

Perbandingan Algoritma *Clustering K-Means, Gaussian Mixture Model, dan Spectral Clustering* untuk *Facial Emotion Recognition*

A Comparative Study of K-Means, Gaussian Mixture Model, and Spectral Clustering for Facial Emotion Recognition

¹Francesco Adhimas Totti, ²Nina Setiyawati*

^{1,2}Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana

^{1,2}Jl. Diponegoro No.52-60, Salatiga, Kota Salatiga, Jawa Tengah 50711, Indonesia

*e-mail: nina.setiyawati@uksw.edu

(**received:** 26 August 2025, **revised:** 8 September 2025, **accepted:** 9 September 2025)

Abstrak

Pengenalan emosi wajah merupakan salah satu bidang dalam kecerdasan buatan yang banyak diterapkan dalam interaksi manusia dan komputer. Penelitian ini membandingkan tiga algoritma *clustering* K-Means, Gaussian Mixture Model (GMM), dan Spectral Clustering untuk mengelompokkan ekspresi wajah dari dataset FER-2013 yang telah disampling menjadi 1.820 gambar. Fitur wajah diekstraksi menggunakan filter Gabor dan direduksi dimensinya dengan Principal Component Analysis (PCA) guna mempertahankan informasi penting secara efisien. Evaluasi dilakukan menggunakan *Silhouette Score*, *Davies-Bouldin Index*, dan estimasi akurasi clustering. Hasil menunjukkan bahwa Spectral Clustering memberikan performa terbaik dengan *Davies-Bouldin Index* sebesar 0.464 dan akurasi 95.36%, diikuti oleh GMM (*Silhouette Score* 0.526) dan K-Means (*Silhouette Score* 0.523). Selain itu, PCA dengan retensi 80% varian menghasilkan representasi fitur yang efektif dalam bentuk 1D, sehingga hasil *clustering* divisualisasikan secara sederhana namun informatif. Temuan ini menunjukkan bahwa pemilihan algoritma *clustering* perlu disesuaikan dengan kebutuhan akurasi atau efisiensi sistem.

Kata kunci: pengenalan emosi wajah, *clustering*, filter gabor, PCA

Abstract

Facial emotion recognition is one of the key areas of artificial intelligence that has been widely applied in human-computer interaction. This study compares three clustering algorithms—K-Means, Gaussian Mixture Model (GMM), and Spectral Clustering—to group facial expressions from the FER-2013 dataset, which was sampled down to 1,820 images. Facial features were extracted using Gabor filters and reduced in dimensionality using Principal Component Analysis (PCA) to efficiently preserve essential information. The evaluation was conducted using Silhouette Score, Davies-Bouldin Index, and clustering accuracy estimation. The results show that Spectral Clustering achieved the best performance, with a Davies-Bouldin Index of 0.464 and an accuracy of 95.36%, followed by GMM (Silhouette Score 0.526) and K-Means (Silhouette Score 0.523). Furthermore, PCA with 80% variance retention produced an effective 1D feature representation, allowing clustering results to be visualized in a simple yet informative manner. These findings suggest that the choice of clustering algorithm should be aligned with the desired trade-off between system accuracy and efficiency.

Keywords: facial emotion recognition, clustering, gabor filter, PCA

1 Pendahuluan

Pengenalan emosi wajah merupakan bidang penelitian yang berkembang pesat dalam kecerdasan buatan, interaksi manusia-komputer, dan psikologi komputasional. Sistem pengenalan emosi wajah banyak digunakan dalam berbagai aplikasi, seperti deteksi kebohongan, analisis perilaku pengguna, monitoring psikologis, dan keamanan berbasis biometrik. Dengan meningkatnya penggunaan

teknologi berbasis wajah dalam kehidupan sehari-hari, pengelompokan ekspresi wajah menjadi aspek penting dalam memahami pola emosi tanpa perlu anotasi manual dari data.

Salah satu metode yang digunakan dalam pengelompokan ekspresi wajah adalah *unsupervised clustering*, yang dapat mengelompokkan gambar berdasarkan kesamaan pola tanpa memerlukan label eksplisit. Teknik ini sangat bermanfaat pada situasi dengan keterbatasan data berlabel atau untuk studi eksploratif. Penelitian oleh Delazeri et al.[1] menunjukkan bahwa pendekatan *unsupervised*, khususnya melalui *self-taught representation learning* dan *clustering*, efektif dalam mengidentifikasi struktur alami dari ekspresi wajah pada dataset seperti FER-2013. Studi lainnya dari Muss dan Liu [2] memperkenalkan *dynamic intra-class clustering* untuk menangani ketidakseimbangan kelas dalam dataset ekspresi wajah, seperti FER-2013 dan AffectNet, serta berhasil meningkatkan kualitas representasi fitur emosi melalui *clustering* adaptif.

Penelitian ini membandingkan tiga algoritma *clustering* dalam pengelompokan ekspresi wajah, yaitu: 1) K-Means merupakan metode *clustering* non-hierarki yang membagi data ke dalam beberapa *cluster* berdasarkan kemiripan karakteristik, terutama melalui jarak *Euclidean*. Metode ini berusaha meminimalkan variasi antar data dalam satu *cluster*, serta memaksimalkan perbedaan antar *cluster* lain. Dalam implementasinya pada citra ekspresi wajah, metode ini bekerja dengan memilih *centroid* awal secara acak, menghitung jarak setiap titik terhadap *centroid*, lalu mengelompokkan data ke *cluster* terdekat. Proses ini diulang hingga posisi *centroid* stabil. Penelitian menggunakan KDEF (Karolinska Directed Emotional Faces) menunjukkan bahwa K-Means mampu mengenali emosi dasar seperti marah, jijik, dan sedih, namun hasil akurasi tergolong rendah (sekitar 38,89%) jika tidak dioptimasi lebih lanjut. [3] 2) *Gaussian Mixture Model* (GMM) merupakan pendekatan berbasis probabilistik yang mengasumsikan bahwa data merupakan campuran dari beberapa distribusi Gaussian. Berbeda dengan metode seperti K-Means yang menerapkan pemisahan keras antar *cluster*, GMM memungkinkan tumpang tindih antar *cluster* dengan menetapkan probabilitas keanggotaan bagi tiap data terhadap semua klaster yang ada. Pendekatan ini menjadikannya lebih fleksibel dalam menangani distribusi data yang kompleks, termasuk dalam konteks ekspresi wajah yang memiliki karakteristik visual yang saling tumpang tindih. Penelitian oleh Betru et al. [4] menunjukkan bahwa GMM mampu secara efektif mengelompokkan pola-pola lalu lintas berdasarkan trajektori kendaraan dari data UAV, bahkan dalam situasi lalu lintas yang kompleks. Dengan menerapkan algoritma *Expectation-Maximization* (EM), GMM menunjukkan performa tinggi dalam mengenali struktur alami dari data spasial-temporal tanpa memerlukan label, menjadikannya relevan dan aplikatif dalam pengelompokan ekspresi wajah berbasis citra. 3) Spectral Clustering merupakan metode berbasis teori graf yang menggunakan *eigen-decomposition* dari matriks kesamaan untuk membentuk *cluster*. Pendekatan ini sangat efektif untuk data yang tidak terpisah secara linier dan memiliki bentuk distribusi yang kompleks. Dalam konteks ekspresi wajah, Spectral Clustering menunjukkan performa yang sangat kompetitif. Penelitian oleh Shen et al. dalam jurnal “Adaptive Facial Imagery Clustering via Spectral Clustering and Reinforcement Learning” menunjukkan bahwa metode ini mencapai akurasi hingga 99,2% pada database wajah (ORL dan CFP), dan mampu mengelompokkan ekspresi wajah dengan baik meskipun terdapat noise atau variabilitas bentuk wajah. [5]

Agar dapat menentukan metode *clustering* terbaik dalam pengelompokan ekspresi wajah, penelitian ini melakukan evaluasi menggunakan beberapa metrik, yaitu: 1) *Silhouette Score*, digunakan untuk menilai seberapa baik data berada dalam *cluster* yang sesuai dibandingkan dengan *cluster* lainnya, dengan nilai antara -1 hingga 1. Semakin tinggi nilainya, semakin baik pemisahan antar *cluster*. Metrik ini efektif untuk evaluasi *clustering* seperti ditunjukkan oleh Januzaj et al. [6] 2) *Davies-Bouldin Index* (DBI), digunakan untuk mengevaluasi kualitas *clustering* berdasarkan seberapa baik *cluster* dipisahkan dan seberapa kompak masing-masing *cluster*. Semakin rendah nilai DBI, semakin baik hasil pengelompokan. Dalam penelitian ini, DBI digunakan untuk menilai performa pemisahan ekspresi wajah oleh masing-masing algoritma *clustering*. [7] 3) Akurasi *Clustering* dalam penelitian ini dievaluasi dengan membandingkan hasil pengelompokan terhadap label asli (ground truth) menggunakan metrik seperti *V-Measure*, *Adjusted Rand Index* (ARI), dan *Purity score*. Ketiga metrik ini digunakan secara kombinatorial untuk memberikan estimasi akurasi *clustering*. Pendekatan ini didukung oleh Agughasi dan Murali [8], yang menunjukkan efektivitas strategi semi-supervised dengan kombinasi *clustering* dan *voting* dalam menghasilkan label yang akurat, serta mengevaluasi hasil klasifikasi menggunakan akurasi dari model KNN.

Oleh karena itu, penelitian ini difokuskan pada tiga algoritma clustering, yaitu K-Means, Gaussian Mixture Model (GMM), dan Spectral Clustering. Ketiga metode tersebut dipilih karena mewakili pendekatan berbeda dalam proses pengelompokan data, yakni berbasis centroid, probabilistik, dan graf. Dengan membandingkan ketiga algoritma ini, penelitian bertujuan untuk memperoleh gambaran komprehensif mengenai efektivitas masing-masing metode dalam mengelompokkan ekspresi wajah. Selain itu, pemilihan algoritma ini juga didasarkan pada relevansi dan bukti dari penelitian terdahulu yang menunjukkan performa signifikan, di mana Spectral Clustering tercatat mampu memberikan hasil yang sangat kompetitif pada pengelompokan citra wajah.

Tujuan penelitian ini adalah melakukan eksperimental untuk mendapatkan model *clustering* yang paling optimal dalam pengelompokan ekspresi wajah. Adapun model yang dihasilkan pada penelitian ini akan diterapkan untuk digunakan prediksi emosi korban kekerasan berbasis gender (KBG).

2 Tinjauan Literatur

Beberapa penelitian yang relevan dengan pengenalan emosi wajah berbasis *clustering* telah dilakukan sebelumnya. Penelitian pertama berjudul “Fusi Algoritma K-Means dan CNN untuk Klasifikasi Emosi pada Anak” menggunakan metode K-Means untuk mengklasifikasikan emosi anak berdasarkan detak jantung dan konduktivitas kulit, serta menggabungkannya dengan CNN untuk analisis citra wajah. Penelitian ini menunjukkan bahwa K-Means dapat digunakan untuk mengelompokkan emosi berdasarkan pola fisiologis, meskipun memiliki keterbatasan dalam menangani distribusi yang kompleks [9].

Penelitian kedua berjudul “*Facial Emotion Recognition of Online Learners Using a Hybrid Deep learning Model*” menggunakan metode PCA dalam pengenalan emosi wajah pada pembelajaran daring. Hasil penelitian ini menunjukkan bahwa PCA efektif dalam mengurangi dimensi fitur wajah tanpa kehilangan informasi penting, sehingga meningkatkan efisiensi perhitungan dalam proses *clustering* [10].

Penelitian ketiga berjudul “Klasifikasi Emosi Pada Wajah Dengan Menggunakan K-MEANS Clustering dan KDEP” menggunakan K-Means *Clustering* untuk klasifikasi emosi wajah berbasis dataset KDEP. Penelitian ini menunjukkan bahwa K-Means mampu mengelompokkan tiga jenis ekspresi wajah utama, yaitu marah, jijik, dan sedih, tetapi memiliki keterbatasan dalam akurasi, dengan tingkat keberhasilan hanya 38,89%. Hal ini menunjukkan bahwa K-Means kurang optimal dalam menangani ekspresi wajah yang kompleks dan memerlukan kombinasi dengan metode lain untuk meningkatkan akurasi [3].

Penelitian keempat berjudul “Deep Conditional Gaussian Mixture Model for Constrained Clustering” mengevaluasi Deep Conditional Gaussian Mixture Model (DC-GMM) pada dataset UTKFace, yang digunakan untuk *clustering* berdasarkan gender dan etnis. Penelitian ini menemukan bahwa DC-GMM mencapai akurasi 89% dalam *clustering* berdasarkan gender dan 85% dalam *clustering* berdasarkan etnis, yang lebih tinggi dibandingkan metode *unsupervised* seperti *Variational Autoencoder* (VaDE). Hal ini menunjukkan bahwa pendekatan probabilistik dengan GMM dapat bekerja lebih baik dalam *clustering* data wajah dibandingkan metode berbasis *centroid* seperti K-Means [11].

Penelitian kelima berjudul “A Deep Clustering Algorithm based on Gaussian Mixture Model” mengusulkan *Deep Clustering* berbasis *Gaussian Mixture Model* (GMM) yang menggabungkan *Stacked Autoencoder* (SAE) dengan GMM untuk meningkatkan akurasi *clustering*. Eksperimen yang dilakukan pada dataset COIL-20 dan MNIST menunjukkan bahwa kombinasi *deep learning* dan GMM dapat meningkatkan akurasi *clustering* sebesar 6-20% dibandingkan GMM konvensional. Hasil ini menunjukkan bahwa GMM dapat bekerja lebih baik dibandingkan K-Means dalam *clustering* data wajah jika dikombinasikan dengan teknik ekstraksi fitur berbasis *deep learning* [12].

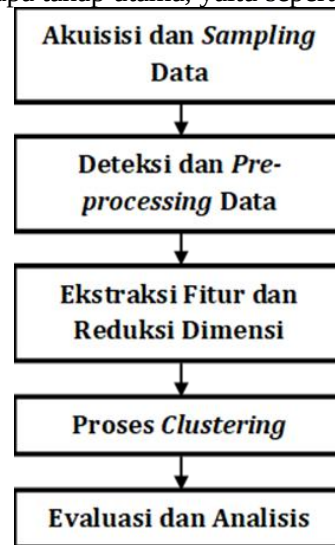
Penelitian keenam berjudul “Adaptive Facial Imagery Clustering via Spectral Clustering and Reinforcement Learning” menggunakan Spectral Clustering yang dikombinasikan dengan Reinforcement Learning (Q-learning) untuk *clustering* wajah. Metode ini diterapkan pada ORL Face Database dan CFP Face Database, dengan hasil yang menunjukkan bahwa Spectral Clustering mencapai akurasi 99.2% dalam *clustering* wajah pada ORL Face Database. Hasil ini menunjukkan

bahwa Spectral Clustering lebih unggul dibandingkan metode berbasis *centroid* seperti K-Means dalam mengelompokkan data wajah yang kompleks [5].

3 Metode Penelitian

Penelitian ini mengusulkan pendekatan pengenalan emosi wajah (*Facial Emotion Recognition/FER*) berbasis kombinasi antara teknik *deep learning* untuk deteksi wajah dan *unsupervised machine learning* untuk pengelompokan ekspresi.

Proses penelitian ini terdiri dari beberapa tahap utama, yaitu seperti terlihat pada Gambar 1:



Gambar 1 Alur penelitian

3.1 Akuisisi dan Sampling Data

Dataset yang digunakan adalah FER-2013 (*Facial Expression Recognition 2013*) yang diperoleh dari Kaggle. Dataset ini berisi gambar wajah dengan berbagai ekspresi seperti marah, jijik, takut, bahagia, netral, sedih dan terkejut sebagaimana terlihat pada Gambar 2.



Gambar 2 Dataset FER-2013

Untuk mengoptimalkan proses komputasi serta menjaga representasi kategori yang seimbang, dilakukan proses sampling acak secara proporsional terhadap data dalam folder test. Dari total 3.589 gambar, dipilih sebanyak 2.000 gambar, sebagaimana distribusinya terlihat pada Tabel 1, menggunakan skrip *Python split_dataset()* yang diimplementasikan dalam file *sampling.py*.

Tabel 1 Distribusi dataset hasil sampling (2000 gambar)

Emosi	Jumlah Gambar
Marah	246
Jijik	38
Takut	298
Bahagia	521
Netral	339
Sedih	342
Terkejut	216

Total	2000
-------	------

Proses ini menyaring data berdasarkan kategori secara merata untuk menghindari dominasi kelas tertentu. Pendekatan ini didukung oleh studi Mejia Escobar et al. [13], yang menunjukkan bahwa reduksi dataset secara terkontrol dapat meningkatkan efisiensi pelatihan dan kinerja model dalam sistem pengenalan ekspresi wajah.

3.2 Deteksi dan *pre-processing* Data

Dalam penelitian ini, pendekatan *deep learning* digunakan untuk membantu proses deteksi dan pemrosesan awal gambar wajah secara otomatis. Model yang digunakan adalah YOLOv8 (*You Only Look Once version 8*), yang merupakan salah satu arsitektur object detection terkini dengan performa tinggi dalam hal kecepatan dan akurasi deteksi.

YOLOv8 dipilih dalam penelitian ini karena memiliki sejumlah keunggulan yang mendukung kebutuhan sistem deteksi wajah. Salah satu keunggulan utamanya adalah efisiensi tinggi, di mana YOLO dirancang untuk mendeteksi objek hanya dalam satu proses inferensi (*single forward pass*), sehingga lebih cepat dibandingkan metode deteksi lainnya. Selain itu, YOLOv8 juga memiliki presisi deteksi yang tinggi, bahkan saat diterapkan pada gambar dengan resolusi rendah seperti yang terdapat dalam dataset FER-2013. Kemampuan ini membuatnya sangat cocok untuk mendeteksi wajah secara akurat dalam kondisi variatif. YOLOv8 juga kompatibel dengan teknologi CUDA dan GPU, dan dalam penelitian ini dijalankan menggunakan GPU NVIDIA GTX 1650, yang secara signifikan mempercepat proses deteksi. Kemudahan integrasi menjadi nilai tambah, karena YOLOv8 dapat digunakan langsung melalui pustaka *Ultralytics* dalam *Python*, sehingga memungkinkan integrasi yang lancar dengan pipeline *pre-processing* dataset secara keseluruhan. Keunggulan-keunggulan ini juga ditunjukkan oleh Yisihak dan Li [14], yang berhasil menerapkan YOLOv8 ke dalam sistem AI *real-time* untuk deteksi wajah, dengan hasil presisi, *recall*, dan *F1-score* yang sangat tinggi bahkan dalam kondisi lingkungan yang kompleks.

Model YOLOv8 tidak digunakan sebagai model klasifikasi emosi dalam penelitian ini, melainkan hanya sebagai alat bantu untuk melokalisasi wajah dalam gambar. Dengan demikian, perannya adalah sebagai tahap awal dari *pre-processing* yang akan memberikan input bersih dan valid ke tahap selanjutnya yaitu ekstraksi fitur dan *clustering*.

Setelah proses pengambilan sampel dataset, dilakukan tahapan *pre-processing* data dengan menggunakan model YOLOv8, khususnya *yolov8l-face.pt*, untuk mendeteksi wajah pada setiap gambar dari total 2000 sampel. Model ini menghasilkan output berupa koordinat *bounding box* dan nilai *confidence* untuk setiap deteksi wajah. Dari hasil deteksi tersebut, hanya gambar yang memiliki satu wajah valid yang dipertahankan. Gambar tanpa wajah atau dengan lebih dari satu wajah secara otomatis diabaikan agar kualitas data tetap terjaga. Wajah yang terdeteksi kemudian dipotong (*cropping*) sesuai dengan area *bounding box*, lalu diubah ukurannya menjadi 48x48 piksel dan dikonversi ke dalam skala abu-abu (*grayscale*) agar sesuai dengan kebutuhan pada tahap ekstraksi fitur menggunakan Gabor filter. Terakhir, gambar-gambar hasil *pre-processing* distandarisasi dalam hal ukuran dan format, sehingga siap digunakan dalam proses ekstraksi fitur dan analisis lebih lanjut.

Dari total 2000 gambar yang digunakan dalam penelitian ini, sebanyak 1971 gambar berhasil terdeteksi memiliki satu wajah yang valid setelah diproses menggunakan YOLOv8. Hal ini menunjukkan tingkat keberhasilan deteksi wajah sebesar 98.55%, yang mencerminkan kinerja deteksi yang sangat baik. Selain itu, proses deteksi berlangsung dengan efisien, dengan rata-rata waktu pemrosesan per gambar sebesar 0.18 detik atau sekitar 176.22 milidetik, menjadikan metode ini sangat ideal untuk digunakan dalam sistem yang memerlukan pemrosesan cepat dan akurat.

Distribusi hasil deteksi wajah dapat dilihat pada Tabel 2, yang memperlihatkan jumlah gambar valid pada tiap kelas emosi. Dari Tabel 2 terlihat bahwa kelas Bahagia memiliki distribusi terbanyak dengan 518 gambar, sedangkan kelas Jijik menjadi kelas dengan jumlah paling sedikit, yaitu 38 gambar. Sementara itu, kelas lainnya seperti Marah (243 gambar), Takut (288 gambar),

Netral (337 gambar), Sedih (333 gambar), dan Terkejut (214 gambar) tetap terdistribusi secara proporsional. Distribusi ini penting karena memastikan setiap kelas emosi tetap terwakili dengan baik setelah proses deteksi wajah.

Tabel 2 Distribusi dataset setelah deteksi YOLO (1971 gambar valid)

Emosi	Jumlah Gambar
Marah	243
Jijik	38
Takut	288
Bahagia	518
Netral	337
Sedih	333
Terkejut	214
Total	1971

Proses ini memberikan dataset wajah yang bersih dan siap untuk tahap ekstraksi fitur, sekaligus memastikan bahwa input yang digunakan untuk *clustering* benar-benar fokus pada satu ekspresi wajah yang valid.

3.3 Ekstraksi Fitur dan Reduksi Dimensi

Setelah wajah berhasil dipotong, dilakukan proses ekstraksi fitur menggunakan Gabor Filter, yang dikenal efektif dalam menangkap informasi tekstur dari citra wajah. Setiap gambar wajah terlebih dahulu dikonversi ke dalam format *grayscale* untuk menjaga konsistensi data. Selanjutnya, diterapkan delapan kernel Gabor yang merupakan kombinasi dari empat orientasi utama (0° , 45° , 90° , dan 135°) serta dua nilai skala sigma (1 dan 3), sehingga menghasilkan variasi pola tekstur. Dari setiap hasil konvolusi, diambil dua parameter statistik, yaitu nilai mean dan *standard deviation*, yang mencerminkan intensitas dan keragaman pola. Dengan demikian, dihasilkan total 16 fitur untuk setiap gambar wajah. Proses ini diimplementasikan dalam fungsi `extract_gabor_features()` pada file `feature_extraction.py`.

Untuk menyederhanakan representasi data dan mengurangi kompleksitas komputasi, dilakukan reduksi dimensi menggunakan *Principal Component Analysis* (PCA). Metode ini bertujuan untuk mempertahankan informasi paling signifikan dari hasil ekstraksi fitur, sekaligus menghilangkan dimensi yang kurang relevan. Dalam penelitian ini, PCA diterapkan dengan tingkat *variance retained* sebesar 88%, sehingga fitur yang dihasilkan tetap menggambarkan karakteristik utama dari ekspresi wajah secara optimal. Hasil transformasi PCA kemudian disimpan dalam format `features.npy`, dan digunakan sebagai input pada tahap pemodelan *clustering*.

Penggunaan PCA dalam penelitian ini didukung oleh studi yang dilakukan oleh Aris Budi et al. [15], yang menunjukkan bahwa teknik PCA mampu mereduksi dimensi citra wajah secara signifikan, sekaligus mempertahankan akurasi identifikasi hingga 80%. Hal ini membuktikan efektivitas PCA dalam menyederhanakan fitur tanpa mengorbankan kualitas informasi yang terkandung di dalamnya.

Dari total 1971 gambar wajah valid hasil deteksi YOLO, hanya 1115 gambar yang lolos tahap seleksi ekstraksi fitur. Gambar-gambar yang tidak memenuhi kriteria—misalnya hasil cropping tidak sesuai, kualitas tekstur rendah, atau gagal diproses pada tahap ekstraksi—dikeluarkan agar kualitas data tetap terjaga. Distribusi akhir dari dataset setelah tahap *pre-processing* ini ditunjukkan pada Tabel 3, yang memperlihatkan jumlah data yang siap digunakan dalam proses *clustering*. Terlihat bahwa kelas Bahagia tetap mendominasi dengan 282 gambar, sedangkan

kelas Jijik memiliki jumlah paling sedikit dengan 23 gambar. Proporsi distribusi antar kelas tetap konsisten dengan hasil sebelumnya, sehingga representasi tiap emosi masih terjaga meskipun terjadi pengurangan jumlah gambar.

Tabel 3 Distribusi dataset setelah *pre-processing* (1115 gambar valid untuk *clustering*)

Emosi	Jumlah Gambar
Marah	152
Jijik	23
Takut	162
Bahagia	282
Netral	204
Sedih	184
Terkejut	108
Total	1115

Proses ini menghasilkan dataset wajah yang lebih bersih, terstandarisasi, dan optimal, sehingga layak digunakan untuk tahap pengelompokan (*clustering*) menggunakan algoritma yang dibandingkan dalam penelitian ini.

3.4 Proses *Clustering*

Setelah fitur wajah berhasil diekstraksi dan direduksi dimensinya menggunakan PCA, proses selanjutnya adalah melakukan pengelompokan ekspresi wajah menggunakan pendekatan *unsupervised learning*, khususnya dengan menerapkan tiga algoritma *clustering*, yaitu K-Means, *Gaussian Mixture Model* (GMM), dan *Spectral Clustering*.

Pemilihan tiga algoritma ini didasarkan pada pertimbangan efisiensi komputasi serta fleksibilitas dalam menangani data citra wajah. K-Means dipilih karena kesederhanaan dan kecepatannya dalam melakukan pengelompokan berbasis *centroid*, sehingga sering dijadikan *baseline* dalam penelitian *clustering*. *Gaussian Mixture Model* (GMM) dipilih karena kemampuannya menangani distribusi data yang kompleks serta memungkinkan overlap antar *cluster*, yang lebih sesuai dengan karakteristik ekspresi wajah yang saling mirip. Sementara itu, *Spectral Clustering* digunakan karena efektivitasnya dalam mengelompokkan data non-linear berbasis *eigen-decomposition* dari *similarity graph*, sehingga mampu menghasilkan pemisahan *cluster* yang lebih baik meskipun dengan biaya komputasi lebih tinggi. Metode *clustering* lain tidak digunakan dalam penelitian ini karena fokus utama adalah membandingkan tiga pendekatan representatif yang mewakili model *centroid-based*, *probabilistic*, dan *graph-based*, yang juga banyak digunakan dalam literatur *clustering* citra wajah.

Eksperimen pertama menggunakan K-Means, algoritma berbasis *centroid* yang populer karena kesederhanaan dan kecepatannya. Parameter yang digunakan adalah $n_clusters = 7$, $n_init = 20$, dan $max_iter = 500$. K-Means mengelompokkan data dengan menghitung jarak *Euclidean* antara setiap titik data dengan *centroid* terdekat, dan memperbarui posisi *centroid* secara iteratif hingga mencapai konvergensi. Penerapan K-Means dalam pengolahan citra wajah juga telah digunakan secara luas, seperti yang ditunjukkan oleh Nuraeni et al. [16] dalam studi segmentasi wajah menggunakan pendekatan *centroid-based clustering*.

Eksperimen kedua menggunakan *Gaussian Mixture Model* (GMM), yaitu algoritma probabilistik yang mengasumsikan bahwa data berasal dari campuran beberapa distribusi Gaussian. Parameter yang digunakan mencakup $n_components = 7$, $covariance_type = 'tied'$, dan $reg_covar = 1e-4$. GMM memiliki fleksibilitas yang lebih tinggi dibanding K-Means karena

<http://sistemasi.ftik.unisi.ac.id>

memperbolehkan adanya tumpang tindih antar kluster dan mampu menangani bentuk distribusi yang tidak selalu sferis. Penerapan GMM untuk *clustering* fitur wajah berdasarkan tekstur juga relevan dengan pendekatan yang dilakukan oleh Prayogi [17], yang menggabungkan fitur tekstur dengan K-Means dalam segmentasi citra visual.

Eksperimen terakhir adalah Spectral Clustering, sebuah metode yang memanfaatkan *eigen decomposition* terhadap *similarity graph* untuk menemukan kluster dalam struktur data *non-linear*. Parameter yang digunakan adalah *n_clusters* = 7, *affinity* = 'rbf', dan *assign_labels* = 'discretize'. Meskipun metode ini memiliki kompleksitas komputasi yang lebih tinggi dibanding dua metode sebelumnya, hasil eksperimen menunjukkan bahwa Spectral Clustering unggul dalam hal akurasi serta pemisahan antar kluster pada data ekspresi wajah yang kompleks.

3.5 Evaluasi dan Analisis

Evaluasi dilakukan untuk mengukur kualitas hasil *clustering* terhadap fitur wajah yang telah dikelompokkan ke dalam tujuh *cluster* emosi. Dalam penelitian ini digunakan tiga metrik utama, yaitu *Silhouette Score*, *Davies-Bouldin Index* (DBI), dan Akurasi *Clustering* (Estimasi). Ketiganya dihitung menggunakan fungsi *evaluate_clustering()* yang diimplementasikan dalam file *evaluation.py*.

Silhouette Score mengukur seberapa baik sebuah titik data berada dalam *cluster* yang sesuai dibandingkan dengan *cluster* terdekat lainnya. Nilai skor berada pada rentang -1 hingga 1, di mana nilai mendekati 1 menunjukkan bahwa data sangat sesuai dengan *clusternya*, nilai mendekati 0 menunjukkan data berada di batas antara dua *cluster*, dan nilai negatif menunjukkan kemungkinan bahwa data salah dikelompokkan. Penerapan *Silhouette Score* dalam evaluasi *clustering*, termasuk pada K-Means dan DBSCAN, telah banyak dilakukan dalam studi-studi lokal seperti oleh Hasan (2024), yang membandingkan metrik ini dengan DBI pada data kinerja karyawan. [18]

Sementara itu, *Davies-Bouldin Index* (DBI) digunakan untuk mengevaluasi pemisahan antar *cluster* dan kepadatan internal dalam setiap *cluster*. Semakin rendah nilai DBI, semakin baik kualitas pemisahan antar *cluster* tersebut. Pendekatan serupa juga digunakan oleh Utami et al. (2023) dalam optimasi jumlah kluster, yang menggunakan DBI bersama metode *Silhouette* dan Elbow untuk menentukan jumlah kluster yang optimal dalam analisis data Kesehatan. [19]

Adapun Akurasi *Clustering* (Estimasi) digunakan apabila tersedia label *ground truth* dari dataset. Evaluasi dilakukan dengan menghitung kombinasi dari beberapa metrik evaluasi supervised, yaitu V-Measure (mengukur harmoni antara *homogeneity* dan *completeness*), *Adjusted Rand Index* (ARI), serta *Purity Score*. Ketiganya digunakan untuk memberikan estimasi akurasi *clustering* terhadap hasil pelabelan asli.

Tujuan utama dari proses evaluasi ini adalah untuk membandingkan performa dari tiga algoritma *clustering* yaitu K-Means, Gaussian Mixture Model (GMM), dan Spectral Clustering dalam mengelompokkan ekspresi wajah secara otomatis dari dataset FER-2013. Seluruh proses ini didukung dengan deteksi wajah otomatis berbasis YOLOv8 sebagai tahap awal *pre-processing* data.

4 Hasil dan Pembahasan

4.1 Statistik Deteksi Wajah

Dataset yang digunakan dalam penelitian ini diperoleh dari FER-2013 (*Facial Expression Recognition* 2013), yang tersedia di Kaggle dan dikembangkan oleh Manas Sambare. Dataset asli memiliki total 3.589 gambar dalam folder test, namun untuk efisiensi pemrosesan, jumlah gambar dikurangi menjadi 2.000 gambar dengan teknik pemilihan secara acak menggunakan kode sampling yang telah dikembangkan.

Dalam tahap deteksi wajah, model YOLOv8 diterapkan untuk mendeteksi wajah pada gambar yang telah disampling. Dari hasil eksperimen, wajah berhasil terdeteksi pada 1.971 dari 2.000 gambar, menghasilkan per-sentase keberhasilan sebesar 98.55%.

Dari segi efisiensi waktu, total waktu yang dibutuhkan untuk deteksi adalah 352.45 detik, dengan rata-rata 0.18 detik per gambar (176.22 ms). Hal ini menegaskan bahwa YOLOv8

mampu melakukan deteksi wajah dengan kecepatan tinggi, menjadikannya cocok untuk aplikasi *real-time*.

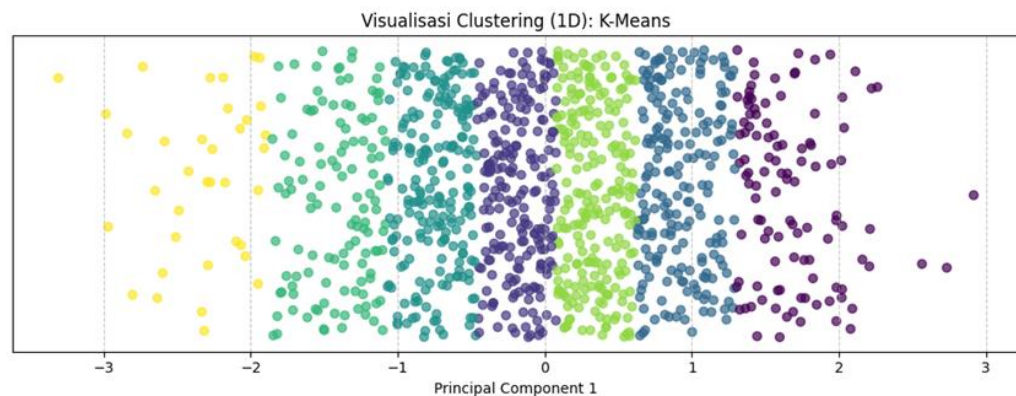
Selanjutnya proses ekstraksi fitur, PCA (*Principal Component Analysis*) digunakan untuk mereduksi dimensi fitur dengan *variance retained* sebesar 88%. Hal ini berarti bahwa meskipun ada sedikit kehilangan informasi, fitur utama tetap terjaga dengan baik untuk proses *clustering*.

4.2 Hasil Evaluasi *Clustering*

Setelah fitur wajah diekstrak dan direduksi dimensinya menggunakan PCA, tiga algoritma *clustering* diterapkan: K-Means, Gaussian Mixture Model (GMM), dan Spectral Clustering.

a. Evaluasi K-Means

Algoritma K-Means pada penelitian ini menunjukkan performa yang cukup baik dengan akurasi sebesar **94.78%**. Nilai *Silhouette Score* yang diperoleh adalah 0.523, yang mengindikasikan bahwa kluster yang terbentuk sudah cukup terpisah dengan baik. Meskipun demikian, nilai *Davies-Bouldin Index* (0.522) masih menunjukkan adanya beberapa kluster yang memiliki kemiripan satu sama lain, sehingga pemisahan antar kluster belum sepenuhnya optimal. Secara keseluruhan, K-Means terbukti mampu mengelompokkan data dengan tingkat konsistensi yang baik, meskipun masih terdapat ruang perbaikan dalam hal kualitas pemisahan kluster.

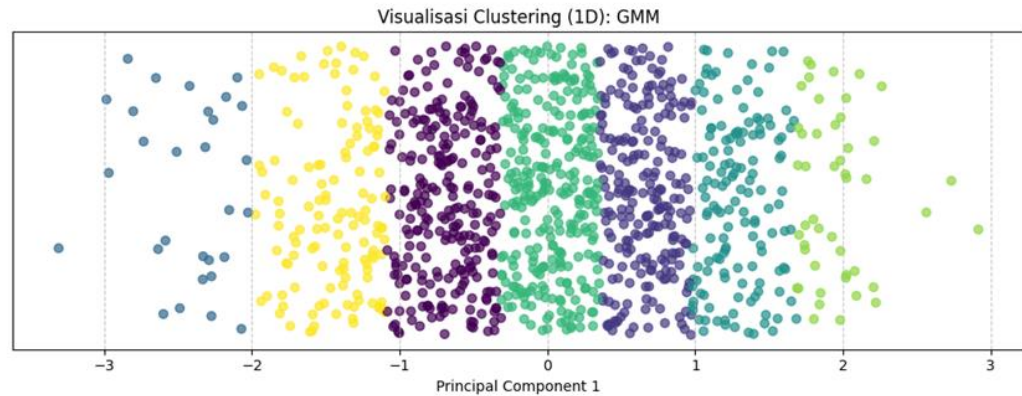


Gambar 3 Hasil *clustering* algoritma K-Means

Gambar 3 memperlihatkan hasil visualisasi *clustering* menggunakan K-Means setelah reduksi dimensi dengan PCA. Titik-titik data tersebar merata sepanjang sumbu *Principal Component 1* dengan warna yang merepresentasikan masing-masing kluster. Terlihat ada sedikit tumpang tindih antar warna, namun distribusinya cukup stabil.

b. Evaluasi Gaussian Mixture Model (GMM)

Algoritma Gaussian Mixture Model (GMM) menunjukkan performa yang sedikit lebih baik dibandingkan K-Means, dengan akurasi sebesar **94.93%**. Nilai *Davies-Bouldin Index* (0.507) yang lebih rendah menandakan bahwa kluster yang terbentuk lebih jelas terpisah, sementara nilai *Silhouette Score* (0.526) yang sedikit lebih tinggi dibandingkan K-Means menunjukkan stabilitas *clustering* yang lebih baik. Hasil ini mengindikasikan bahwa GMM lebih mampu menangkap kompleksitas distribusi data ekspresi wajah, terutama pada kasus di mana terdapat tumpang tindih antar kluster.

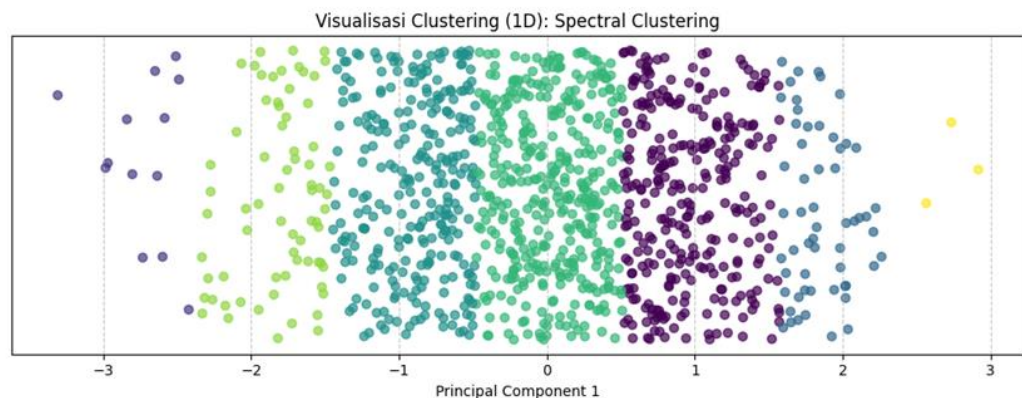


Gambar 4 Hasil clustering algoritma GMM

Gambar 4 adalah hasil visualisasi *clustering* GMM dalam 1 dimensi yang menunjukkan sebaran kluster yang lebih padat dan be-beberapa zona transisi yang lebih jelas dibanding K-Means. GMM memungkinkan adanya overlap antar *cluster*, namun tetap mampu mengidentifikasi batas *cluster* secara baik.

c. Evaluasi Spectral Clustering

Algoritma Spectral Clustering menunjukkan performa terbaik di antara ketiga metode, dengan akurasi tertinggi sebesar 95.36% dan nilai Davies-Bouldin Index (0.464) yang paling rendah, menandakan bahwa kluster yang terbentuk lebih terpisah dibandingkan metode lainnya. Namun demikian, nilai Silhouette Score (0.495) relatif lebih rendah dibandingkan K-Means maupun GMM, yang mengindikasikan bahwa sebagian sampel dalam satu kluster masih memiliki kedekatan dengan kluster lain. Temuan ini menunjukkan bahwa meskipun Spectral Clustering unggul dalam hal pemisahan antar kluster, stabilitas internal klasternya tidak selalu lebih baik daripada metode probabilistik seperti GMM.



Gambar 5 Hasil clustering algoritma spectral clustering

Gambar 5 menampilkan hasil visualisasi *clustering* menggunakan algoritma Spectral Clustering, yang menunjukkan distribusi kluster yang relatif rapi dan padat. Beberapa kluster tampak terpisah dengan jelas, meskipun masih terdapat area tertentu yang menunjukkan tumpang tindih antar data. Kondisi ini dapat terjadi karena sifat Spectral Clustering yang sangat bergantung pada representasi struktur graf dari data, sehingga pada pola fitur yang memiliki kemiripan tinggi, batas antar kluster menjadi kurang tegas.

4.3 Perbandingan Algoritma Clustering

Tabel 4 menunjukkan perbandingan hasil ketiga algoritma *clustering* yang digunakan.

Tabel 4 Hasil metode clustering

Algoritma	Akurasi	Silhouette Score	Davies-Bouldin Index
K-Means	94.78%	0.523	0.522
GMM	94.93%	0.526	0.507
Spectral Clustering	95.36%	0.495	0.464

Dari hasil ini, dapat disimpulkan bahwa:

1. Spectral Clustering menghasilkan akurasi terbaik (95.36%) dengan *Davies-Bouldin Index* paling rendah (0.464), menandakan bahwa *cluster* lebih terpisah dibanding metode lain.
2. GMM memberikan keseimbangan yang baik antara akurasi (94.93%) dan *Davies-Bouldin Index* (0.507), membuatnya pilihan yang baik untuk *clustering* berbasis probabilistik.
3. K-Means tetap menjadi metode tercepat dan memiliki *Silhouette Score* yang cukup tinggi (0.523), tetapi performanya sedikit lebih rendah dibanding metode lainnya.

Dari segi efisiensi, K-Means unggul karena memiliki kompleksitas komputasi yang rendah dan tidak memerlukan operasi matriks berat. Sebaliknya, Spectral Clustering memerlukan *eigen decomposition* terhadap matriks kesamaan (*similarity matrix*), yang berdampak pada tingginya beban komputasi, terutama untuk dataset berukuran besar. Hal ini sesuai dengan temuan pada studi terbaru, yang menyebutkan bahwa *eigen decomposition* memiliki kompleksitas hingga $O(n^3)$, sehingga metode ini cenderung kurang cocok untuk aplikasi real-time tanpa optimasi tambahan [20] [21].

4.4 Analisis dan Implikasi Hasil

Beberapa temuan utama dari eksperimen ini adalah:

1. YOLOv8 menunjukkan performa tinggi dalam deteksi wajah, dengan akurasi 98.55% dan rata-rata waktu deteksi 176.22 ms per gambar, menjadikannya ideal untuk aplikasi *real-time*.
2. Spectral Clustering menunjukkan hasil terbaik dalam akurasi, tetapi memerlukan lebih banyak waktu komputasi dibanding metode lain.
3. GMM memiliki keseimbangan yang baik antara akurasi dan pemisahan kluster, membuatnya alternatif menarik untuk *clustering* berbasis probabilistik.
4. K-Means tetap efisien tetapi sedikit kurang optimal dalam memisahkan kluster dibanding metode lainnya.

Hasil ini menunjukkan bahwa pemilihan metode *clustering* tergantung pada kebutuhan sistem, apakah lebih mengutamakan kecepatan atau akurasi. Jika tujuan utama adalah akurasi tinggi, maka Spectral Clustering adalah pilihan terbaik. Namun, jika kecepatan dan efisiensi lebih diutamakan maka K-Means atau GMM lebih disarankan.

5 Kesimpulan

Penelitian ini mengkaji dan membandingkan tiga algoritma *clustering*—K-Means, Gaussian Mixture Model (GMM), dan Spectral Clustering— untuk pengelompokan ekspresi wajah dari dataset FER-2013 menunjukkan bahwa Spectral Clustering memiliki performa terbaik dengan akurasi 95.36% dan Davies-Bouldin Index (DBI) terendah 0.464. GMM menempati posisi kedua dengan akurasi 94.93% dan Silhouette Score tertinggi 0.526, menunjukkan keseimbangan baik antara akurasi dan pemisahan kluster. K-Means, meskipun efisien secara komputasi, menunjukkan performa sedikit lebih rendah dengan akurasi 94.78%. Proses penelitian ini diawali dengan deteksi wajah menggunakan YOLOv8 dengan tingkat keberhasilan 98.55%, diikuti oleh ekstraksi fitur menggunakan Gabor Filter dan reduksi dimensi menggunakan PCA yang mempertahankan 88% variansi. Kesimpulannya, pemilihan algoritma sebaiknya disesuaikan dengan kebutuhan sistem: Spectral Clustering ideal untuk akurasi tinggi, sementara K-Means dan GMM lebih cocok untuk aplikasi yang memerlukan efisiensi komputasi, seperti aplikasi *real-time*.

Referensi

- [1] B. R. Delazeri, L. L. Vera, J. P. Barddal, A. L. Koerich, and S. B. Alceu De, "Evaluation of Self-Taught Learning-based Representations for Facial Emotion Recognition," *Proc. Int. Jt. Conf. Neural Networks*, Vol. 2022-July, No. ii, 2022, DOI: 10.1109/IJCNN55064.2022.9891956.
- [2] Q. Li et al., "Optimizing Class Imbalance in Facial Expression Recognition using Dynamic Intra-Class Clustering," *Biomimetics*, Vol. 10, No. 5, pp. 1–19, 2025, DOI: 10.3390/biomimetics10050296.
- [3] N. K. Zuhal, D. P. Pamungkas, and R. Wulaningrum, "Klasifikasi Emosi pada Wajah dengan menggunakan K-MEANS Clustering dan KDEP," *Pros. SEMNAS INOTEK (Seminar Nas. Inov. Teknol.)*, Vol. 5, No. 1, pp. 243–248, 2021.
- [4] A. M. Betru, T. M. D. Tran, and W. Ectors, "Automating Composition of Origin-Destination Flows of Intersections based on UAV Data," *Procedia Comput. Sci.*, Vol. 257, pp. 233–240, 2025, DOI: 10.1016/j.procs.2025.03.032.
- [5] C. Shen, L. Qian, and N. Yu, "Adaptive Facial Imagery Clustering via Spectral Clustering and Reinforcement Learning," *Appl. Sci.*, Vol. 11, No. 17, 2021, DOI: 10.3390/app11178051.
- [6] M. Piao Tan and C. A. Floudas, "Determining the Optimal Number of Clusters," *Encycl. Optim.*, Vol. 1, pp. 687–694, 2008, DOI: 10.1007/978-0-387-74759-0_123.
- [7] H. Hendrik, K. Kusriani, and K. Kusnawi, "Optimasi Penentuan Sentroid Awal pada K-Means untuk Meningkatkan Hasil Evaluasi Davies-Bouldin Index," *J. Inform. Teknol. dan Sains*, Vol. 6, No. 1, pp. 52–57, 2024, DOI: 10.51401/jinteks.v6i1.3873.
- [8] V. I. Agughasi and M. Srinivasiah, "Semi-Supervised Labelling of Chest X-Ray Images using Unsupervised Clustering for Ground-Truth Generation," *Appl. Eng. Technol.*, Vol. 2, No. 3, pp. 188–202, 2023, DOI: 10.31763/aet.v2i3.1143.
- [9] F. A. G. Zhafirah, R. Rokhana, R. Sigit, and B. S. B. Dewantara, "Fusi Algoritma K-Means dan CNN untuk Klasifikasi Emosi pada Anak," *Techno.Com*, Vol. 22, No. 3, pp. 622–634, 2023, DOI: 10.33633/tc.v22i3.8667.
- [10] D. Evangeline and A. Parkavi, "Facial Emotion Recognition of Online Learners using a Hybrid Deep Learning Model," *Int. J. Intell. Eng. Syst.*, Vol. 17, No. 6, pp. 735–751, 2024, DOI: 10.22266/ijies2024.1231.56.
- [11] L. Manduchi, K. Chin-Cheong, H. Michel, S. Wellmann, and J. E. Vogt, "Deep Conditional Gaussian Mixture Model for Constrained Clustering," *Adv. Neural Inf. Process. Syst.*, Vol. 14, No. NeurIPS 2021, pp. 11303–11314, 2021.
- [12] X. Lin, X. Yang, and Y. Li, "A Deep Clustering Algorithm based on Gaussian Mixture Model," *J. Phys. Conf. Ser.*, Vol. 1302, No. 3, 2019, DOI: 10.1088/1742-6596/1302/3/032012.
- [13] C. Mejia-Escobar, M. Cazorla, and E. Martinez-Martin, "Towards a Better Performance in Facial Expression Recognition: A Data-Centric Approach," *Comput. Intell. Neurosci.*, Vol. 2023, No. 1, 2023, DOI: 10.1155/2023/1394882.
- [14] H. M. Yisihak and L. Li, "Advanced Face Detection with YOLOv8: Implementation and Integration into AI Modules," Vol. 11, 2024, DOI: 10.4236/oalib.1112474.
- [15] A. Budi, S. Suma'inna, and H. Maulana, "Pengenalan Citra Wajah sebagai Identifier menggunakan Metode Principal Component Analysis (PCA)," *J. Tek. Inform.*, Vol. 9, No. 2, pp. 166–175, 2018, DOI: 10.15408/jti.v9i2.5608.
- [16] F. Nuraeni, H. Susilawati, and Y. Handoko Agustin, "Perbandingan Implementasi Algoritma K-Means++ dan Fuzzy C-Means pada Segmentasi Citra Wajah," *JuTI "Jurnal Teknol. Informasi"*, Vol. 1, No. 2, p. 47, 2023, DOI: 10.26798/juti.v1i2.722.
- [17] Agyztia Premana, Raden Mohamad Herdian Bhakti, and Dimas Prayogi, "Segmentasi K-Means Clustering pada Citra menggunakan Ekstraksi Fitur Warna dan Tekstur," *J. Ilm. Intech Inf. Technol. J. UMUS*, 2020.
- [18] K. Dbscan and Y. Hasan, "Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster," Vol. 06, No. 01, pp. 60–74, 2024.
- [19] I. T. Utami, F. Suryaningrum, and D. Ispriyanti, "K-Means Cluster Count Optimization with Silhouette Index Validation and Davies Bouldin Index (Case Study: Coverage of Pregnant

- Women, Childbirth, and Postpartum Health Services in Indonesia in 2020*),” *BAREKENG J. Ilmu Mat. dan Terap.*, Vol. 17, No. 2, pp. 0707–0716, 2023, DOI: 10.30598/barekengvol17iss2pp0707-0716.
- [20] Y. Wang, “Improving Spectral Clustering using Spectrum-Preserving Node Aggregation,” *Proc. - Int. Conf. Pattern Recognit.*, Vol. 2022-Augus, No. 30, pp. 3063–3068, 2022, DOI: 10.1109/ICPR56361.2022.9956605.
- [21] M. Alshammari and M. Takatsuka, “Approximate Spectral Clustering with Eigenvector Selection and Self-Tuned K,” *Pattern Recognit. Lett.*, Vol. 122, pp. 31–37, 2019, DOI: 10.1016/j.patrec.2019.02.006.