

ResNet50-Based Mobile Application for Big Five Personality Detection Using Handwriting

Adelia Salsabila Arifin ^{1*}, Salamun Rohman Nudin ^{2*}

* Manajemen Informatika, Universitas Negeri Surabaya
adeliasalsabila.22022@mhs.unesa.ac.id¹, salamunrohman@unesa.ac.id²

Article Info

Article history:

Received 2026-04-30

Revised 2026-06-03

Accepted 2026-06-11

Keyword:

*Deep Learning,
Grafologi,
Big Five personality,
ResNet.*

ABSTRACT

Handwriting reflects a person's unique traits and has long been studied in the field of graphology to uncover personality characteristics. However, traditional graphological analysis is subjective, time-consuming, and prone to inter-rater differences. This study aims to develop a PenaKepribadian mobile application using ResNet50 transfer learning to automatically identify Big Five personality traits (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism). The HiEnWrite dataset contains 327 English handwritten images annotated by a certified graphologists, was utilized with an 80:20 train-test split. Three optimizers such as SGD, RMSprop, and Adam were comparatively evaluated. Adam achieved the best performance with a training PCC of 0.5872 with 91.90% accuracy and a testing PCC of 0.4719 with 89.95% accuracy, outperforming both SGD and RMSprop. A testing PCC of 0.4719 indicates moderate correlation, suggesting promising yet improvable results. Robustness testing across varying lighting conditions, paper backgrounds, and writing media showed consistent performance, with mean prediction deviations ranging from 0.070 to 0.126. All Black Box Testing scenarios returned valid results. These findings confirm that ResNet50 transfer learning effectively extracts handwriting features for personality prediction, though further improvements remain necessary before high-stakes deployment. This research contributes to personality computing and opens avenues for efficient, automated, and accessible personality assessment systems.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Tulisan tangan merupakan salah satu bentuk ekspresi unik manusia yang dapat merefleksikan aspek kepribadian, emosi, serta pola berpikir individu. Hubungan antara tulisan tangan dan kepribadian telah lama dikaji dalam bidang grafologi, yang berasumsi bahwa karakteristik tulisan seperti tekanan pena, kemiringan huruf, ukuran, dan jarak antar huruf dapat mengungkap sifat psikologis seseorang [1]. Pendekatan grafologi telah dimanfaatkan dalam berbagai bidang, mulai dari rekrutmen karyawan, pendidikan dan bimbingan karier, hingga psikologi klinis dan forensik [2]. Hal ini menunjukkan luasnya potensi tulisan tangan sebagai media penilaian kepribadian. Namun demikian, analisis grafologi secara tradisional masih dilakukan secara manual sehingga bersifat

subjektif, memerlukan waktu yang relatif lama, serta rentan terhadap perbedaan interpretasi antarpenilai [3].

Kepribadian merupakan faktor penting yang memengaruhi perilaku, pengambilan keputusan, dan interaksi sosial individu. Salah satu model kepribadian yang diakui secara luas dalam kajian ilmiah adalah Big Five Personality Traits yang mencakup lima dimensi utama, yaitu *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness*, dan *Neuroticism* [4]. Setiap kepribadian menggambarkan cakupan luas mengenai karakteristik individu, sehingga terdapat perbedaan sifat mengenai tingkat rendah dan tinggi masing-masing kepribadian tersebut [5], yang dapat dilihat di Tabel I. Model ini dinilai memiliki validitas empiris yang kuat serta stabilitas yang paling baik dalam menggambarkan perilaku manusia terutama konteks pekerjaan dan hubungan sosial dibandingkan model kepribadian lainnya, sehingga banyak

digunakan dalam penelitian psikologi maupun analisis berbasis komputasi [6][7]. Meskipun demikian, metode pengukuran kepribadian secara konvensional umumnya menggunakan kuesioner atau wawancara yang memerlukan partisipasi langsung responden dan berpotensi menimbulkan bias jawaban [2].

TABEL I
TINGKATAN KEPERIBADIAN BIG FIVE

Dimensi	Rendah	Tinggi
Openness	konvensional, praktis, skeptis	imajinatif, ingin tahu, kreatif
Conscientiousness	ceroboh, spontan, hedonistik	rajin, teratur, tepat waktu
Extraversion	menikmati waktu sendiri, acuh tak acuh, pendiam	aktif, supel, mencari kesenangan
Agreeableness	tidak peduli, tidak kooperatif	empatik, percaya, patuh
Neuroticism	puas dengan diri sendiri, stabil secara emosional	emosi negatif, cenderung stres

Seiring perkembangan kecerdasan buatan, muncul pendekatan otomatis untuk menganalisis tulisan tangan menggunakan teknik *machine learning* dan *deep learning*. Penelitian awal dalam bidang *computer-aided graphology* mengadopsi tahapan pengenalan pola yang meliputi pra-pemrosesan, ekstraksi fitur, dan klasifikasi [3]. Meskipun validitas ilmiah grafologi konvensional masih memicu perdebatan dalam psikologi modern karena minimnya dukungan bukti empiris yang konsisten [8], analisis tulisan tangan kini mengalami pergeseran paradigma dalam ranah *personality computing*. Pendekatan berbasis *deep learning* memanfaatkan citra tulisan tangan sebagai sinyal observasional (*behavioral residue*) yang diproses secara otomatis. Metode ini tidak bertujuan memvalidasi klaim grafologi tradisional, melainkan mengekstraksi pola statistik visual yang berkorelasi dengan label kepribadian dari instrumen psikometrik tervalidasi [9].

Metode *machine learning* seperti Support Vector Machine (SVM) dan Random Forest telah digunakan untuk klasifikasi kepribadian, namun masih bergantung pada proses ekstraksi fitur manual yang kompleks [10]. Convolutional Neural Network (CNN) mampu mengekstraksi fitur citra secara otomatis, tetapi arsitektur yang sangat dalam berpotensi mengalami masalah *vanishing gradient* yang mengganggu stabilitas pelatihan [11]. Untuk mengatasi kendala tersebut, dikembangkan arsitektur Residual Network (ResNet) yang menggunakan mekanisme *skip connection* sehingga proses pelatihan menjadi lebih stabil dan efisien [12]. Kemampuan ResNet dalam mengenali pola visual kompleks menjadikannya potensial sebagai ekstraktor fitur citra tulisan tangan untuk mengidentifikasi pola yang berkaitan dengan dimensi kepribadian.

Beberapa penelitian terdahulu telah mengeksplorasi pendekatan otomatis untuk mengenali kepribadian dari citra

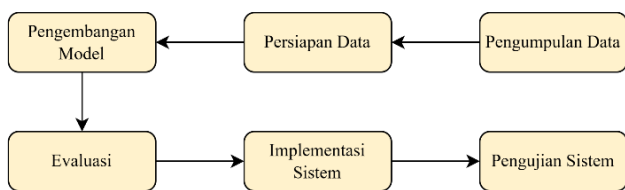
tulisan tangan. Penelitian oleh [13] mengusulkan arsitektur CNN berbasis *patch* untuk mengidentifikasi kepribadian Big Five dari 418 sampel, namun hanya memperoleh AUC 0,5023, menandakan perlunya arsitektur model yang lebih kuat. Di sisi lain, penelitian oleh [14] mengembangkan model Fine DenseNet yang dikombinasikan dengan AMDLSTM-CTC dan berhasil memperoleh akurasi 97,6%, meskipun *dataset* yang digunakan hanya 118 citra dan belum diuji pada kondisi *real-time*. Dan penelitian oleh [15], mengambil pendekatan berbeda dengan mengidentifikasi kepribadian melalui sidik jari menggunakan Minutiae dan SVM dengan akurasi 80,9%, namun metode ini memerlukan perangkat keras khusus sehingga kurang praktis untuk perangkat seluler umum.

Keunggulan arsitektur ResNet50 sendiri telah dibuktikan oleh studi di luar domain kepribadian. Penelitian oleh [16] menggabungkan *pre-trained* ResNet50 ImageNet dengan SVM untuk pengenalan karakter tulisan tangan Amharic dan mencapai akurasi 91,89%, menegaskan kapabilitas model dalam menghasilkan representasi fitur pada tulisan tangan yang kompleks. Selanjutnya, penelitian oleh [17] juga membuktikan keunggulan ResNet50 dibandingkan EfficientNet-B1 dan VGG16 dalam klasifikasi citra medis dengan akurasi 98,15%, menunjukkan performa arsitektur dalam menangkap pola visual yang rumit secara konsisten. Melihat berbagai keberhasilan tersebut menunjukkan bahwa model ResNet50 sangat potensial untuk diterapkan dan digunakan dalam mengenali karakteristik unik pada pola tulisan tangan.

Berdasarkan uraian tersebut, terdapat kesenjangan yang signifikan, yaitu belum adanya sistem yang mengintegrasikan *transfer learning* ResNet50, penggunaan *dataset* tulisan tangan yang divalidasi pakar grafologi, serta implementasi *end-to-end* dalam bentuk aplikasi *mobile*. Kesenjangan inilah yang menjadi tujuan utama penelitian ini, sekaligus menjembatani batas antara metode grafologi tradisional yang subjektif dengan kebutuhan sistem identifikasi kepribadian yang objektif, otomatis, dan aksesibel. Selain menghadirkan sistem terintegrasi dari pengambilan citra hingga visualisasi hasil prediksi, penelitian ini juga membandingkan pengaruh berbagai jenis optimizer terhadap kinerja model. Melalui pendekatan ini, penelitian diharapkan dapat berkontribusi pada bidang *personality computing* sekaligus membuka peluang pengembangan sistem asesmen kepribadian yang lebih efisien dan adaptif.

II. METODE

Metode penelitian yang digunakan pada penelitian ini mengacu pada CRISP-DM (*Cross-Industry Standard Process for Data Mining*), yang menyediakan kerangka kerja sistematis untuk mengelola tahapan penelitian berbasis data [18]. Seperti yang diilustrasikan pada Gambar 1.

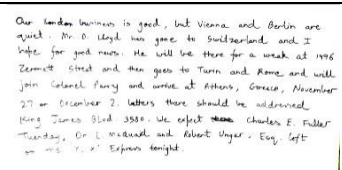
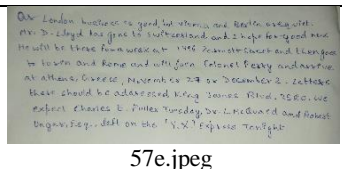
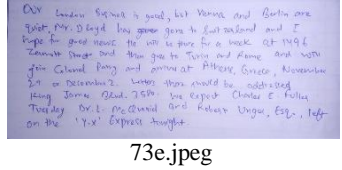
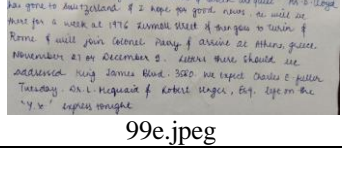


Gambar 1. Bagan Alur Penelitian

A. Pengumpulan Data

Pada penelitian ini, dataset yang digunakan berjudul ‘HiEnWrite Dataset’ dari repository GitHub milik Saksham Checker, dkk. [19]. Dataset tersebut mencakup citra tulisan tangan bahasa Inggris dan Hindi, namun hanya bahasa Inggris yang digunakan. Hal ini dikarenakan fokus analisis terletak pada bentuk huruf alfabet atau latin, yang memiliki perbedaan struktural dengan aksara Hindi. Total citra tulisan tangan bahasa Inggris yang diperoleh adalah 327 dengan label anotasi kepribadian Big Five yang divalidasi oleh pakar grafologi. Berdasarkan dokumentasi dataset HiEnWrite, proses anotasi melibatkan Dr. Sonam Arora, seorang pakar grafolog bersertifikat yang bertindak memberikan skor untuk setiap dimensi OCEAN. Contoh rincian dataset tertera di Tabel II.

TABEL II
CONTOH RINCIAN DATASET

File Gambar	Anotasi				
	O	C	E	A	N
 4e.jpeg	20	29	22	27	19
 57e.jpeg	33	36	3	13	25
 73e.jpeg	9	7	23	14	37
 99e.jpeg	27	35	26	33	6

Anotasi tersebut berdasarkan Big Five Personality Traits dengan dimensi OCEAN (*Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism*). Nilai anotasi kemudian dinormalisasi dengan menjadikannya skala 0-1, sehingga menghasilkan rentang nilai yang lebih terstandarisasi dan mudah dianalisis.

B. Persiapan Data

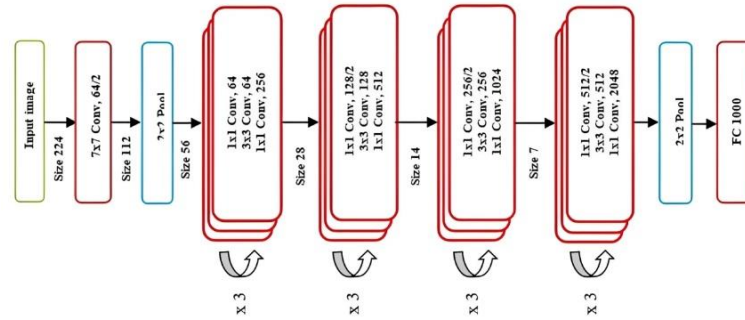
Tahap persiapan data meliputi seluruh proses pengolahan data mentah menjadi data yang siap dilakukan proses pengembangan model.

- 1) *Pembersihan Data*: Dilakukan proses eliminasi secara manual semua sampel nonalfabet (*dataset* Hindi), serta menghapus gambar yang *error*. Menghasilkan 327 gambar bersih yang digunakan dalam proses pengembangan dan pengujian model.
- 2) *Pembagian & Penyesuaian Ukuran Data*: Dataset dibagi menggunakan parameter *test_size=0.2* dengan rasio 80:20, sehingga diperoleh 261 data latih dan 66 data uji. Guna menghindari potensi *data leakage*, pembagian dilakukan secara acak pada tingkat sampel individu. Karena setiap baris data merepresentasikan satu individu yang berbeda sesuai dengan struktur asli dataset, maka tidak terdapat *overlap* karakteristik subjek antara data latih dan uji. Selanjutnya, penyesuaian ukuran gambar disesuaikan dengan model ResNet yaitu 224x224 piksel.
- 3) *Augmentasi Data*: Dilakukan proses peningkatan variasi dan penyeimbangan *dataset* dengan menggunakan transformasi geometris yang bersifat minimal, antara lain; Rotasi=2°, Pergeseran lebar=0.2, Pergeseran tinggi=0.2, dan Perbesar/Perkecil=0.1.

C. Pengembangan Model

Tahap ini adalah fase inti yang bertujuan membangun, melatih, dan mengoptimalkan model *deep learning* untuk klasifikasi kepribadian yang menggunakan Google Colab. Proses ekstraksi fitur dilakukan dengan model ResNet50 untuk memperoleh fitur-fitur paling relevan dari citra tulisan tangan, sesuai dengan arsitektur modelnya Gambar 2 [20]. Dalam proses ini, karakteristik yang khas dan relevan, diekstrak dari piksel gambar dari dataset besar berupa *pre-trained* ImageNet [19]. Selanjutnya, bobot yang sudah dipelajari ditransfer ke arsitektur dengan seluruh lapisan konvolusi ResNet50 dibuka agar model belajar dengan baik pada data baru.

Paramater yang digunakan pada pelatihan model penelitian ini adalah *learning rate* 0.0001 dengan *batch size* 32. Fungsi *Dropout* (0.4) dan (0.2) juga dimasukkan di antara lapisan *dense* untuk mencegah *overfitting*. Augmentasi data dilakukan secara langsung pada data latih untuk membuat dataset lebih kuat dan bervariasi. Selanjutnya, model dilatih selama 100 *epoch* dengan *early stopping* 10 guna mendapatkan model pada kondisi paling optimal dan efisien waktu pelatihan.



Gambar 2. Arsitektur Model ResNet50

Penggunaan *optimizer* dilakukan untuk mencari nilai parameter yang menghasilkan prediksi paling akurat dengan menurunkan nilai kesalahan secara bertahap. Menurut [21], pemilihan *optimizer* yang tepat berpengaruh signifikan terhadap performa jaringan, karena setiap algoritma memiliki strategi yang berbeda-beda. *Optimizer* yang digunakan antara lain:

- 1) *SGD*: Metode iteratif dalam *machine learning* yang digunakan untuk mengoptimalkan penurunan *gradient* dengan memperbarui bobot model menggunakan satu sampel data pada setiap iterasi. Pendekatan ini memungkinkan proses pelatihan menjadi lebih efisien serta membantu mempercepat konvergensi [22].
- 2) *RMSprop*: Dikembangkan untuk mengatasi kekurangan SGD dengan memperkenalkan mekanisme *adaptive learning rate*, sehingga lebih efisien untuk data seperti citra atau teks, karena *learning rate* setiap parameter dapat menyesuaikan secara otomatis [23].
- 3) *Adam*: Kombinasi dari keunggulan RMSprop dan momentum dengan menghitung rata-rata momentum pertama dan kedua dari gradien, sering digunakan sebagai optimasi utama karena mampu menyesuaikan *learning rate* secara spesifik pada tiap parameter model, sehingga proses pelatihan lebih efisien [24].

Pemilihan *optimizer* dan *hyperparameter* dilakukan melalui proses eksperimen dan penyesuaian nilai secara bertahap hingga mendapatkan performa yang terbaik. Pendekatan iteratif ini sangat krusial karena setiap arsitektur model memiliki sensitivitas yang berbeda terhadap parameter seperti *learning rate* maupun *batch size*. Oleh karena itu, diperlukan ketelitian dalam menganalisa hasil setiap uji coba guna memastikan model tersebut stabil dan berjalan optimal serta terhindar dari risiko *overfitting*.

D. Evaluasi

Evaluasi dilakukan untuk mengukur seberapa efektif model dalam mengidentifikasi kepribadian Big Five menggunakan metrik evaluasi regresi.

- 1) *MSE*: Metrik populer dalam *machine learning* untuk menilai keakuratan model regresi. Berdasarkan konsep menghitung rata-rata kuadrat selisih antara nilai prediksi

dan nilai aktual, metode ini memberikan gambaran umum tingkat kesalahan model [25].

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

- 2) *MAE*: Metrik yang mengukur rata-rata selisih absolut antara nilai prediksi dan nilai sebenarnya, sehingga menunjukkan besar nilai kesalahan model [26].

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2)$$

- 3) *RMSE*: Akar kuadrat dari MSE yang digunakan untuk menggambarkan besarnya rata-rata nilai kesalahan dalam prediksi. Nilai RMSE yang rendah menunjukkan nilai yang dihasilkan suatu model perkiraan lebih dekat dengan nilai aslinya, begitupun sebaliknya [27].

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

- 4) *PCC*: Metrik untuk mengukur tingkat korelasi linear antara dua variabel. Koefisien PCC berada pada rentang nilai -1 hingga 1, yang menunjukkan arah hubungan linear yang kuat. Nilai yang mendekati 1 menandakan korelasi positif, mendekati -1 menunjukkan korelasi negatif, sedangkan nilai 0 mengindikasikan tidak adanya hubungan linear [28].

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4)$$

- 5) *Accuracy (approximately)*: Metrik untuk menunjukkan seberapa dekat prediksi model terhadap nilai sebenarnya dalam skala 100%. Perlu ditegaskan bahwa metrik ini merupakan pendekatan berbasis MAE, bukan akurasi klasifikasi standar. Digunakan pendekatan ini agar lebih mudah memahami performa model.

$$Acc_{approx} = (1 - MAE) \times 100 \quad (5)$$

III. HASIL DAN PEMBAHASAN

Pada bagian ini menjawab tujuan utama dari penelitian dengan memberikan wawasan tentang kinerja model untuk memprediksi kepribadian Big Five pada tulisan tangan seseorang sekaligus implementasinya dalam bentuk aplikasi.

A. Hasil Evaluasi Perbandingan Optimizer

Pada tahap ini, ketiga *optimizer* (SGD, RMSprop, dan Adam) diuji untuk membandingkan kinerjanya dalam memprediksi dimensi kepribadian Big Five dari tulisan tangan. Hasil pelatihan dan pengujian masing-masing *optimizer* ditunjukkan pada Tabel III dan Tabel IV.

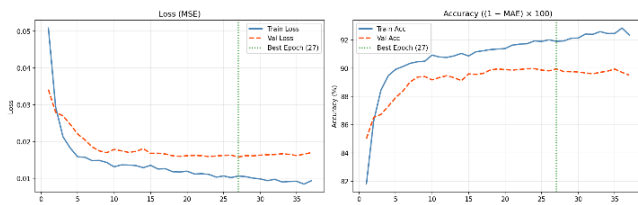
TABEL III
HASIL PELATIHAN

Optimizer	Loss	MAE	RMSE	PCC	Acc
SGD	0.0390	0.1596	0.1926	0.1311	84.04%
RMSprop	0.0139	0.0938	0.1229	0.3760	90.62%
Adam	0.0106	0.0810	0.1067	0.5872	91.90%

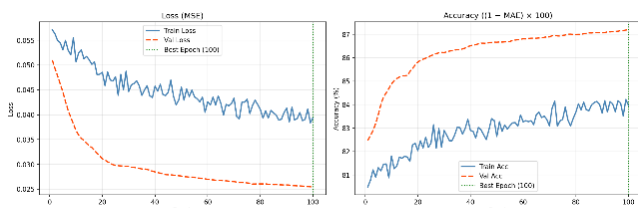
TABEL IV
HASIL PENGUJIAN

Optimizer	Loss	MAE	RMSE	PCC	Acc
SGD	0.0283	0.1365	0.1547	0.3142	86.35%
RMSprop	0.0175	0.1060	0.1311	0.3917	89.40%
Adam	0.0158	0.1005	0.1271	0.4719	89.95%

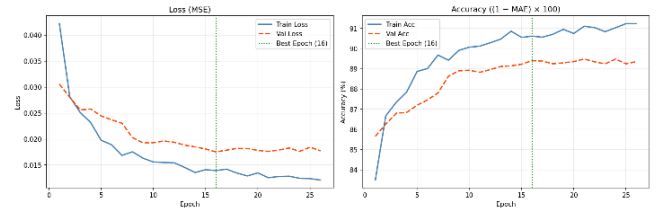
Model dengan *optimizer* Adam menunjukkan hasil yang lebih baik dibandingkan SGD dan RMSprop. Pada Gambar 2 merupakan grafik *loss* dan *accuracy* dari model dengan *optimizer* Adam, menunjukkan tren penurunan *loss* stabil dan nilai akurasi yang konsisten meningkat selama proses pelatihan. Hal ini mengindikasikan bahwa model sudah mampu menyesuaikan proses pembelajaran secara adaptif pada data tulisan tangan dalam penelitian ini. Gambar 3 dan 4 menampilkan grafik *loss* dan *accuracy* dari penggunaan *optimizer* SGD dan RMSprop sebagai pembandingan dengan *optimizer* Adam.



Gambar 2. Grafik Loss dan Accuracy Optimizer Adam



Gambar 3. Grafik Loss dan Accuracy Optimizer SGD



Gambar 4. Grafik Loss dan Accuracy Optimizer RMSprop

Pada grafik residual Gambar 5 menunjukkan sebaran nilai prediksi dan aktual yang terpusat di sekitar nol, mengindikasikan model tidak memiliki bias yang signifikan [29]. Meskipun demikian, terdapat beberapa titik dengan penyimpangan yang lebih besar, terutama pada dimensi *Extraversion*, yang kemungkinan disebabkan oleh variasi karakteristik tulisan tangan yang tinggi pada dimensi tersebut.

B. Skenario Perbandingan Ground Truth

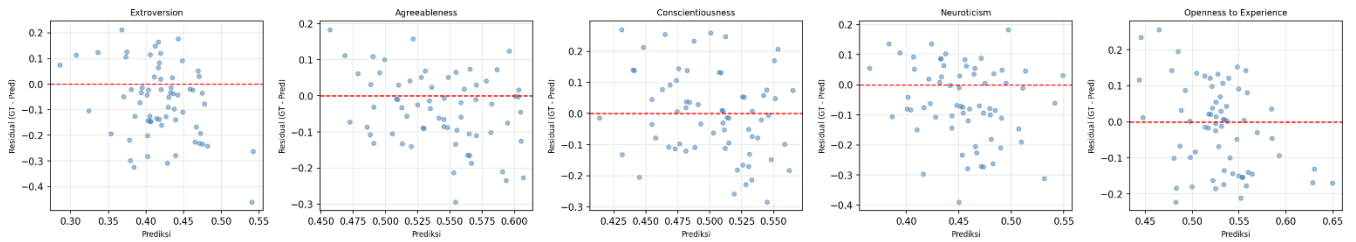
Skenario ini dilakukan untuk mengevaluasi seberapa dekat prediksi model terhadap nilai anotasi kepribadian yang telah divalidasi oleh pakar grafologi. Dua sampel gambar tulisan tangan dipilih secara acak untuk dibandingkan antar nilai prediksi dan nilai *ground truth* pada kelima dimensi kepribadian Big Five (OCEAN), yang ditampilkan pada Tabel V.

TABEL V
HASIL PERBANDINGAN GROUND TRUTH

Nama File	Dimensi Kepribadian	Anotasi	Prediksi	Selisih
77e.jpeg	Openness	0.62	0.47	0.15
	Conscientiousness	0.40	0.41	0.01
	Extraversion	0.28	0.41	0.13
	Agreeableness	0.38	0.51	0.13
	Neuroticism	0.36	0.50	0.14
Rata-Rata Selisih				0.112
382e.jpeg	Openness	0.46	0.44	0.02
	Conscientiousness	0.38	0.47	0.09
	Extraversion	0.46	0.33	0.13
	Agreeableness	0.36	0.49	0.13
	Neuroticism	0.34	0.41	0.07
Rata-Rata Selisih				0.088

C. Skenario Perbandingan Pencahayaan

Skenario ini bertujuan untuk mengetahui kemampuan model terhadap variasi kondisi pencahayaan pada saat pengambilan gambar tulisan tangan. Pengujian dilakukan menggunakan dua kondisi pencahayaan, yaitu gambar dengan pencahayaan terang dan gelap, untuk mengamati pengaruhnya terhadap selisih nilai prediksi dan anotasi. Perbedaan kondisi pencahayaan saat pengambilan gambar dengan kamera memengaruhi kualitas visual, seperti kontras, saturasi warna, dan level *noise* yang secara langsung berdampak pada efektivitas ekstraksi fitur ResNet50. Hasil pengujiannya seperti pada Tabel VI dibawah ini.



Gambar 5. Grafik Residual Plot

TABEL VI
HASIL PERBANDINGAN PENCAHAYAAN

Nama File	Dimensi Kepribadian	Anotasi	Prediksi	Selisih
118e.jpeg (T)	Openness	0.50	0.46	0.04
	Conscientiousness	0.58	0.41	0.17
	Extraversion	0.40	0.46	0.06
	Agreeableness	0.58	0.52	0.06
	Neuroticism	0.58	0.44	0.14
Rata-Rata Selisih				0.094
293e.jpeg (G)	Openness	0.64	0.54	0.10
	Conscientiousness	0.58	0.44	0.14
	Extraversion	0.12	0.40	0.28
	Agreeableness	0.62	0.60	0.02
	Neuroticism	0.56	0.47	0.09
Rata-Rata Selisih				0.126

Catatan: T = Terang, G = Gelap

D. Skenario Perbandingan Latar Belakang

Skenario ini dirancang untuk menganalisis sensitivitas model terhadap perbedaan latar belakang tempat menulis, yaitu kertas polos dan kertas bergaris. Perbedaan latar belakang dapat memengaruhi proses ekstraksi fitur oleh ResNet, sehingga perlu diketahui sejauh mana variasi tersebut berdampak pada hasil prediksi kepribadian yang terlihat pada Table VII.

TABEL VII
HASIL PERBANDINGAN LATAR BELAKANG

Nama File	Dimensi Kepribadian	Anotasi	Prediksi	Selisih
40e.jpeg (P)	Openness	0.62	0.49	0.13
	Conscientiousness	0.60	0.53	0.07
	Extraversion	0.34	0.47	0.13
	Agreeableness	0.66	0.51	0.15
	Neuroticism	0.50	0.39	0.11
Rata-Rata Selisih				0.118
31e.jpeg (B)	Openness	0.58	0.53	0.05
	Conscientiousness	0.46	0.50	0.04
	Extraversion	0.32	0.47	0.15
	Agreeableness	0.54	0.50	0.04
	Neuroticism	0.54	0.47	0.07
Rata-Rata Selisih				0.070

Catatan: P = Polos, B = Bergaris

E. Skenario Perbandingan Media Tulis

Skenario ini menguji apakah perbedaan media penulisan, yaitu tulisan di atas kertas fisik dan tulisan di layar digital memengaruhi akurasi prediksi model. Pengujian ini penting mengingat pengguna aplikasi nantinya dapat menggunakan berbagai media yang tersedia, sehingga model diharapkan mampu bekerja secara konsisten pada kedua kondisi tersebut, sebagaimana hasilnya yang terlihat pada Tabel VIII.

TABEL VIII
HASIL PERBANDINGAN MEDIA TULIS

Nama File	Dimensi Kepribadian	Anotasi	Prediksi	Selisih
310e.jpeg (K)	Openness	0.40	0.54	0.14
	Conscientiousness	0.38	0.51	0.13
	Extraversion	0.36	0.43	0.07
	Agreeableness	0.48	0.55	0.07
	Neuroticism	0.40	0.43	0.03
Rata-Rata Selisih				0.088
320e.jpeg (D)	Openness	0.56	0.44	0.12
	Conscientiousness	0.52	0.49	0.03
	Extraversion	0.26	0.38	0.12
	Agreeableness	0.56	0.49	0.07
	Neuroticism	0.34	0.45	0.11
Rata-Rata Selisih				0.090

Catatan: K = Kertas, D = Digital

F. Analisis Komparatif

Untuk mengevaluasi kestabilan performa model secara menyeluruh, dilakukan analisis komparatif terhadap rata-rata selisih prediksi pada seluruh skenario pengujian, yang dirangkum pada Tabel IX.

TABEL IX
REKAPITULASI RATA-RATA SELISIH PREDIKSI ANTAR SKENARIO

Skenario	Kondisi	Rata-Rata Selisih	Selisih Maksimal (Dimensi)
Pencahayaan	Terang	0.094	0.17 (C)
	Gelap	0.126	0.28 (E)
Latar Belakang	Polos	0.118	0.15 (A)
	Bergaris	0.070	0.15 (E)
Media Tulis	Kertas	0.088	0.14 (O)
	Digital	0.090	0.12 (O, E)

Secara umum, model ResNet50 dengan optimizer Adam menunjukkan performa yang relatif stabil di beberapa kondisi, dengan rata-rata selisih berkisar antara 0,070 hingga

0,126. Hal tersebut masih berada dalam rentang yang wajar untuk tugas regresi kepribadian berbasis citra.

Pada skenario pencahayaan, kondisi terang menghasilkan rata-rata selisih lebih rendah (0,094) dibandingkan kondisi gelap (0,126). Intensitas cahaya yang rendah mereduksi kontras dan *signal-to-noise ratio* (SNR) citra, sehingga membatasi efektivitas ekstraksi fitur oleh ResNet50. Deviasi tertinggi terdeteksi pada dimensi *Extraversion* dalam kondisi gelap dengan selisih 0,28. Fenomena ini selaras dengan grafik *residual plot* yang menunjukkan variasi prediksi tertinggi pada dimensi tersebut, mengindikasikan sensitivitas fitur karakteristik visual *Extraversion* terhadap penurunan kualitas pencahayaan perangkat.

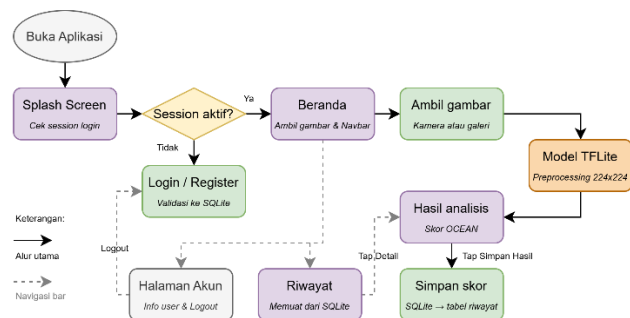
Skenario latar belakang menunjukkan pola yang unik, di mana media kertas bergaris menghasilkan rata-rata selisih lebih rendah (0,070) daripada kertas polos (0,118). Garis horizontal pada kertas terbukti tidak menjadi *noise* dominan yang mengubah bentuk asli atau struktur tulisan tangan setelah melalui *preprocessing*. Hal ini membuktikan bahwa arsitektur ResNet50 mampu mengabaikan gangguan latar belakang yang polanya konsisten dan tetap fokus mengeksplorasi ciri khas tulisan tangan secara akurat.

Pada skenario media tulis, perbedaan antara kertas fisik (0,088) dan media layar digital (0,090) terbilang sangat kecil dan dapat dianggap tidak signifikan secara praktis. Konsistensi performa ini dipengaruhi oleh tahap *resizing* ke 224 x 224 piksel yang menyamakan resolusi *input*, sehingga karakteristik unik masing-masing media seperti, tekstur kertas maupun luminansi layar digital, tereduksi pada tahap *preprocessing*. Ini merupakan indikasi bahwa model berpotensi digunakan secara fleksibel oleh pengguna dengan media yang berbeda-beda.

G. Hasil Pengembangan Sistem

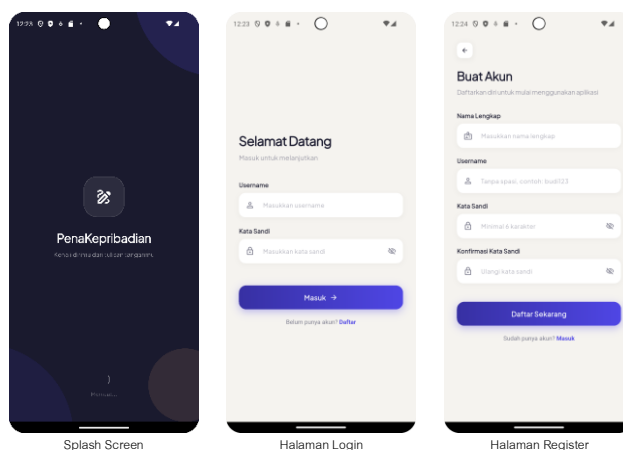
Model dengan kinerja terbaik yang akan diimplementasikan ke dalam aplikasi *mobile* adalah ResNet50 dengan *optimizer* Adam. PenaKepribadian merupakan nama aplikasi untuk memprediksi kepribadian Big Five menggunakan tulisan tangan. Aplikasi tersebut berbasis Android yang dibuat melalui platform Android Studio dengan mengintegrasikan Flutter sebagai *framework* dan SQLite sebagai *database*.

Berdasarkan alur kerja sistem pada Gambar 6, aplikasi PenaKepribadian dirancang untuk dapat digunakan secara terstruktur mulai dari membuka aplikasi hingga penyimpanan hasil analisis ke dalam basis data lokal. Setiap tahapan dalam proses tersebut diimplementasikan secara sistematis guna memastikan alur penggunaan yang konsisten serta memudahkan pengguna dalam mengoperasikan aplikasi.

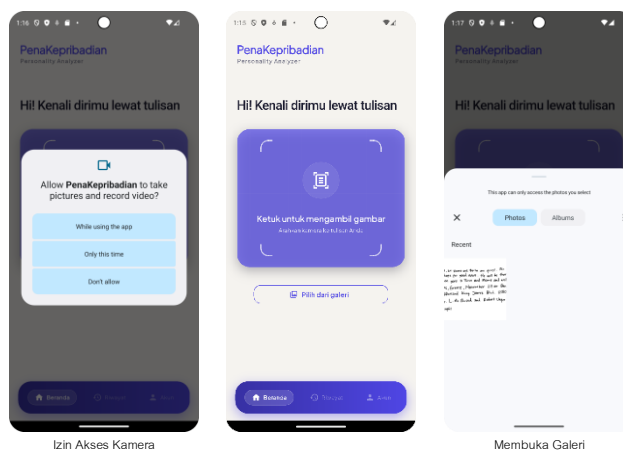


Gambar 6. Bagan Alur Kerja Sistem

Aplikasi ini dirancang dengan antarmuka yang modern untuk memudahkan pengguna dalam melakukan prediksi kepribadian. Saat membuka aplikasi, sistem akan menampilkan ‘*Splash Screen*’ diikuti proses pengecekan sesi login pengguna secara otomatis, seperti pada Gambar 7.

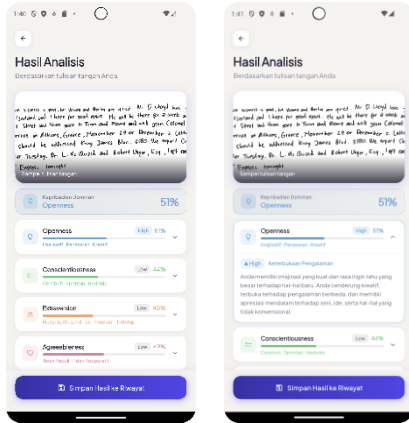


Gambar 7. Halaman Splash Screen, Login, dan Register



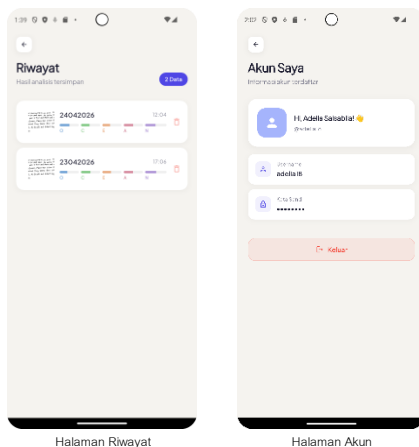
Gambar 8. Halaman Beranda

Setelah berhasil masuk, pengguna diarahkan ke halaman ‘Beranda’ pada Gambar 8. Dari halaman tersebut, pengguna dapat langsung memulai proses analisis baru dengan menekan tombol ‘Ambil Gambar’ untuk membuka kamera atau ‘Pilih dari Galeri’.



Gambar 9. Halaman Hasil Analisis

Gambar tulisan tangan yang dipilih akan diproses oleh model ResNet50 dengan preprocessing berukuran 224x224 piksel menggunakan TensorFlow Lite yang telah diintegrasikan ke dalam aplikasi. Setelah proses prediksi selesai, hasil analisis ditampilkan pada halaman ‘Hasil Analisis’ Gambar 9 yang memuat skor kelima dimensi kepribadian dalam bentuk persentase disertai penjelasan untuk tingkatan rendah atau tinggi nya setiap dimensi. Tingkatan rendah dan tinggi pada setiap dimensi mencerminkan kecenderungan kepribadian pengguna berdasarkan pola tulisan tangan yang terdeteksi oleh model. Dan pengguna dapat memahami karakteristik kepribadiannya melalui deskripsi singkat yang tersedia pada setiap dimensi.



Gambar 10. Halaman Riwayat dan Akun

Pengguna kemudian dapat memilih untuk menyimpan hasil tersebut ke dalam basis data lokal sistem. Seluruh riwayat analisis yang tersimpan dapat diakses melalui halaman ‘Riwayat’ Gambar 10, yang memuat rekap skor kepribadian untuk dapat dipantau oleh pengguna. Informasi akun juga dapat dilihat melalui halaman ‘Akun’, serta dapat melakukan logout untuk mengakhiri sesi dari penggunaan aplikasi.

H. Pengujian Black Box

Black Box Testing merupakan metode pengujian yang memastikan hasil eksekusi aplikasi berdasarkan input yang diberikan untuk memastikan fungsionalitas fitur-fitur aplikasi berjalan sesuai dengan kebutuhan yang telah ditentukan [30]. Pengujian ini dilakukan menggunakan perangkat dengan spesifikasi Chipset MediaTek Dimensity 6100+, GPU Mali-G57 MC2, dan RAM 8/256. Skenario pengujian Black Box dapat dilihat pada Tabel X.

TABEL X
PENGUJIAN BLACK BOX

No	Skenario Pengujian	Hasil	Status
1	Login & Register	Pengguna masuk ke halaman Beranda	Valid
2	Pengambilan gambar (kamera & galeri)	Gambar berhasil ditampilkan	Valid
3	Analisis tulisan tangan	Skor OCEAN tampil	Valid
4	Simpan hasil ke riwayat	Hasil tersimpan ke halaman Riwayat	Valid
5	Logout	Sesi berakhir, pengguna diarahkan ke halaman Login	Valid

I. Pengujian End-User Aplikasi

Pengujian *end-user* dilakukan menggunakan *dataset* uji mandiri yang dikumpulkan melalui penyebaran formulir daring kepada responden. Penelitian ini telah mengutamakan aspek privasi responden dan mendapatkan persetujuan etik dari Komisi Etik Penelitian Universitas Negeri Surabaya dengan Nomor Surat: 057/UN38.16/KEP/EC/2026. Formulir tersebut memuat kuesioner kepribadian BFI-2-S berjumlah 30 pertanyaan dengan skala 1–5 sebagai label acuan (*ground truth*), serta permintaan sampel tulisan tangan. Dari tahapan ini, berhasil dikumpulkan 20 data uji dengan karakteristik media yang beragam, meliputi penggunaan kertas polos, kertas bergaris, hingga tulisan digital dari perangkat layar sentuh. Variasi ini sengaja diterapkan untuk menguji ketahanan dan konsistensi model dalam memprediksi kepribadian pada berbagai kondisi nyata yang berbeda.

Proses evaluasi dilakukan dengan memasukkan seluruh sampel tulisan tangan ke dalam aplikasi, lalu membandingkan hasil prediksi model dengan label acuan dari kuesioner. Guna mempermudah analisis, perbandingan dilakukan menggunakan representasi biner, di mana nilai 0 mewakili

kategori *Low* dan nilai 1 mewakili kategori *High* pada setiap dimensi kepribadian OCEAN. Ketidaksesuaian antara hasil kuesioner dan prediksi model ditandai dengan warna merah untuk mempermudah identifikasi *misclassification*. Hasil perbandingan komparatif dari seluruh 20 data uji tersebut disajikan pada Tabel XI.

TABEL XI
HASIL PERBANDINGAN END-USER

File	Kuisisioner					Model				
	O	C	E	A	N	O	C	E	A	N
User1	0	1	0	0	0	1	1	0	0	0
User2	1	1	0	0	1	1	1	0	0	0
User3	1	1	0	0	0	1	1	0	1	0
User4	1	1	0	1	0	1	1	0	0	0
User5	1	1	0	1	0	0	0	0	1	0
User6	1	1	0	0	1	1	0	0	1	1
User7	0	1	0	0	1	0	1	0	1	0
User8	0	0	0	1	0	1	1	0	1	0
User9	1	1	0	0	0	0	0	0	0	0
User10	1	0	0	0	0	1	1	0	1	0
User11	1	1	0	0	0	1	0	0	1	0
User12	1	1	0	0	1	1	1	0	0	0
User13	1	1	0	0	0	1	0	0	1	0
User14	1	1	0	0	1	0	1	0	0	0
User15	1	0	0	1	0	0	0	0	1	0
User16	1	1	0	0	1	1	0	0	0	0
User17	1	1	0	0	1	1	1	0	0	0
User18	0	1	0	0	0	1	1	0	0	0
User19	1	0	0	0	0	1	0	0	1	0
User20	0	1	0	0	0	1	0	0	0	0

Secara keseluruhan, hasil pengujian *end-user* ini menunjukkan bahwa sistem yang dikembangkan mampu berjalan dengan baik dalam kondisi nyata, meskipun masih terdapat ruang untuk peningkatan terutama pada kemampuan model dalam memprediksi dimensi kepribadian yang memiliki tingkat ketidaksesuaian lebih tinggi dibandingkan dimensi lainnya.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem asesmen kepribadian yang otomatis berbasis aplikasi *mobile* Android bernama PenaKepribadian untuk mendeteksi kepribadian Big Five dari tulisan tangan menggunakan arsitektur model ResNet50. Dari hasil eksperimen perbandingan *optimizer*, Adam terbukti menghasilkan performa terbaik dengan nilai PCC 0,5872 dan akurasi 91,90% pada fase pelatihan, serta PCC 0,4719 dan akurasi 89,95% pada fase pengujian, lebih unggul dari SGD dan RMSprop. Hasil tersebut juga menunjukkan peningkatan yang signifikan dibandingkan penelitian sebelumnya oleh [19] yang memiliki selisih nilai PCC pelatihan lebih tinggi sebesar 0,319 dan PCC pengujian lebih tinggi sebesar 0,173 pada tugas serupa, mengindikasikan bahwa pendekatan *transfer learning*

ResNet50 yang digunakan mampu menghasilkan representasi fitur tulisan tangan yang efektif. Nilai PCC pengujian sebesar 0,4719 menunjukkan korelasi sedang antara prediksi model dengan *ground truth*, yang mengindikasikan bahwa pendekatan ini menjanjikan namun masih memerlukan peningkatan lebih lanjut, terutama untuk mencapai tingkat akurasi yang memadai bagi penggunaan di konteks bertaraf tinggi (*high-stakes*). Selain itu, model tetap menunjukkan performa yang relatif konsisten pada berbagai kondisi pengambilan gambar, meskipun masih ditemukan tantangan pada dimensi Extraversion. Aplikasi yang dikembangkan juga telah berjalan dengan baik berdasarkan pengujian fungsional menggunakan Black Box Testing.

Penelitian selanjutnya disarankan untuk dapat meningkatkan jumlah *dataset* guna memperoleh representasi data yang lebih beragam. Selain itu, eksplorasi arsitektur *deep learning* serta teknik optimasi lainnya diharapkan dapat lebih meningkatkan akurasi prediksi kelima dimensi kepribadian secara merata. Pengembangan sistem ke depan juga harus mengintegrasikan evaluasi terhadap bias model (*fairness*), analisis dampak salah klasifikasi, dan penerapan protokol perlindungan privasi bagi data pengguna. Selanjutnya, pengembangan fitur aplikasi seperti fitur edukatif mengenai literasi kepribadian dapat memberikan dampak yang lebih optimal bagi pengguna.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Dr. Diana Rahmasari, S.Psi., M.Si., Psikolog. atas kesediaan waktu, diskusi, dan saran konstruktifnya terkait aspek psikologi dalam pengembangan sistem serta penyusunan artikel ini.

DAFTAR PUSTAKA

- [1] G. Marano, G. Traversi, E. Gaetani, G. Sani, S. Mazza, dan M. Mazza, "Graphology: An Interface Between Biology, Psychology and Neuroscience," *Psychological Disorders and Research*, vol. 3, no. 3, hlm. 2–13, 2021.
- [2] A. G. Purwanto, R. Y. Wijaya, Timotius, I. J. Agustian, dan I. B. Trisno, "Analisis dan Desain Intelligent Agent Menentukan Kepribadian Berdasarkan Tulisan Tangan," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 2, hlm. 243–247, 2022.
- [3] G. Sheikholeslami, S. N. Srihari, dan V. Govindaraju, "Computer Aided Graphology," Center of Excellence for Document Analysis and Recognition, State University of New York at Buffalo, 1999.
- [4] O. P. John dan S. Srivastava, "The Big Five Traits Taxonomy: History, Measurement, and Theoretical Perspectives," dalam *HANDBOOK OF PERSONALITY Theory and Research*, THE GUILFORD PRESS, 2024, hlm. 102–138.
- [5] S. R. Nudin, R. Gernowo, M. Somantri, dan A. Wibowo, "Multi Task Classification Using Deep Learning Approaches for Big Five Personality Traits Prediction: A Review," dalam *2024 11th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, IEEE, 2024, hlm. 37–42.
- [6] E. Dimitriou dan M. Galanakis, "Organizational Psychology Re-Invented—The Big Five Personality Traits Model as a Reliable Behavior Framework in the Workplace," *Psychology*, vol. 13, no. 05, hlm. 798–804, 2022.
- [7] F. Celli dan B. Lepri, "Is Big Five better than MBTI?," dalam *Proceedings of the Fifth Italian Conference on Computational*

- Linguistics CLiC-it 2018*, Accademia University Press, 2018, hlm. 93–98.
- [8] B. Gawda, "Lack of evidence for the assessment of personality traits using handwriting analysis," *Polish Psychological Bulletin*, vol. 45, no. 1, hlm. 73–79, 2014, doi: 10.2478/ppb-2014-0011.
- [9] A. Vinciarelli dan G. Mohammadi, "A Survey of Personality Computing," *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, hlm. 273–291, 2014, doi: 10.1109/TAFFC.2014.2330816.
- [10] Samsuryadi, R. Kurniawan, J. Supardi, Sukemi, dan F. S. Mohamad, "A Framework for Determining the Big Five Personality Traits Using Machine Learning Classification through Graphology," *Journal of Electrical and Computer Engineering*, vol. 2023, no. 1, hlm. 1–15, 2023.
- [11] S.-H. Noh, "Performance Comparison of CNN Models Using Gradient Flow Analysis," *Informatics*, vol. 8, no. 3, hlm. 53, 2021.
- [12] S. Patnaik, S. Kumari, dan S. Das Mahapatra, "Comparison of deep CNN and ResNet for Handwritten Devanagari Character Recognition," dalam *2020 IEEE International Conference for Convergence in Engineering*, IEEE, 2020, hlm. 235–238.
- [13] J. E. Valdez-Rodriguez, H. Calvo, dan E. M. Felipe-Riverón, "Handwritten Texts for Personality Identification Using Convolutional Neural Networks," dalam *Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, 2019, hlm. 140–145.
- [14] B. S. Puttaswamy dan N. Thillaiarasu, "Fine DenseNet based human personality recognition using english hand writing of non-native speakers," *Biomed. Signal Process. Control*, vol. 99, hlm. 106910, 2025.
- [15] L. Hakim dan Hendro, "Penentuan Tes Kepribadian Calon Mahasiswa Berdasarkan Sidik Jari Menggunakan Minutiae dan Support Vector Machine," *Journal of Applied Informatics and Computing (JAIC)*, vol. 4, no. 1, hlm. 28–32, 2020.
- [16] N. A. Shiferaw *dkk.*, "Handwritten Amharic Character Recognition Through Transfer Learning: Integrating CNN Models and Machine Learning Classifiers," *IEEE Access*, vol. 13, hlm. 52134–52148, 2025.
- [17] I. Santoso, A. M. Manurung, dan E. R. Subhiyako, "Comparison of ResNet-50, EfficientNet-B1, and VGG-16 Algorithms for Cataract Eye Image Classification," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 2, hlm. 284–294, 2025.
- [18] C. Schröder, F. Kruse, dan J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," dalam *Procedia Computer Science 181*, Elsevier B.V., 2021, hlm. 526–534.
- [19] S. Checker, M. Yadav, dan R. Katarya, "HiEnWrite: A Hindi-English Bilingual Dataset for Big Five Personality Detection," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 24, no. 9, hlm. 1–19, 2025, doi: 10.1145/3756010.
- [20] S. Mukherjee, "The Annotated ResNet-50," Towards Data Science.
- [21] R. Zaheer dan H. Shaziya, "A Study of the Optimization Algorithms in Deep Learning," dalam *Proceedings of the 3rd International Conference on Inventive Systems and Control, ICISC 2019*, Institute of Electrical and Electronics Engineers Inc., 2019, hlm. 536–539.
- [22] B. Nugroho, E. Y. Puspaningrum, dan M. S. Munir, "Kinerja Algoritma Optimasi Root-Mean-Square Propagation dan Stochastic Gradient Descent pada Klasifikasi Pneumonia Covid-19 Menggunakan CNN," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 7, no. 3, hlm. 420–425, 2021.
- [23] D. Altunel, "Development of Deep Learning Optimizers: Approaches, Concepts, and Update Rules," *arXiv preprint*, 2025.
- [24] G. A. Saputra dan I. M. A. Agastya, "Beta Fish Identification System Based On Convolutional Neural Network," *Journal of Applied Informatics and Computing (JAIC)*, vol. 8, no. 2, hlm. 443–452, 2024.
- [25] M. O. G. Sibannang dan I. G. L. Wijayakusuma, "Implementation of LSTM for Gold Price Prediction in Indonesia," *Journal of Applied Informatics and Computing (JAIC)*, vol. 10, no. 1, hlm. 2548–6861, 2026.
- [26] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," *Geosci. Model Dev.*, vol. 15, no. 14, hlm. 5481–5487, 2022.
- [27] V. K. A. Zebua dan Rusdah, "Forecasting Modeling of Imported Goods Release Waiting Time in Importer Logistics Operations Using Multiple Linear Regression," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 4, hlm. 1878–1885, 2025.
- [28] P. Schober, C. Boer, dan L. A. Schwarte, "Correlation coefficients: Appropriate use and interpretation," *Anesth. Analg.*, vol. 126, no. 5, hlm. 1763–1768, 2018.
- [29] C. Putry, V. Paramita, dan I. Hartati, "Implementation of Ambon Banana Peel (*Musa Paradisiaca*) as Foaming Agent of Banana Fruitghurt using Foam Mat Drying Method," *Journal of Vocational Studies on Applied Research*, vol. 4, no. 2, hlm. 40–46, 2022.
- [30] G. S. Mahendra dan I. K. A. Asmarajaya, "Evaluation Using Black Box Testing and System Usability Scale in the Kidung Sekar Madya Application," *Sinkron: Jurnal dan Penelitian Teknik Informatika*, vol. 6, no. 4, hlm. 2292–2302, 2022.