporsi bibir)?
2. Bagaimana performa algoritma K-Nearest Neighbor berbasis 12 fitur geometris dalam mengklasifikasikan citra wajah wanita ke dalam tiga tingkatan kategori kecantikan, yaitu Sangat Cantik, Cantik, dan Cukup Cantik?
3. Seberapa besar tingkat akurasi, presisi, recall, dan F1-score yang dicapai oleh sistem klasifikasi yang dibangun berdasarkan pengujian 10-fold stratified cross-validation terhadap dataset primer?

1.3 Tujuan Penelitian

1. Mengimplementasikan modul ekstraksi 12 fitur geometris wajah (simetri, golden ratio, proporsi hidung, dan proporsi bibir) menggunakan MediaPipe Face Mesh dan bahasa pemrograman Python.
2. Membangun model klasifikasi K-Nearest Neighbor berbobot jarak (distance-weighted) untuk mengklasifikasikan tingkat kecantikan wajah wanita ke dalam tiga kategori.
3. Mengevaluasi kinerja model klasifikasi menggunakan metrik akurasi, presisi, recall, F1-score, dan confusion matrix melalui skema 10-fold stratified cross-validation.

1.4 Kerangka Kerja Penelitian (*Research Framework*)

Kerangka kerja penelitian ini disusun secara sistematis dan berurutan mengikuti pendekatan eksperimental berbasis pengolahan citra digital, mulai dari tahap pengumpulan data primer hingga evaluasi kinerja model. Gambar 1.1 menyajikan diagram alur tahapan penelitian secara keseluruhan.

Tabel 1. 1 Framework Penelitian

Tahap	Nama Tahap	Aktivitas	Metode/Alat	Output
1	Pengumpulan Data	Pengumpulan dataset citra wajah wanita	Data Primer	Dataset citra wajah (100 citra)
2	Deteksi Landmark Wajah	Deteksi wajah dan 468 titik landmark wajah	MediaPipe Face Mesh	Peta landmark wajah
3	Ekstraksi Fitur & Pelabelan	Perhitungan 12 fitur geometris, normalisasi fitur, pelabelan semi-otomatis + tinjauan manual	StandardScaler, skor komposit	Vektor fitur (12D) dan label kategori kecantikan
4	Klasifikasi KNN & Evaluasi	Training KNN berbobot jarak, optimasi nilai K, evaluasi performa	scikit-learn KNN, 10-fold stratified CV, Euclidean distance	Model KNN terlatih, akurasi, presisi, recall, F1-score, confusion matrix

Tahap pertama dimulai dengan pengumpulan data primer berupa 100 citra wajah wanita yang dikumpulkan langsung oleh peneliti. Tahap kedua melibatkan deteksi wajah dan lebih dari 468 titik landmark menggunakan MediaPipe Face Mesh, yang bekerja langsung pada citra RGB tanpa memerlukan tahap preprocessing grayscale maupun resize eksplisit. Tahap ketiga adalah ekstraksi 12 fitur geometris dari landmark yang terdeteksi, dilanjutkan dengan pelabelan kategori kecantikan secara semi-otomatis berdasarkan skor komposit, serta normalisasi fitur menggunakan StandardScaler. Tahap terakhir adalah pelatihan model KNN berbobot jarak, optimasi nilai K, serta evaluasi kinerja model

menggunakan metrik akurasi, presisi, recall, F1-score, dan confusion matrix melalui skema 10-fold stratified cross-validation. Setiap tahap menghasilkan output yang menjadi input bagi tahap berikutnya, sehingga keseluruhan proses membentuk alur kerja yang runtut dan dapat direproduksi.

