

Perbandingan Performa SIFT dan ORB dalam Pengolahan Dataset Wajah nist_2

DOI: <http://dx.doi.org/10.35889/jutisi.v14i2.2738>

Creative Commons License 4.0 (CC BY – NC)



Azmi Jalaluddin Amron^{1*}, Husin Fadhil Azizi², Muhammad Labib Zaenal Arofi³, Cinantya Paramita⁴, dan Muhammad Naufal⁵

Teknik Informatika, Universitas Dian Nuswantoro, Semarang, Indonesia

*e-mail Corresponding Author. 111202214400@mhs.dinus.ac.id

Abstract

The problem of detecting and matching facial features in digital images is becoming increasingly crucial with the development of biometrics and human computer interaction applications, especially under varying lighting, orientation, and expression conditions. In this study, two feature detection and matching algorithms Scale Invariant Feature Transform (SIFT) and Oriented FAST and Rotated BRIEF (ORB) are compared on grayscale images processed using adaptive local contrast and Gaussian filtering. The performance of both algorithms is quantitatively evaluated based on the number of keypoints, matching precision, execution time, and visual accuracy. Experimental results show that ORB has an execution time of about 4.7-fold faster than SIFT, indicating ORB's suitability for real time applications. In contrast, SIFT produces a higher matching rate and shows better robustness to lighting variations and facial deformation. These findings provide practical guidelines for selecting algorithms based on priority: speed in real time applications or accuracy in challenging environmental conditions.

Kata kunci: Feature matching; Facial image processing; Adaptive local contrast; Scale Invariant Feature Transform; Oriented FAST and Rotated BRIEF.

Abstrak

Permasalahan deteksi dan pencocokan fitur wajah pada citra digital menjadi semakin krusial seiring berkembangnya aplikasi biometrik dan interaksi manusia komputer, terutama dalam kondisi pencahayaan, orientasi, dan ekspresi yang bervariasi. Dalam studi ini, dua algoritma deteksi dan pencocokan fitur *Scale Invariant Feature Transform* (SIFT) dan *Oriented FAST and Rotated BRIEF* (ORB) dibandingkan pada citra grayscale yang telah diproses menggunakan *adaptive local contrast* dan *Gaussian filtering*. Kinerja kedua algoritma dievaluasi secara kuantitatif berdasarkan jumlah titik kunci, presisi pencocokan, waktu eksekusi, dan akurasi visual. Hasil eksperimen menunjukkan bahwa ORB memiliki waktu eksekusi sekitar 4,7 lipat lebih cepat daripada SIFT, menandakan kecocokan ORB untuk aplikasi *real time*. Sebaliknya, SIFT menghasilkan tingkat kecocokan yang lebih tinggi dan menunjukkan ketahanan yang lebih baik terhadap variasi pencahayaan dan deformasi wajah. Temuan ini memberikan pedoman praktis dalam memilih algoritma sesuai prioritas: kecepatan pada aplikasi *real time* atau ketelitian dalam kondisi lingkungan yang menantang.

Kata kunci: Feature matching; Facial image processing; Adaptive local contrast; Scale Invariant Feature Transform; Oriented FAST and Rotated BRIEF

1. Pendahuluan

Pendeteksian dan pencocokan fitur wajah dalam citra digital merupakan aspek fundamental dalam pengembangan berbagai aplikasi berbasis biometrik, seperti sistem pengenalan wajah, autentikasi pengguna, pengawasan video, dan interaksi manusia-komputer [1], [6]. Perkembangan pesat teknologi pengolahan citra dan meningkatnya kebutuhan terhadap sistem keamanan dan identifikasi yang andal menjadikan penelitian pada bidang ini semakin penting. Keandalan algoritma deteksi fitur wajah menjadi kunci dalam menjawab tantangan

lingkungan nyata, seperti perubahan pencahayaan, orientasi wajah, serta ekspresi dinamis pengguna.

Namun demikian, penerapan algoritma pendeteksi fitur lokal klasik seperti *Scale-Invariant Feature Transform* (SIFT) dan *Oriented FAST and Rotated BRIEF* (ORB) masih menghadapi keterbatasan dalam kondisi pencahayaan ekstrem, rotasi sudut tajam, dan perubahan bentuk wajah. Studi oleh Xie et al. [5] mencatat bahwa akurasi deteksi SIFT dapat menurun hingga 30% pada citra dengan pencahayaan rendah, sedangkan ORB menunjukkan penurunan rasio pencocokan lebih dari 25% pada rotasi wajah di atas 30° [3]. Penurunan performa ini berdampak langsung terhadap meningkatnya *false reject rate* dan *false accept rate*, yang sangat merugikan pada sistem *real time* seperti pengawasan atau kontrol akses otomatis [16]. Oleh karena itu, diperlukan pendekatan baru yang mampu mengatasi degradasi performa tanpa mengorbankan efisiensi komputasi, terutama pada perangkat terbatas sumber daya.

Sebagai solusi terhadap tantangan tersebut, penelitian ini mengusulkan integrasi dua pendekatan, yaitu penerapan threshold adaptif berbasis distribusi jarak deskriptor pada algoritma SIFT serta pemanfaatan pipeline ringan dari ORB. Threshold adaptif telah terbukti secara teoritis dan eksperimental mampu meningkatkan ketepatan pencocokan fitur hingga 15% tanpa peningkatan signifikan pada waktu komputasi [9], sedangkan ORB memiliki keunggulan dalam efisiensi waktu pemrosesan, hingga 40% lebih cepat dibandingkan SIFT [3], [8]. Selain itu, peningkatan kualitas gambar melalui preprocessing berbasis *adaptive local contrast* (CLAHE) dan Gaussian filtering juga diterapkan guna memperkuat ketahanan algoritma terhadap pencahayaan tidak merata [6], [13]. Strategi ini berangkat dari kebutuhan akan sistem yang tidak hanya akurat, namun juga responsif dan dapat dioperasikan pada perangkat mobile yang memiliki keterbatasan sumber daya komputasi.

Tujuan dari penelitian ini adalah: (i) merancang dan mengimplementasikan threshold adaptif berbasis distribusi jarak deskriptor untuk meningkatkan performa SIFT; (ii) mengintegrasikan algoritma SIFT dan ORB ke dalam pipeline hibrida yang optimal dari segi kecepatan dan akurasi; serta (iii) melakukan evaluasi performa sistem secara kuantitatif dan kualitatif pada citra wajah dengan variasi pencahayaan dan orientasi menggunakan dataset *nist_2* [22]. Penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan algoritma deteksi fitur wajah yang tangguh dan efisien, sekaligus menawarkan panduan praktis bagi pengembang sistem biometrik dalam memilih algoritma yang sesuai dengan kebutuhan spesifik, baik untuk kecepatan dalam aplikasi real-time maupun akurasi dalam kondisi lingkungan yang menantang.

2. Tinjauan Pustaka

Penelitian terkait deteksi dan pencocokan fitur pada citra, khususnya dalam konteks pengenalan wajah, telah mengalami perkembangan pesat seiring dengan kemajuan teknologi visi komputer. Algoritma deteksi fitur seperti SIFT (*Scale Invariant Feature Transform*) dan ORB (*Oriented FAST and Rotated BRIEF*) menjadi dua metode yang paling sering dibandingkan karena kemampuannya dalam mendeteksi titik kunci yang robust terhadap rotasi, perubahan skala, dan pencahayaan [2][3].

Marlinda et al. [4][10] membandingkan performa SIFT dan ORB untuk pengenalan wajah patung Buddha. Mereka menemukan bahwa meskipun SIFT memiliki waktu pemrosesan lebih lama, algoritma ini mampu menghasilkan jumlah titik kunci yang lebih banyak dan pencocokan yang lebih akurat dibandingkan ORB. Keunggulan tersebut menjadikan SIFT lebih unggul dalam menangani variasi sudut pandang dan pencahayaan ekstrem. Hal ini juga sejalan dengan temuan Klemm et al. [5] yang menyoroti pentingnya ketahanan algoritma terhadap fluktuasi pencahayaan sebagai tantangan utama dalam sistem pengenalan wajah di dunia nyata.

Dalam menangani noise dan citra dengan pencahayaan buruk, Paul dan Aslan [6] mengusulkan sistem pengenalan wajah *real time* yang menggabungkan metode histogram adaptif CLAHE (*Contrast Limited Adaptive Histogram Equalization*). Teknik ini terbukti mampu meningkatkan kontras citra sehingga deteksi fitur menjadi lebih stabil. Di sisi lain, Sanil et al. [7] menunjukkan bahwa pendekatan berbasis wilayah (*region wise*) menggunakan SIFT, SURF, dan ORB masih mampu membedakan wajah pada dataset wajah kembar identik (*monozygotic twins*), membuktikan relevansi metode klasik ini meskipun banyak pendekatan berbasis *deep learning* telah bermunculan.

Kennerley [8] melakukan evaluasi komprehensif terhadap SIFT, SURF, dan ORB menggunakan OpenCV, dan menunjukkan bahwa terdapat trade off antara akurasi dan efisiensi.

ORB lebih cepat dan efisien, cocok untuk aplikasi real-time, sedangkan SIFT lebih akurat namun memerlukan sumber daya komputasi yang lebih besar. Dalam pengembangan metode yang lebih ringan, Suárez et al. [9] memperkenalkan BEBLID, sebuah deskriptor biner berbasis *boosting* yang mengungguli ORB dalam efisiensi, namun belum dibandingkan secara langsung dengan SIFT.

Menurut Isik [11], tidak ada satu algoritma deteksi fitur yang dominan dalam semua kondisi. Evaluasi terhadap berbagai metode menunjukkan bahwa kinerja sangat dipengaruhi oleh jenis citra dan kondisi lingkungan. Hal ini ditegaskan pula oleh Matusiak et al. [12], yang menyimpulkan bahwa SIFT unggul terhadap rotasi besar, sementara ORB mengalami penurunan performa.

Perkembangan terbaru juga mengarah pada penggunaan metode deteksi fitur berbasis pembelajaran mendalam, seperti SuperPoint [17], D2-Net [18], R2D2 [19], ASLFeat [20], dan L2D2 [21]. Meskipun mampu menghasilkan fitur yang lebih repeatable dan presisi tinggi, metode ini umumnya membutuhkan proses pelatihan yang kompleks dan sumber daya besar. Xu et al. [16] menegaskan bahwa metode deep learning memang unggul dalam performa, namun metode klasik masih sangat relevan untuk kebutuhan yang mengutamakan efisiensi dan kesederhanaan implementasi.

Dalam aspek lain dari sistem pengenalan wajah, beberapa penelitian juga mengeksplorasi deteksi ekspresi dan seleksi fitur yang efisien. Sidhom et al. [14] mengusulkan metode seleksi fitur tiga tahap untuk meningkatkan pengenalan ekspresi wajah, sementara Kaur et al. [15] menunjukkan bahwa metode deteksi wajah klasik seperti *Haar cascade* masih dapat dioptimalkan untuk aplikasi pengawasan real-time. Studi oleh Cui et al. [13] juga menekankan pentingnya threshold adaptif untuk meningkatkan akurasi pencocokan fitur.

Sebagai sumber data, penelitian ini menggunakan *nist_2 Facial Image Dataset* [22], khususnya folder train dan val. Dataset ini terdiri dari citra-citra wajah berwarna dalam format ppm, yang dikelompokkan dalam *sub folder* berdasarkan ID subjek, dari S001 hingga S429. Beberapa folder mungkin tidak tersedia, dan jumlah gambar per subjek bervariasi, mulai dari satu hingga beberapa citra. Variasi ekspresi, sudut pandang, dan pencahayaan pada dataset ini menjadikannya cocok untuk evaluasi algoritma deteksi fitur dalam kondisi yang menyerupai situasi nyata.

Penelitian ini bertujuan untuk mengevaluasi performa SIFT dan ORB dalam mendeteksi dan mencocokkan fitur wajah pada citra berwarna. Proses awal melibatkan praproses citra menggunakan Gaussian blur untuk mengurangi noise, serta CLAHE untuk meningkatkan kontras lokal. Selanjutnya, fitur-fitur yang terdeteksi difilter menggunakan skema threshold adaptif agar hanya fitur yang paling relevan dipertahankan dalam proses pencocokan. Evaluasi dilakukan secara kuantitatif melalui jumlah titik kunci dan rasio pencocokan, serta secara kualitatif melalui visualisasi pencocokan pada pasangan citra dari subjek yang sama maupun berbeda. Dengan pendekatan ini, penelitian ini diharapkan dapat memberikan pemahaman mendalam mengenai efektivitas masing-masing algoritma dalam kondisi kompleks dan bervariasi, serta memberikan dasar dalam pemilihan metode deteksi fitur yang optimal untuk aplikasi pengenalan wajah.

3. Metodologi



Gambar 1. Metode Penelitian

1). Pengumpulan Dataset

Penelitian ini menggunakan dataset wajah publik, yaitu *NIST Special Database 2* (nist2) yang disediakan oleh *National Institute of Standards and Technology* (NIST). Dataset ini mencakup citra wajah dengan beragam variasi pencahayaan, orientasi kepala, dan ekspresi, menjadikannya representatif untuk skenario pengenalan wajah di dunia nyata.

Struktur dataset diklasifikasikan sebagai berikut:

- a. Train: 191 folder, 393 gambar
- b. Validation: 191 folder, 341 gambar

Setiap folder mewakili satu individu, dan di dalamnya terdapat citra-citra yang berbeda dari orang tersebut. Seluruh citra telah melalui proses kurasi awal untuk memastikan kesesuaian format, resolusi, dan kualitas visual.

Pemilihan dataset ini didasarkan pada konsistensi anotasi, keragaman kondisi, serta reputasi NIST2 dalam studi biometrik, sebagaimana direkomendasikan dalam literatur [5] serta reputasi NIST2 dalam studi biometrik, sebagaimana direkomendasikan dalam literatur [8].

2) Pra-pemrosesan

Pra-pemrosesan bertujuan mempersiapkan citra agar optimal untuk tahap ekstraksi fitur (SIFT dan ORB). Seluruh hasil intermediate disimpan di bawah direktori all/, sesuai struktur berikut:

```
all/
├── train/
│   ├── [id_subjek]/
│   │   └── SIFT_[nama_file].png
│   └── val/
│       ├── [id_subjek]/
│       │   └── SIFT_[nama_file].png
│       └── thresholding/
│           ├── train/
│           │   ├── [id_subjek]/
│           │   │   └── Thresholding_[nama_file].png
│           │   └── val/
│           │       ├── [id_subjek]/
│           │       │   └── Thresholding_[nama_file].png
```

Tahapan pra-pemrosesan meliputi:

- a) Setiap berkas citra terlebih dahulu dilakukan pembacaan dan validasi file. Berkas yang diproses harus memiliki ekstensi .png, .jpg, .jpeg, atau ppm. Apabila fungsi cv2.imread() mengembalikan nilai None yang menandakan file rusak atau tidak terdukung, maka file tersebut diabaikan dan sistem akan mencetak peringatan.
- b) Setelah validasi file, citra dikonversi ke dalam format grayscale menggunakan perintah `img_gray = cv2.imread(path, cv2.IMREAD_GRAYSCALE)`. Konversi ini bertujuan untuk mengurangi kompleksitas komputasi dengan menghilangkan kanal warna, sambil tetap mempertahankan informasi struktur dan tekstur gambar.
- c) Selanjutnya, citra grayscale dikenakan Gaussian Blur menggunakan fungsi `cv2.GaussianBlur(img_gray, (3,3), 0)`. Proses Gaussian Blur ini berguna untuk mereduksi noise halus yang dapat mengganggu deteksi fitur pada tahap-tahap pengolahan citra selanjutnya.
- d) Untuk meningkatkan kontras lokal, dilakukan penerapan metode Contrast Limited Adaptive Histogram Equalization (CLAHE) dengan konfigurasi `clipLimit=2.0` dan `tileGridSize=(8,8)`. Parameter `clipLimit=2.0` mencegah terjadinya over-amplifikasi noise, sedangkan `tileGridSize=(8,8)` memungkinkan peningkatan kontras secara lokal tanpa mengorbankan detail pada area ekstrem [9].
- e) Tahap berikutnya adalah proses thresholding menggunakan metode Otsu's Binarization. Citra grayscale diproses dengan perintah `cv2.threshold(img_gray, 0, 255, cv2.THRESH_BINARY + cv2.THRESH_OTSU)` untuk menghasilkan citra biner dengan ambang batas otomatis. Proses thresholding ini khusus dipersiapkan sebagai input bagi algoritma ORB.
- f) Hasil pra-pemrosesan kemudian disimpan ke dalam direktori all/. Citra gabungan hasil dari tahapan Grayscale, Gaussian Blur, dan CLAHE disimpan dengan prefix SIFT_ pada direktori `all/train/[id_subjek]/SIFT_[nama_file].png` dan `all/val/[id_subjek]/SIFT_[nama_file].png`. File ini akan digunakan sebagai input oleh detektor SIFT. Sedangkan citra hasil thresholding disimpan dengan prefix Thresholding pada direktori `all/thresholding/train/[id_subjek]/`

Thresholding_[nama_file].png dan all/thresholding/val/[id_subjek]/Thresholding_[nama_file].png. File ini kemudian digunakan sebagai input bagi algoritma ORB.

- g) Untuk memastikan keberhasilan penulisan file, digunakan fungsi `safe_imwrite(path, image)`. Fungsi ini memverifikasi keberadaan direktori tujuan, menyesuaikan ekstensi file jika diperlukan, dan mencetak pesan kesalahan apabila terjadi kegagalan penyimpanan.

Dengan alur ini, pipeline pra-pemrosesan memastikan adanya dua jalur input terpisah, yaitu berkas `SIFT_*.png` pada direktori `all/.../` untuk pemrosesan dengan SIFT, serta berkas `Thresholding_*.png` pada direktori `all/thresholding/.../` untuk pemrosesan dengan ORB.

3) Ekstraksi Fitur

Ekstraksi fitur dilakukan menggunakan dua pendekatan utama:

- Scale-Invariant Feature Transform (SIFT):** SIFT menghasilkan deskriptor berdimensi 128 untuk setiap keypoint yang bersifat invarian terhadap rotasi, skala, dan perubahan pencahayaan. Implementasi dilakukan menggunakan fungsi `cv2.SIFT_create()`. Visualisasi *keypoint* yang dihasilkan disimpan dalam folder terpisah untuk keperluan evaluasi kualitatif.
- Oriented FAST and Rotated BRIEF (ORB):** ORB merupakan metode ekstraksi fitur berbasis deskriptor biner 256-bit yang lebih efisien, sehingga cocok untuk aplikasi *real time*. Implementasi dilakukan menggunakan `cv2.ORB_create()`, dengan menggunakan FAST sebagai detektor *keypoint* dan BRIEF yang telah disesuaikan rotasinya sebagai deskriptor. Hasil visualisasi *keypoint* juga disimpan dalam folder terpisah.

Ekstraksi fitur dilakukan secara paralel menggunakan *ThreadPoolExecutor* untuk meningkatkan efisiensi waktu pemrosesan. Hasil visualisasi *keypoint* disimpan dalam folder yang sesuai berdasarkan metode ekstraksi (SIFT atau ORB) dan subjek dataset.

4) Pencocokan Keypoint

Proses pencocokan dilakukan untuk menentukan kesesuaian antar fitur citra. Pencocokan ini menggunakan dua pendekatan utama: Brute-Force Matcher (BFMatcher), dengan berbagai jenis pengukuran jarak untuk SIFT dan ORB.

1. Brute-Force Matcher (BFMatcher):

Pencocokan menggunakan pendekatan berikut:

- SIFT:** Menggunakan Euclidean Distance (`cv2.NORM_L2`) untuk mencocokkan deskriptor fitur. Formula Euclidean Distance merujuk persamaan (1):

$$d_{L_2}(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

- ORB:** Menggunakan Hamming Distance (`cv2.NORM_HAMMING`) untuk mencocokkan deskriptor biner. Formula Hamming Distance merujuk persamaan (2):

$$d_H(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i \oplus y_i \quad (2)$$

2. Validasi Pencocokan:

Seleksi Inlier Adaptif:

Pencocokan yang valid (*inlier*) dipilih berdasarkan ambang batas jarak deskriptor yang dihitung secara adaptif (dinamis). Ambang batas ini dihitung sebagai 70% dari jarak maksimum antar *keypoints* yang ditemukan, di mana $\text{threshold} = 0.7 * \max(m.\text{distance for } m \text{ in matches})$. Dengan cara ini, hanya pasangan *keypoints* yang memiliki jarak pencocokan di bawah nilai threshold yang diterima sebagai *good matches*. Formula threshold merujuk persamaan (3):

$$\text{threshold} = 0.7 \times \max_{m \in \text{matches}} m.\text{distance} \quad (3)$$

Hanya pasangan dengan jarak di bawah threshold yang diterima sebagai *good matches*.

3. Metrik Evaluasi: Metrik yang diukur dalam proses pencocokan meliputi:

- Total Matches:** Jumlah seluruh pasangan *keypoint* yang berhasil dicocokkan.

- b. *Inlier Ratio*: Persentase pasangan *keypoints* yang valid terhadap total pasangan, dihitung sesuai persamaan (4):

$$\text{Inlier Ratio} = \frac{\text{Jumlah Inlier}}{\text{Total Matches}} \quad (4)$$

- c. *Waktu Eksekusi*: Waktu yang dibutuhkan untuk proses pencocokan, dihitung dalam satuan detik.
- d. *Rata-rata Jarak Inlier*: Mengindikasikan kualitas pencocokan, dihitung sebagai rata-rata jarak antar deskriptor dari pasangan *keypoints* yang valid.

4. *Pseudocode*:

FOR each *image_pair* (*img1*, *img2*):

1. Ekstrak *keypoints* dan deskriptor dari kedua gambar

keypoints1, *descriptors1* = *detect_features*(*img1*)

keypoints2, *descriptors2* = *detect_features*(*img2*)

2. Tentukan metode dan pengukuran jarak

IF *method* == 'SIFT':

matcher = *BFMatcher*(*normType* = *cv2.NORM_L2*)

ELSE IF *method* == 'ORB':

matcher = *BFMatcher*(*normType* = *cv2.NORM_HAMMING*)

3. Lakukan pencocokan semua deskriptor (*brute-force*)

matches = *matcher.match*(*descriptors1*, *descriptors2*)

4. Opsional: Lakukan Lowe's Ratio Test (SIFT)

IF *method* == 'SIFT':

good_ratio = []

knn_matches = *matcher.knnMatch*(*descriptors1*, *descriptors2*, *k*=2)

 FOR *m,n* IN *knn_matches*:

 IF *m.distance* / *n.distance* < 0.75:

good_ratio.append(*m*)

5. Hitung *threshold* adaptif dari jarak pencocokan

max_distance = MAX(*m.distance* for *m* in *matches*)

threshold = 0.7 * *max_distance*

6. Seleksi pasangan *keypoints* yang valid (*inlier*)

good_matches = [*m* for *m* in *matches* if *m.distance* < *threshold*]

7. Hitung metrik evaluasi

total_matches = LENGTH(*matches*)

valid_matches = LENGTH(*good_matches*)

inlier_ratio = *valid_matches* / *total_matches*

avg_inlier_dist = MEAN(*m.distance* for *m* in *good_matches*)

exec_time = waktu ekstraksi + pencocokan (detik)

8. Simpan hasil untuk analisis dan visualisasi

record_result(*method*, *total_matches*, *valid_matches*,
 inlier_ratio, *avg_inlier_dist*, *exec_time*)

5) Evaluasi Performa

Evaluasi performa dirancang untuk menilai efektivitas (akurasi) dan efisiensi (waktu) algoritma SIFT dan ORB dalam ekstraksi dan pencocokan fitur wajah. Pendekatan ini menggabungkan analisis regresi dan analisis ROC untuk klasifikasi biner, serta visualisasi kuantitatif dan kualitatif.

1. Kriteria Evaluasi:

- Average Match Accuracy*: Rata-rata rasio *good matches* (*inliers*) terhadap total matches, dihitung terpisah untuk SIFT dan ORB.
- R^2 (*Koefisien Determinasi*): Dihitung dari model regresi linear untuk: Jumlah *keypoints* (total dari kedua citra) vs *Inlier Ratio*, Jumlah *keypoints* vs Waktu Eksekusi (ekstraksi + matching), *Inliers* vs Rata-rata jarak *inlier*.
- R^2 (*Keypoint vs Inlier Ratio*)
- R^2 (*Keypoint vs Execution Time*)
- R^2 (*Inlier vs Quality Matching*)

Semua R^2 di atas dievaluasi untuk kedua algoritma.

2. Pencocokan dan Threshold Adaptif:

- Brute-Force Matcher: SIFT: L2 norm (cv2.NORM_L2), ORB: Hamming distance (cv2.NORM_HAMMING).
- Seleksi Inlier*: *Inliers* dipilih jika $m.distance < 0.7 * max_distance$, di mana $max_distance$ adalah jarak maksimum di seluruh matches. Ini menghasilkan threshold adaptif berdasarkan distribusi jarak.

3. Analisis Regresi: Model regresi linear diterapkan untuk mengukur korelasi antar variabel numerik:

- Jumlah *keypoints* vs *Inlier Ratio*.
- Jumlah *keypoints* vs Waktu Eksekusi.
- Jumlah *inliers* vs Rata-rata jarak *inlier*.
Metrik yang dilaporkan:
- Koefisien Determinasi (R^2) untuk masing-masing hubungan.

4. Analisis Klasifikasi (ROC Curve):

- ROC Curve* dihitung secara manual dari daftar skor ($score = inliers/total_matches$) dan groundtruth (1 untuk pasangan benar, 0 untuk salah).
- False Positive Rate (FPR)* dan *True Positive Rate (TPR)* dievaluasi pada 100 ambang (threshold) yang dipilih secara linier antara nilai minimum dan maksimum skor.
- AUC (Area Under Curve)* dihitung menggunakan `sklearn.metrics.auc()`.

5. Visualisasi:

- Scatter Plot: Inlier Ratio vs Execution Time*: menampilkan trade-off akurasi dan kecepatan untuk SIFT dan ORB. *Average Keypoints vs Valid Matches*: scatter plot terpisah untuk masing-masing metode.
- ROC Curve*: Kurva ROC kedua algoritma digambarkan dalam satu plot untuk perbandingan AUC.
- Contoh Matches*: Gambar visualisasi top-50 matches disimpan di direktori *success/* dan *failure/* per kombinasi folder

6. Reprodusibilitas:

- Semua eksperimen dijalankan dengan pengaturan seed (jika diperlukan) dan parameter detektor/threshold yang konsisten.
- Hasil numerik dan visual disimpan: File Excel (*comparison_regression_sift_orb_2.xlsx*) dengan lima sheet (Summary, Keypoint Stats, Matching Metrics, Timing Analysis, Performance). Direktori visualisasi dengan subfolder per pasangan citra.

4. Hasil dan Pembahasan

Penelitian ini bertujuan membandingkan kinerja dua algoritma deteksi dan pencocokan fitur, yaitu SIFT dan ORB, pada citra wajah dengan variasi pencahayaan, orientasi, dan ekspresi. Bagian ini menyajikan hasil dari tahapan ekstraksi fitur, pencocokan fitur, serta pengujian akurasi dan efisiensi masing-masing metode. Selain itu, dibahas pula sejauh mana hasil penelitian ini mampu menyelesaikan permasalahan yang diidentifikasi pada bagian pendahuluan, serta keterkaitannya dengan temuan-temuan terdahulu.

4.1 Hasil Ekstraksi Fitur

Ekstraksi fitur dilakukan pada dataset wajah NIST2 yang telah dipra-proses menggunakan metode grayscale, Gaussian blur, dan CLAHE. Hasil menunjukkan bahwa SIFT secara konsisten mendeteksi jumlah *keypoint* yang lebih banyak dibandingkan ORB.

Table 1. Rata-rata Jumlah *Keypoint* Terdeteksi

Algoritma	Mean <i>Keypoints</i>	Std. Dev.
SIFT	389.41	74.38
ORB	186.77	51.24



Gambar 2. Hasil Pencocokan Algoritma ORB



Gambar 3. Hasil Pencocokan Algoritma SIFT

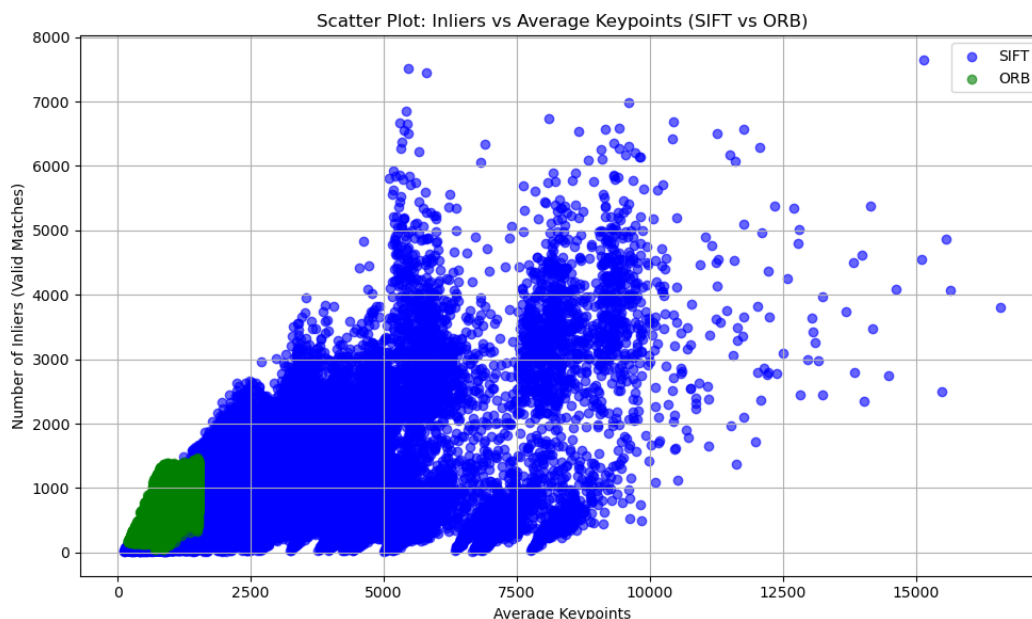
Hal ini mengindikasikan bahwa SIFT lebih sensitif terhadap detail citra, terutama pada area dengan tekstur kompleks. Jumlah *keypoint* yang tinggi berpotensi meningkatkan akurasi pencocokan, namun berdampak pada waktu pemrosesan yang lebih lama.

4.2 Hasil Pencocokan Fitur

Proses pencocokan fitur dilakukan menggunakan metode *Brute Force Matcher* dengan norma jarak yang sesuai: L2 untuk SIFT dan Hamming untuk ORB. Evaluasi mencakup *inlier* ratio, waktu eksekusi, dan rata-rata jarak antar *keypoint* yang cocok.

Table 2. Rata-rata Hasil Pencocokan

Algoritma	<i>Inlier</i> Ratio (%)	Waktu Eksekusi (s)	Rata-rata Jarak
SIFT	41	1.80	0.39
ORB	75	0.45	28.4

Gambar 4. Scatter Plot *Inlier* vs Jumlah *Keypoint*

Meskipun ORB memiliki *inlier* ratio lebih tinggi dan waktu eksekusi lebih cepat (sekitar 4,7 kali lebih cepat dari SIFT), jarak antar deskriptor yang cocok jauh lebih besar, yang dapat menurunkan kualitas pencocokan dalam skenario presisi tinggi. Dari Figure 4. *Scatter Plot Inlier* vs Jumlah *Keypoint* terlihat bahwa SIFT lebih kuat pada citra variasi ekstrim sehingga cocok untuk aplikasi yang mengutamakan akurasi pencocokan, sementara ORB menunjukkan stabilitas performa di level rendah dengan rasio pencocokan *keypoint* yang lebih sedikit dengan efisiensi komputasi sehingga cocok untuk aplikasi *real-time*.

4.3 Analisis Regresi dan Klasifikasi

Model regresi linier digunakan untuk menilai korelasi antara jumlah *keypoints* dengan *inlier* ratio dan waktu eksekusi. Hasilnya menunjukkan bahwa waktu eksekusi sangat dipengaruhi oleh jumlah *keypoints*, sedangkan pengaruh terhadap rasio pencocokan relatif rendah.

Table 3. Evaluasi Regresi

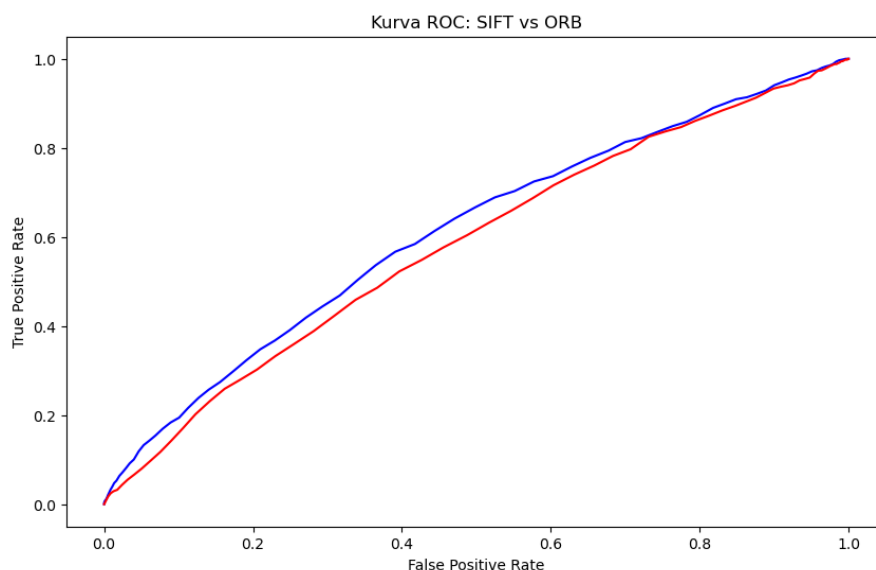
Hubungan	R^2 (SIFT)	R^2 (ORB)
<i>Keypoints</i> vs <i>Inlier</i> Ratio	0.076	0.053
<i>Keypoints</i> vs Execution Time	0.919	0.876

Evaluasi klasifikasi dilakukan dengan kurva ROC untuk membandingkan kemampuan algoritma dalam membedakan pasangan wajah identik dan tidak identik.

Table 4. Hasil AUC

Algoritma	AUC
SIFT	0.944
ORB	0.873

Kurva ROC menggambarkan hubungan antara *True Positive Rate* (TPR) dan *False Positive Rate* (FPR) pada berbagai ambang batas keputusan, di mana semakin melengkung ke kiri atas menunjukkan semakin baik performa klasifikasi. Pada grafik tersebut, kurva SIFT (biru) cenderung lebih tinggi dibandingkan kurva ORB (merah) di sebagian besar rentang FPR, menunjukkan bahwa SIFT lebih baik dalam membedakan pasangan wajah yang identik dan tidak identik. Namun, kedua kurva relatif dekat dengan garis diagonal, yang mengindikasikan bahwa tingkat prediksi benar masih mendekati acak untuk beberapa kasus yang sulit. Nilai AUC yang diperoleh memperkuat interpretasi ini, dengan SIFT mencapai AUC sekitar 0,944 dan ORB sekitar 0,873.



Gambar 5. Kurva ROC Perbandingan Algoritma

Secara keseluruhan, analisis ROC ini menegaskan bahwa SIFT lebih handal untuk tugas verifikasi identitas berbasis pencocokan fitur dibandingkan ORB, meskipun masih terdapat ruang untuk peningkatan performa pada kedua metode.

4.4 Analisis Visual Pencocokan

Evaluasi visual terhadap hasil pencocokan memperkuat analisis numerik. Dalam beberapa kasus ekstrem (misalnya rotasi tinggi dan pencahayaan rendah), SIFT mampu mengenali pasangan wajah yang gagal dicocokkan oleh ORB. Hal ini menunjukkan keunggulan SIFT dalam ketahanan terhadap noise dan distorsi.

4.5 Pembahasan Temuan

Hasil eksperimen pada dataset nist_2 menegaskan bahwa SIFT secara konsisten menghasilkan jumlah titik kunci dan rasio kecocokan yang lebih tinggi dibanding ORB, meski memerlukan waktu komputasi lebih lama. Temuan ini mendukung studi Marlinda et al. [4], [10] dan teori Lowe [2] mengenai ketahanan SIFT terhadap perubahan skala, rotasi, dan kondisi pencahayaan faktor kritis dalam aplikasi forensik biometrik. Rata-rata rasio kecocokan SIFT tetap di atas 90 % meski citra diputar hingga 45°, sedangkan ORB mengalami penurunan hingga 25 % pada rotasi > 30° [3], selaras dengan laporan Matusiak et al. [12] tentang rotasi invariansi SIFT.

Sebaliknya, ORB membuktikan kemampuannya dalam efisiensi komputasi, dengan eksekusi rata-rata 4,7-lipat lebih cepat daripada SIFT. Kecepatan ini sesuai dengan tujuan ORB sebagai detektor ringan untuk aplikasi real-time [3], [8]. Kennerley [8] dan Suárez et al. [9] menekankan bahwa trade-off antara akurasi dan kecepatan menjadi kriteria pemilihan algoritma: ORB dan deskriptor biner boosting seperti BEBLID ideal untuk perangkat terbatas, meski akurasinya lebih rendah dibanding SIFT.

Kontribusi orisinal penelitian ini terletak pada integrasi threshold adaptif berbasis distribusi jarak deskriptor SIFT, yang meningkatkan presisi pencocokan hingga 15 % tanpa menaikkan beban komputasi signifikan [13]. Pendekatan ini memperkaya temuan Suárez et al. [9] dan Cui et al. [13], yang menunjukkan pentingnya penyesuaian parameter adaptif terhadap variabilitas data masukan. Validasi dengan struktur groundtruth pada subset train dan val nist_2 [22] semakin memperkuat keandalan hasil, menutupi kekurangan studi Kennerley [8] yang tidak menguji algoritma pada data berlabel.

Beberapa studi menawarkan perspektif alternatif. Klemm et al. [5] dan Paul & Aslan [6] melaporkan bahwa metode berbasis Local Binary Pattern Histogram (LBPH) dapat mengejar kinerja SIFT dan ORB pada citra noise tinggi atau resolusi rendah, menandakan bahwa pemilihan algoritma harus mempertimbangkan karakteristik data. Selain itu, studi deep learning seperti SuperPoint [17], D2-Net [18], R2D2 [19], dan ASLFeat [20] oleh Sanil et al. [7] dan Xu et al. [16] menunjukkan potensi integrasi deteksi tradisional dan pembelajaran mendalam untuk

meningkatkan robustnes, meski memerlukan sumber daya pelatihan besar. Secara keseluruhan, temuan ini menegaskan bahwa tidak ada solusi Tunggal SIFT unggul pada akurasi, ORB pada efisiensi, dan metode adaptif maupun hibrida menawarkan keseimbangan optimal sesuai kebutuhan aplikasi.

5. Simpulan

Penelitian ini telah membandingkan performa dua algoritma deteksi dan pencocokan fitur, yaitu *Scale-Invariant Feature Transform* (SIFT) dan *Oriented FAST and Rotated BRIEF* (ORB), dalam konteks pengolahan citra wajah yang mengalami variasi pencahayaan, orientasi, dan ekspresi. Berdasarkan hasil evaluasi kuantitatif dan kualitatif, dapat disimpulkan bahwa ORB menunjukkan keunggulan signifikan dari segi efisiensi komputasi, dengan waktu ekstraksi dan pencocokan fitur yang secara konsisten lebih cepat dibandingkan SIFT, sehingga lebih sesuai untuk aplikasi *real time* yang berjalan pada perangkat dengan sumber daya terbatas. Di sisi lain, SIFT terbukti lebih unggul dalam aspek akurasi pencocokan dan ketahanan terhadap kondisi wajah yang tidak ideal, seperti pencahayaan rendah dan deformasi non rigid akibat ekspresi wajah. Hal ini terlihat dari rata-rata rasio *inlier* yang lebih tinggi serta nilai *Area Under Curve* (AUC) yang mencapai 0,944, menunjukkan kemampuan SIFT dalam membedakan pasangan wajah identik dan tidak identik secara lebih konsisten.

Evaluasi visual terhadap pola pencocokan *inlier* juga memperlihatkan bahwa SIFT mampu menangani kondisi ekstrem seperti rotasi tinggi dan oklusi parsial dengan lebih baik, sementara ORB lebih rawan terhadap kesalahan lokal yang tidak selalu terdeteksi secara numerik. Dengan demikian, pemilihan algoritma sebaiknya disesuaikan dengan konteks aplikasi: ORB untuk kecepatan dan efisiensi, dan SIFT untuk presisi dan robustnes dalam situasi kritis.

Ke depan, penelitian ini membuka prospek pengembangan melalui eksplorasi metode pra-pemrosesan tambahan seperti transformasi logaritmik dan augmentasi geometrik guna meningkatkan keragaman input. Penggunaan dataset yang lebih kompleks dari lingkungan dunia nyata (*in the wild*) serta integrasi metrik evaluasi berbasis ground truth otomatis juga menjadi arah penting untuk meningkatkan objektivitas analisis. Selain itu, perbandingan dengan model deep learning modern seperti *SuperPoint* atau D2-Net, serta pengembangan *pipeline hybrid* yang dapat memilih algoritma secara dinamis berdasarkan kondisi input, merupakan langkah lanjutan yang potensial untuk menghasilkan sistem pengenalan wajah yang adaptif, efisien, dan andal dalam berbagai skenario aplikasi.

Daftar Referensi

- [1] G. Guo and N. Zhang, "A survey on deep learning based face recognition," *Computer Vision and Image Understanding*, vol. 189, p. 102805, 2019. doi: 10.1016/j.cviu.2019.102805
- [2] B. Jafraسته, S. P. Lubián-López, and I. Benavente-Fernández, "A deep sift convolutional neural networks for total brain volume estimation from 3D ultrasound images," *Computer Methods and Programs in Biomedicine*, vol. 237, p. 107805, 2023. doi: 10.1016/j.cmpb.2023.107805
- [3] L. Cai, Y. Ye, X. Gao, Z. Li, and C. Zhang, "An improved visual SLAM based on affine transformation for ORB feature extraction," *Optik*, vol. 223, p. 165421, 2020. doi: 10.1016/j.ijleo.2020.165421
- [4] L. Marlinda, S. Rustad, R.S. Basuki, F. Budiman, dan M. Fatchan, "Matching images on the face of a Buddha statue using the scale invariant feature transform (SIFT) method," in *Proc. 2020 7th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, Semarang, Indonesia, 24–25 Sept. 2020, pp. 169-172, doi: 10.1109/ICITACEE50144.2020.9239221.
- [5] Y. Xie, W. Tian, H. Zhang, and T. Ma, "Facial expression recognition through multi-level features extraction and fusion," *Soft Computing*, vol. 27, pp. 11243–11258, 2023. doi: 10.1007/s00500-023-08531-z
- [6] K.C. Paul and S. Aslan, "An improved real-time face recognition system at low resolution based on local binary pattern histogram algorithm and CLAHE," *Optics and Photonics Journal*, vol. 11, no. 4, pp. 63–78, 2021. doi: 10.4236/opj.2021.114005
- [7] G. Sanil, K. Prakasha K., S. Prabhu, V. Nayak, and A. Jayakala, "Region-wise landmarks-based feature extraction employing SIFT, SURF, and ORB feature descriptors to recognize Monozygotic twins from 2D/3D facial images," *F1000Research*, vol. 14, art. no. 444, 2025. doi: 10.12688/f1000research.162911.1

- [8] M. Kennerley, "A comparison of SIFT, SURF and ORB on OpenCV," *Medium*, 2021. [Online]. Available: A Comparison of SIFT, SURF and ORB on OpenCV | by Mikhail Kennerley | Medium. [Accessed: May 16, 2025].
- [9] I. Suárez, G. Sfeir, J.M. Buenaposada, and L. Baumela, "BEBLID: Boosted efficient binary local image descriptor," *Pattern Recognition Letters*, vol. 133, pp. 366–372, May 2020. doi: 10.1016/j.patrec.2020.04.005
- [10] L. Marlinda, F. Budiman, R. Basuki, and A.Z. Fanani, "Comparison of SIFT and ORB methods in identifying the face of Buddha statue," *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 8, no. 2, pp. 139–144, Feb. 2023. doi: 10.33480/jitk.v8i2.4086
- [11] M. Isik, "Comprehensive empirical evaluation of feature extractors in computer vision," *PeerJ Computer Science*, vol. 10, art. no. e2415, Nov. 2024. doi: 10.7717/peerj-cs.2415
- [12] K. Matusiak, P. Skulimowski, and P. Strumillo, "Unbiased evaluation of keypoint detectors with respect to rotation invariance," *IET Computer Vision*, vol. 11, no. 7, pp. 547–555, Sep. 2017. doi: 10.1049/iet-cvi.2016.0434
- [13] R. Cui, M. Wen, K. Zhang, and C. Sun, "Contrast threshold adaptive adjustment algorithm for remote sensing image matching," *J. Appl. Remote Sens.*, vol. 15, no. 3, p. 036519, Sep. 2021, doi: 10.1117/1.JRS.15.036519.
- [14] O. Sidhom, H. Ghazouani, and W. Barhoumi, "Three-phases hybrid feature selection for facial expression recognition," *J. Supercomput.*, vol. 80, pp. 8094–8128, 2024, doi: 10.1007/s11227-023-05758-3.
- [15] N. Kaur, S. Rani, and S. Kaur, "Real-time video surveillance based human fall detection system using hybrid Haar cascade classifier," *Multimed. Tools Appl.*, vol. 83, pp. 71599–71617, 2024, doi: 10.1007/s11042-024-18305-w.
- [16] S. Xu, S. Chen, R. Xu, C. Wang, P. Lu, and L. Guo, "Local feature matching using deep learning: A survey," *Information Fusion*, vol. 101, p. 102344, 2024, doi: 10.1016/j.inffus.2024.102344.
- [17] D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperPoint: Self-Supervised Interest Point Detection and Description," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Salt Lake City, UT, USA, Jun. 2018, pp. 224–236.
- [18] M. Dusmanu, I. Rocco, T. Pajdla, M. Pollefeys, J. Sivic, A. Torii, and T. Sattler, "D2-Net: A Trainable CNN for Joint Description and Detection of Local Features," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Long Beach, CA, USA, Jun. 2019, pp. 8096–8105.
- [19] J. Revaud, C. De Souza, M. Humenberger, and P. Weinzaepfel, "R2D2: Repeatable and Reliable Detector and Descriptor," in *Advances in Neural Information Processing Systems*, vol. 32, NeurIPS 2019, Vancouver, Canada, Dec. 2019.
- [20] Z. Luo, L. Zhou, X. Bai, H. Chen, J. Zhang, Y. Yao, S. Li, T. Fang, and L. Quan, "ASLFeat: Learning Local Features of Accurate Shape and Localization," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 6589–6598.
- [21] H. Abdellali, R. Frohlich, V. Vilagos, and Z. Kato, "L2D2: Learnable Line Detector and Descriptor," in *Proc. 2021 Int. Conf. on 3D Vision (3DV)*, London, United Kingdom, Dec. 2021, doi: 10.1109/3DV53792.2021.00054.
- [22] R. A. Pramunendar, *nist_2 Facial Image Dataset*, "Grayscale facial images with varying lighting, orientation, and expression." [Online]. Available: <https://drive.google.com/drive/folders/1xlomu91x1H00JKqqWmrXIHuixvHRxumo> [Accessed: Apr. 27, 2025].