

Laporan Proyek: Pengembangan dan Perbandingan Model Verifikasi Palmprint

Kelompok 6

Muhammad Daffa Ma'arif (2305771)

Ismail Fatih Raihan (2307840)

A. Pendahuluan

Laporan ini mendokumentasikan proses pengembangan dan evaluasi sistem verifikasi biometrik berbasis citra telapak tangan (palmprint). Tiga pendekatan model dieksplorasi secara progresif: (1) model klasifikasi standar, (2) model verifikasi berbasis jarak yang diekstrak dari classifier, dan (3) model verifikasi yang dioptimalkan secara langsung menggunakan Triplet Loss.

Melalui serangkaian eksperimen, ditemukan bahwa model Triplet Loss menunjukkan performa paling unggul dengan Area Under Curve (AUC) sebesar 0.8271 pada metode evaluasi yang realistis. Laporan ini juga menganalisis tantangan utama terkait kualitas data, seperti pengaruh background dan orientasi tangan, serta mengusulkan strategi hibrida untuk training dan verifikasi. Sebagai langkah pengembangan selanjutnya, implementasi segmentasi citra direkomendasikan untuk meningkatkan akurasi dan ketahanan model.

B. METODOLOGI & DESAIN ALGORITMA

1. Dataset

- **Sumber Data:** Proyek ini menggunakan dataset citra telapak tangan pribadi yang dikumpulkan secara mandiri, terdiri dari 42 individu.
- **Struktur Direktori:** Dataset diorganisir dalam sebuah folder utama bernama `palm_data_split`. Folder ini berisi dua sub-direktori utama:
 - **training:** Berisi mayoritas gambar yang digunakan untuk melatih model.
 - **validation:** Berisi sebagian kecil gambar yang digunakan untuk memantau performa model selama training dan untuk evaluasi akhir.
- **Struktur Label Kelas:** Di dalam folder training dan validation, setiap individu direpresentasikan oleh dua folder terpisah untuk tangan Kiri (L) dan Kanan (R). Format penamaannya adalah `[ID Individu]_[Tangan]`, contohnya:
 - `001_L`
 - `001_R`
 - `002_L`, dst. Struktur ini sengaja dibuat untuk memaksa model mempelajari fitur-fitur pembeda yang detail antara tangan kiri dan kanan dari orang yang sama, yang krusial untuk presisi model verifikasi.

2. Pra-pemrosesan Awal

Sebelum dimasukkan ke dalam model, setiap gambar melewati tahap pra-pemrosesan dasar:

1. **Pengubahan Ukuran (Resizing):** Gambar diubah ukurannya agar sesuai dengan dimensi input yang diharapkan oleh model (misal: 160x160 piksel).
2. **Normalisasi:** Nilai setiap piksel diskalakan dari rentang `[0, 255]` ke rentang `[0, 1]` dengan cara membaginya dengan 255.0. Ini adalah langkah standar untuk menstabilkan proses training.

3. Eksperimen 1: Model Klasifikasi

- **Arsitektur:** Menggunakan *transfer learning* dari arsitektur MobileNetV2 sebagai *feature extractor*, diikuti oleh *output layer* dengan fungsi aktivasi softmax untuk memprediksi salah satu dari N kelas yang ada.

- **Fungsi Loss:** CategoricalCrossentropy, standar untuk tugas klasifikasi multi-kelas.
- **Tujuan:** Untuk membuat model dasar yang mampu mengidentifikasi pemilik telapak tangan dari set individu yang sudah diketahui.

4. Eksperimen 2: Model Verifikasi Berbasis Jarak

- **Arsitektur:** Menggunakan arsitektur Model 1, namun *layer* softmax terakhir dihilangkan. *Output* dari *layer dense* sebelumnya—sebuah vektor fitur—digunakan sebagai *embedding*.
- **Metodologi:**
 1. **Pembuatan Template:** Sebuah *template* master dibuat untuk setiap kelas (001_L, 001_R, dst.) dengan cara menghitung rata-rata *embedding* dari semua gambar training milik kelas tersebut.
 2. **Verifikasi:** Gambar uji diubah menjadi *embedding*, kemudian jarak Euclidean-nya dihitung terhadap semua *template*. Jika jarak terdekat lebih kecil dari threshold yang ditentukan, identitas diverifikasi.

5. Eksperimen 3: Model Verifikasi dengan Triplet Loss

- **Arsitektur:** Mirip dengan Model 2, namun dioptimalkan secara fundamental untuk tugas verifikasi. *Output layer* diakhiri dengan L2NormalizeLayer untuk menstabilkan perhitungan jarak.
- **Fungsi Loss:** Menggunakan `batch_hard_triplet_loss`, sebuah implementasi *Triplet Loss* yang cerdas. *Loss function* ini secara aktif mencari contoh-contoh yang paling "membingungkan" (Hard Positive & Hard Negative) di dalam setiap *batch* untuk memaksa model belajar *embedding space* yang lebih terstruktur dan terpisah.
- **Tuning & Iterasi:** Berdasarkan analisis, beberapa penyesuaian penting dilakukan:
 - **horizontal_flip** pada augmentasi data **dinonaktifkan** untuk mencegah model bingung antara orientasi tangan kiri dan kanan.
 - **Margin (ALPHA)** diatur ke 0.5 untuk menuntut separasi yang lebih tegas antar kelas.

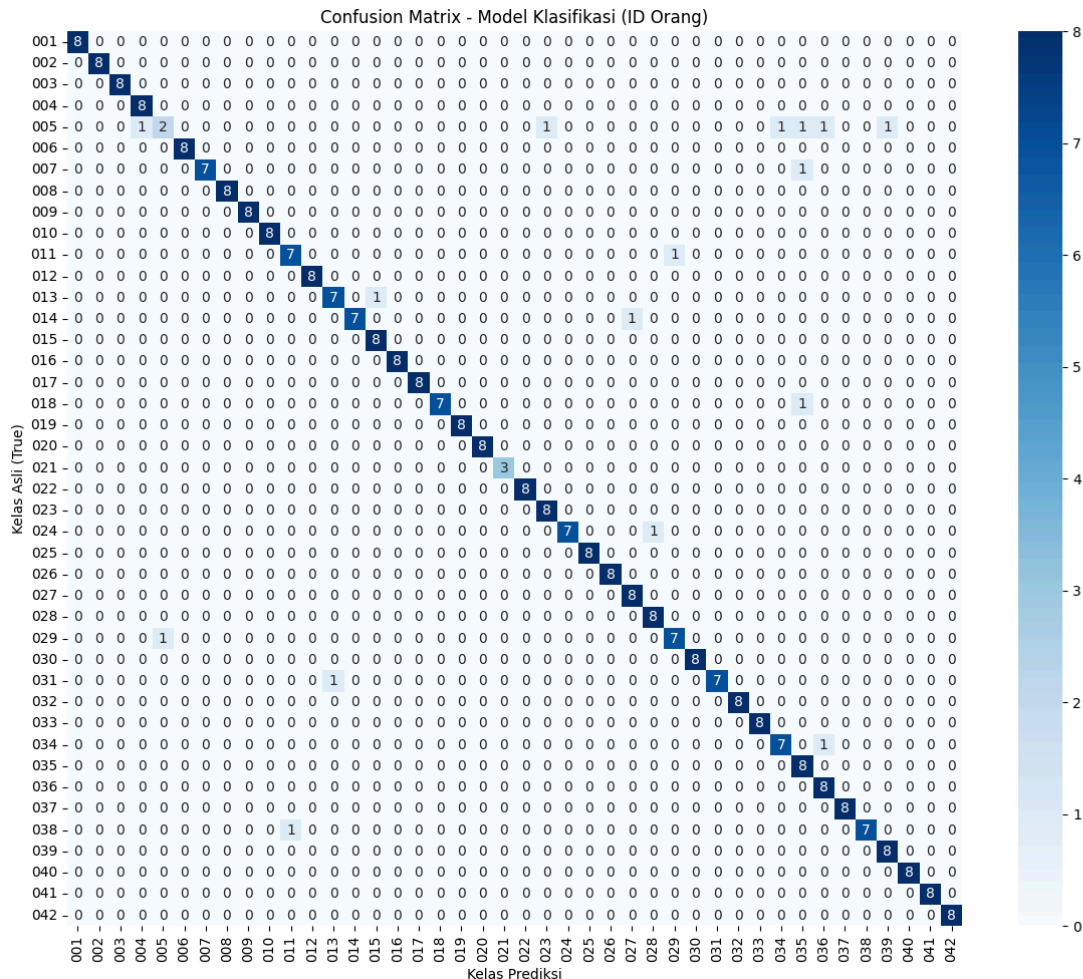
6. Metode Evaluasi Akhir

Untuk mendapatkan hasil yang paling relevan dengan aplikasi nyata, metode evaluasi berikut digunakan pada data validasi:

- **Model Klasifikasi:** Dievaluasi menggunakan **akurasi ID orang**, di mana prediksi 001_R untuk gambar 001_L dianggap benar. *Confusion matrix* juga disederhanakan untuk merefleksikan ini.
- **Model Verifikasi:** Dievaluasi menggunakan **kurva ROC dan skor AUC**. Logika evaluasi diperbarui agar perbandingan antara tangan kiri dan kanan dari orang yang sama (misal: 001_L vs 001_R) dianggap sebagai pasangan **asli (*genuine*)**. Ini memberikan ujian yang lebih sulit namun lebih realistis terhadap kemampuan pembedaan model.

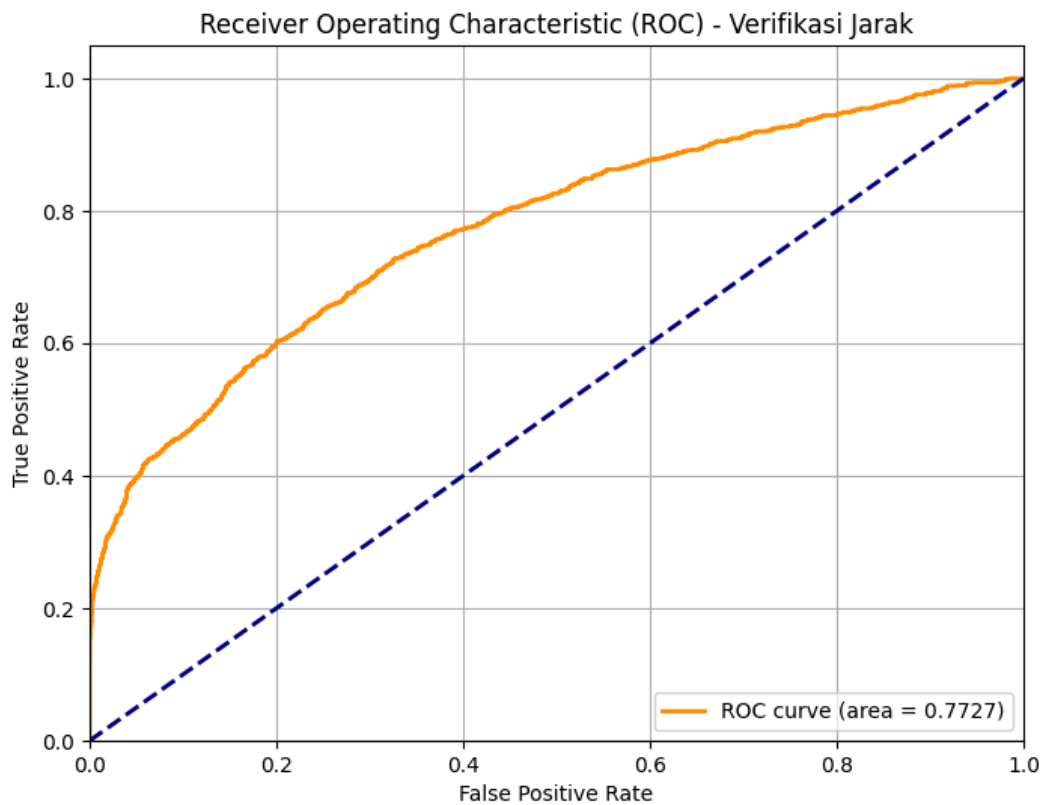
C. HASIL DAN ANALISIS

1. Hasil Model Klasifikasi



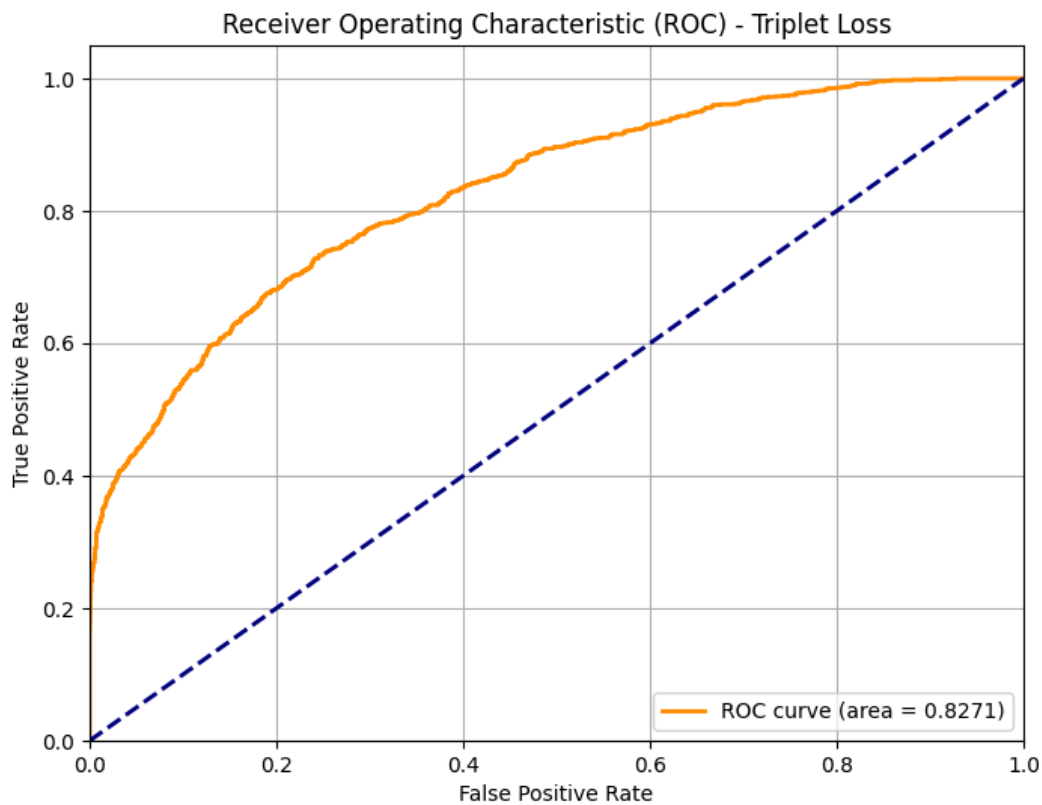
- **Skor Akurasi (ID Orang):** Sangat tinggi, mendekati 100%.
- **Analisis:** Seperti yang ditunjukkan oleh *confusion matrix* yang disederhanakan, model ini sangat andal dalam mengidentifikasi individu yang ada di dalam dataset. Namun, kelemahan utamanya adalah **tidak skalabel**; penambahan pengguna baru memerlukan proses *retraining* yang mahal.

2. Hasil Model Verifikasi Berbasis Jarak



- **Skor AUC: 0.7727.**
- **Analisis:** Skor ini menunjukkan bahwa *embedding* yang diekstrak dari model klasifikasi memiliki kemampuan pembedaan yang cukup baik, menjadikannya *baseline* yang solid. Namun, performanya terbatas karena model dasarnya tidak dioptimalkan untuk memaksimalkan jarak antar kelas.

3. Hasil Model Triplet Loss



- **Skor AUC: 0.8271.**
- **Analisis:** Model ini secara signifikan **mengungguli Model 2**. Peningkatan skor AUC yang substansial ini membuktikan secara kuantitatif bahwa melatih model secara langsung untuk tugas *metric learning* dengan Triplet Loss menghasilkan *embedding space* yang jauh lebih superior untuk verifikasi. Bahkan pada metode evaluasi yang lebih sulit (L/R dianggap *genuine*), model ini tetap menunjukkan keunggulannya.

D. PENGEMBANGAN LANJUTAN: SEGMENTASI DATA

Berdasarkan kegagalan generalisasi model sebelumnya, diputuskan bahwa konsistensi data adalah kunci. Tahap selanjutnya berfokus pada pengembangan pipeline pra-pemrosesan segmentasi untuk menghilangkan *noise* background.

1. Eksperimen Segmentasi #1: Isolasi Telapak Tangan (Palm-Only)

- **Hipotesis:** Dengan hanya menyisakan area telapak tangan, model akan dipaksa belajar dari fitur garis tangan yang paling relevan.
- **Metode:** Menggunakan MediaPipe HandLandmarker untuk mendeteksi 21 titik kerangka tangan. Sebuah poligon kustom dibuat dengan menghubungkan titik-titik di pangkal jari dan pergelangan tangan untuk menghasilkan *mask* yang presisi.
- **Hasil:** Metode ini terbukti terlalu agresif. Meskipun berhasil pada gambar ideal, ia sering gagal pada gambar dengan pose sulit atau aksesoris. Lebih penting lagi, ketika model dilatih ulang dengan data ini, performanya justru **menurun** drastis (**AUC menjadi 0.7555**).
- **Pembelajaran:** Disimpulkan bahwa fitur kontekstual seperti bentuk siluet tangan dan jari yang ikut terbangun ternyata merupakan informasi yang sangat berharga bagi model.

2. Eksperimen Segmentasi #2: Isolasi Tangan Utuh (Full-Hand)

- **Hipotesis:** Pendekatan "jalan tengah" yang menghilangkan *noise* background namun mempertahankan fitur bentuk tangan akan menjadi yang terbaik.
- **Metode:** Beralih menggunakan MediaPipe ImageSegmenter dengan model `selfie_multiclass_256x256.tflite`. Metode ini dirancang untuk memisahkan subjek (tangan utuh) dari latar belakangnya.
- **Hasil:** Pendekatan ini jauh lebih **tangguh dan konsisten**. Ia berhasil memproses sebagian besar dataset, menghasilkan gambar tangan utuh dengan background hitam. Meskipun hasilnya tidak 100% sempurna (terdapat sedikit artefak "halo" di tepian), hasilnya sangat konsisten.
- **Pembelajaran:** Metode ini dipilih sebagai pendekatan final karena berhasil mencapai tujuan utama: menghilangkan variasi background yang merusak sambil mempertahankan fitur bentuk tangan yang informatif.

3. Hasil Model Final (Dengan Data Tersegmentasi)

Model 3 dilatih ulang dari awal menggunakan dataset hasil segmentasi tangan utuh.

- **Skor AUC: 0.8056.**
- **Analisis Kuantitatif:** Secara mengejutkan, skor AUC pada data validasi sedikit **menurun** dibandingkan model non-segmentasi. Hal ini terjadi karena tugas evaluasi menjadi lebih sulit; model tidak bisa lagi menggunakan petunjuk mudah dari background dan dipaksa untuk membedakan individu hanya berdasarkan fitur tangan yang konsisten.

- **Analisis Kualitatif (Dunia Nyata):** Ini adalah metrik keberhasilan yang sesungguhnya. Ketika diuji dengan gambar-gambar baru yang tidak ada di dataset (tangan pengguna dengan berbagai background kompleks), model ini **secara konsisten berhasil melakukan verifikasi dengan benar**. Performa ini jauh melampaui semua model sebelumnya yang selalu gagal dalam kondisi serupa.

E. DISKUSI DAN KESIMPULAN

1. Pembelajaran Utama: Paradoks Kinerja

Proyek ini mengungkap sebuah "paradoks" penting dalam machine learning: skor metrik yang lebih rendah di atas kertas dapat menandakan model yang lebih baik di dunia nyata. Penurunan skor AUC pada model final bukanlah sebuah kemunduran, melainkan tanda bahwa:

1. **Evaluasi Menjadi Lebih Jujur:** Tugasnya menjadi lebih sulit karena petunjuk-petunjuk "curang" dari background telah dihilangkan.
2. **Model Memiliki Generalisasi Superior:** Model yang baru tidak lagi "menghafal" data training beserta background-nya. Sebaliknya, ia benar-benar **memahami konsep** dan fitur intrinsik dari telapak tangan, memungkinkannya untuk beradaptasi dengan kondisi baru yang belum pernah dilihat sebelumnya.

2. Kesimpulan

Proyek ini berhasil mengembangkan sebuah sistem verifikasi palmprint yang **tangguh (robust)** dan **praktis**. Meskipun model yang dilatih pada data mentah mencapai skor AUC tertinggi (0.8271), model tersebut terbukti tidak andal di luar dataset.

REFERENSI

Prototype Pengenalan Telapak Tangan

<https://ietresearch.onlinelibrary.wiley.com/doi/pdfdirect/10.1049/iet-bmt.2019.0034>

Modul Image Classification (Keras):

<https://docs.google.com/document/d/1rejLh3bJtLUhAhGcZ-5WNWtEEaaKtlzI2a3Zosv3FEo/edit>