

Superioritas Optimizer AdaGrad pada Arsitektur Spasio-Temporal PolyFace-LSTM untuk Prediksi Kepribadian Berbasis Video

The Superiority of the AdaGrad Optimizer in the Spatio-Temporal PolyFace-LSTM Architecture for Video-Based Personality Prediction

Muhammad Zidan Alif Oktavian*, Salamun Rohman Nudin

Fakultas Vokasi, Program Studi Manajemen Informatika, Universitas Negeri Surabaya, Surabaya, Indonesia

Email: ^{1,*}muhammadzidan.21045@mhs.unesa.ac.id, ²salamunrohman@unesa.ac.id

Email Penulis Korespondensi: muhammadzidan.21045@mhs.unesa.ac.id

Received: 08/08/2026, Revised: 30/08/2026, Accepted: 02/09/2026, Published: 08/09/2026

Abstrak—Pengukuran kepribadian secara konvensional umumnya menggunakan kuesioner psikometri, seperti Big Five Inventory (BFI). Namun, metode ini rentan terhadap bias penilaian diri (self-report bias) serta membutuhkan waktu evaluasi yang cukup lama. Di sisi lain, pengembangan sistem analisis kepribadian otomatis berbasis video menghadapi tantangan komputasi yang tinggi dalam mengintegrasikan fitur spasial wajah dan dinamika temporal mikro-ekspresi secara bersamaan. Penelitian ini bertujuan untuk mengusulkan solusi berupa arsitektur jaringan saraf spasio-temporal hibrida dengan memanfaatkan teknik transfer learning guna memprediksi kepribadian secara lebih objektif. Sistem ini mengintegrasikan algoritma Multi-Task Cascaded Convolutional Networks (MTCNN) untuk deteksi wajah secara otomatis, model PolyFace sebagai pengekstraksi fitur spasial, dan algoritma Long Short-Term Memory (LSTM) untuk memodelkan hubungan temporal antar-frame. Eksperimen dilakukan dengan membandingkan tiga algoritma optimasi, yaitu Adam, AdaGrad, dan SGD, pada dataset ChaLearn LAP 2017. Hasil pengujian menunjukkan bahwa optimizer AdaGrad memberikan performa generalisasi dan prediksi terbaik pada fase validasi dengan rata-rata akurasi sebesar 89,13%, mengungguli algoritma SGD (88,34%) dan Adam (88,33%). Secara teoritis dan matematis, superioritas AdaGrad pada arsitektur ini didorong oleh kemampuannya menyesuaikan laju pembelajaran secara dinamis berdasarkan akumulasi gradien historis. Mekanisme ini membuat AdaGrad jauh lebih stabil dalam mengekstrak fitur spasio-temporal mikro-ekspresi yang kompleks tanpa terjebak pada minimum lokal (local minima). Kinerja tertinggi yang dihasilkan oleh AdaGrad terdapat pada dimensi Agreeableness yang mencapai tingkat akurasi hingga 90,14%. Kontribusi utama dari penelitian ini adalah menghasilkan model prediksi kepribadian spasio-temporal yang efisien (latensi rendah), presisi, dan terhindar dari jebakan minimum lokal (*local minima*) untuk diimplementasikan pada sistem otomatisasi deteksi psikologis nyata.

Kata Kunci: Kepribadian; MTCNN; PolyFace; LSTM; Optimizer

Abstract—Conventional personality assessment generally relies on psychometric questionnaires, such as the Big Five Inventory (BFI). However, this method is susceptible to self-report bias and requires a relatively long evaluation time. On the other hand, the development of automated video-based personality analysis systems faces high computational challenges in simultaneously integrating facial spatial features and the temporal dynamics of micro-expressions. This study aims to propose a solution in the form of a hybrid spatio-temporal neural network architecture utilizing transfer learning techniques to predict personality more objectively. The proposed system integrates the Multi-Task Cascaded Convolutional Networks (MTCNN) algorithm for automatic face detection, the PolyFace model as a spatial feature extractor, and the Long Short-Term Memory (LSTM) algorithm to model temporal relationships between frames. Experiments were conducted by comparing three optimization algorithms, namely Adam, AdaGrad, and SGD, using the ChaLearn LAP 2017 dataset. The results show that the AdaGrad optimizer achieved the best generalization and prediction performance during the validation phase, with an average accuracy of 89.13%, outperforming SGD (88.34%) and Adam (88.33%). Theoretically and mathematically, the superiority of AdaGrad in this architecture is attributed to its ability to dynamically adjust the learning rate based on the accumulation of historical gradients. This mechanism makes AdaGrad considerably more stable in extracting complex spatio-temporal micro-expression features without becoming trapped in local minima. The highest performance achieved by AdaGrad was observed in the Agreeableness dimension, reaching an accuracy of 90.14%. The main contribution of this study is the development of an efficient (low-latency), precise spatio-temporal personality prediction model that avoids local minima, making it suitable for implementation in real-world automated psychological detection systems.

Keywords: Personality; MTCNN; PolyFace; LSTM; Optimizer

1. PENDAHULUAN

Kepribadian merupakan aspek fundamental dalam disiplin ilmu psikologi yang menggambarkan pola perilaku, pemikiran, dan emosi yang menetap pada seorang individu. Model pengukuran yang paling diakui dan digunakan secara luas di seluruh dunia untuk memahami karakteristik manusia adalah Big Five Personality atau yang sering disingkat dengan OCEAN [1]. Model ini mengklasifikasikan kepribadian ke dalam lima dimensi utama, yakni Openness to Experience (keterbukaan terhadap pengalaman baru), Conscientiousness (ketelitian dan tanggung jawab), Extraversion (tingkat sosialisasi dan energi), Agreeableness (kemudahan bergaul dan empati), serta Neuroticism (stabilitas emosional). Sebelum maraknya pemanfaatan kecerdasan buatan, proses ekstraksi kelima dimensi kepribadian ini secara konvensional bergantung pada instrumen kuesioner psikometri berbasis teks yang dievaluasi secara manual [2].

Meskipun instrumen kuesioner ini memiliki tingkat validitas yang teruji, penerapannya secara manual memiliki kelemahan yang signifikan. Pendekatan konvensional ini sangat rentan terhadap bias penilaian diri sendiri (self-report

bias), di mana individu cenderung memberikan jawaban yang dianggap ideal secara sosial alih-alih merepresentasikan kondisi psikologis mereka yang sebenarnya [3]. Selain itu, pengisian kuesioner memakan waktu dan berisiko menimbulkan kelelahan responden (survey fatigue) [4]. Inilah yang menjadi masalah utama dalam penelitian ini, yaitu kebutuhan akan sistem evaluasi kepribadian yang lebih objektif, otomatis, dan minim intervensi sadar dari subjek.

Sebagai solusi dari permasalahan tersebut, seiring dengan pesatnya perkembangan teknologi komputasi dan kecerdasan buatan, upaya prediksi kepribadian mulai bergeser menggunakan pendekatan Computer Vision dan Deep Learning berbasis data video [5]. Video mampu merekam dinamika mikro-ekspresi wajah secara bawah sadar yang sangat sulit dimanipulasi oleh manusia [6]. Namun, menganalisis data video membawa tantangan komputasi yang jauh lebih besar. Sistem tidak hanya harus mampu mengekstrak informasi spasial (bentuk dan fitur wajah pada satu waktu tertentu), tetapi juga wajib memproses informasi temporal (perubahan dan pergerakan ekspresi wajah dari waktu ke waktu) secara berurutan.

Oleh karena itu, penelitian ini mengusulkan solusi berupa pembangunan arsitektur jaringan saraf spasio-temporal hibrida. Sistem memanfaatkan algoritma Multi-Task Cascaded Convolutional Networks (MTCNN) [7] untuk mendeteksi dan memotong wajah secara presisi di berbagai kondisi [8], serta menggunakan model transfer learning PolyFace sebagai pengekstraksi fitur spasial. Fitur dari PolyFace kemudian didistribusikan ke dalam algoritma Long Short-Term Memory (LSTM) untuk memodelkan dinamika temporalnya. Meskipun LSTM umumnya memakan memori yang besar, sistem ini dirancang ringan (lightweight) dengan melakukan down-sampling ekstrem, di mana sistem hanya mengekstrak 10 frame dari total ratusan frame dalam video berdurasi 15 detik, sehingga beban pemrosesan sekuensial menurun drastis. Lebih dari itu, karena sistem ini dirancang untuk memprediksi lima skor kontinu secara bersamaan (multi-output regression), keberhasilan pelatihan model sangat bergantung pada pemilihan algoritma optimasi (optimizer) yang tepat untuk memperbarui parameter bobot (weights) dan meminimalkan nilai kerugian (loss) [9].

Berbagai penelitian terdahulu telah berupaya mengembangkan sistem otomatisasi prediksi kepribadian, namun sebagian besar memiliki fokus yang berbeda. Penelitian yang dilakukan oleh Suman dkk. (2022) mengembangkan sistem prediksi kepribadian berbasis Deep Learning menggunakan dataset ChaLearn-17. Namun, penelitian tersebut masih mengandalkan backbone ekstraksi fitur wajah konvensional yang masif, seperti VGGFace atau ResNet [10]. Meskipun mencapai akurasi yang tinggi, penggunaan model ekstraksi tersebut membutuhkan biaya komputasi yang teramat besar dan tidak efisien untuk diterapkan pada perangkat keras standar. Sebagai perbandingan, studi membuktikan bahwa arsitektur PolyFace merupakan model yang sangat lightweight secara matematis karena dirancang khusus untuk memenuhi anggaran komputasi ketat di bawah 30 GFlops pada ajang ICCV19 - Lightweight Face Recognition Challenge [11]. Di sisi lain, penelitian yang dilakukan oleh Escalante dkk. (2022) berfokus pada prediksi apparent personality dari video wawancara kerja [12]. Mereka menitikberatkan pada aspek interpretabilitas model agar hasil rekomendasi AI dapat dijelaskan secara visual, namun tidak melakukan kajian mendalam terkait optimalisasi arsitektur secara matematis selama fase pelatihan.

Penelitian selanjutnya yang relevan dilakukan oleh Ouarka dkk. (2024), yang menggunakan pendekatan deep multimodal fusion dengan menggabungkan arsitektur masif seperti Vision Transformer [13] (ViT-B16) dan CNN untuk memprediksi kepribadian. Sayangnya, penggunaan model transformer menuntut spesifikasi perangkat yang sangat ekstrem sehingga sulit diimplementasikan pada aplikasi web yang ringan. Penelitian berikutnya oleh Anggara dkk. (2023) menunjukkan keberhasilan algoritma MTCNN dalam melakukan autentikasi pengenalan wajah pada berbagai kondisi pencahayaan dengan akurasi 84% [14].

Penelitian ini membuktikan ketangguhan MTCNN, namun penerapannya dibatasi hanya untuk klasifikasi biner identitas, bukan prediksi atribut psikologis kontinyu. Berdasarkan tinjauan literatur dari semua penelitian tersebut, terdapat sebuah gap analysis (kesenjangan penelitian) yang sangat jelas. Sebagian besar penelitian kecerdasan buatan sebelumnya berlomba-lomba menggunakan backbone visual yang sangat berat dan kompleks seperti Vision Transformer atau VGGFace semata-mata untuk mengejar nilai akurasi tanpa mempertimbangkan efisiensi komputasi. Mengadopsi keunggulan PolyFace, penelitian ini mengusulkan sebuah pembaruan arsitektur spasio-temporal hibrida dengan mencabut backbone berat dari arsitektur pendahulu (seperti Suman dkk.) dan menggantikannya dengan PolyFace melalui pendekatan transfer learning.

Hingga saat ini, belum ada studi spesifik yang membandingkan efektivitas algoritma optimasi dalam menyelesaikan masalah multi-output regression untuk memprediksi kelima dimensi OCEAN menggunakan arsitektur hibrida ini. Karakteristik data psikologis dan mikro-ekspresi wajah memiliki lanskap kerugian (loss landscape) yang sangat berisik (noisy) dan penuh dengan minimum lokal. Pada kondisi ini, algoritma dengan momentum adaptif tinggi seperti Adam rentan mengalami overshoot (melompat terlalu jauh dari titik optimal), sehingga performanya sering kali stagnan atau gagal. Oleh karena itu, tujuan utama dari penelitian ini adalah untuk melakukan komparasi kinerja secara komprehensif antara tiga algoritma optimizer populer, yaitu Adam (Adaptive Moment Estimation), AdaGrad (Adaptive Gradient Algorithm), dan SGD (Stochastic Gradient Descent) [15] pada arsitektur PolyFace-LSTM.

Adapun kontribusi utama dari penelitian ini adalah menghasilkan evaluasi objektif mengenai optimizer yang paling superior dalam menyeimbangkan stabilitas pelatihan dan kecepatan konvergensi, serta menyediakan model prediksi kepribadian spasio-temporal yang efisien (latensi rendah), akurat, dan terhindar dari jebakan minimum lokal (local minima).

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian



Gambar 1 Skema Penelitian

Alur penelitian ini terbagi ke dalam tiga fase utama sebagaimana diilustrasikan pada Gambar 1. Pertama, fase Persiapan Data meliputi akuisisi 10.000 klip video dari dataset ChaLearn LAP 2017 serta prapemrosesan spasio-temporal berupa pelokalan wajah menggunakan MTCNN (resolusi 112×112 piksel) dan ekstraksi 10 frame.

Kedua, fase Pemodelan & Optimisasi berfokus pada ekstraksi fitur hibrida (frozen weights PolyFace dan LSTM) serta pelatihan model regresi dimensi OCEAN menggunakan tiga optimizer (Adam, AdaGrad, dan SGD). Ketiga, fase Evaluasi & Analisis Akhir dilakukan melalui pengujian metrik MAE dan analisis perbandingan kinerja konvergensi masing-masing optimizer dalam mengatasi jebakan local minima.

2.2 Akuisisi Dataset Video

Penelitian ini mengakuisisi dataset video First Impressions yang dikembangkan untuk tantangan ChaLearn Looking at People (LAP) [16]. Dataset memuat 10.000 klip video HD (durasi rata-rata 15 detik) dari 3.000 video YouTube yang menampilkan individu berbicara menghadap kamera, di mana setiap klip telah dianotasi dengan skor kontinu lima dimensi kepribadian OCEAN sebagai *ground truth*. Dataset dipartisi menjadi data latih, validasi, dan pengujian dengan rasio 3:1:1 (6.000, 2.000, dan 2.000 data). Subjek mencakup keragaman demografis gender, variasi etnis (Asia, Kaukasia, Afrika-Amerika), serta delapan rentang kelompok usia (0 hingga 61+ tahun) untuk menjamin kemampuan generalisasi model terhadap berbagai variasi visual.

2.3 Prapemrosesan Spasio-Temporal

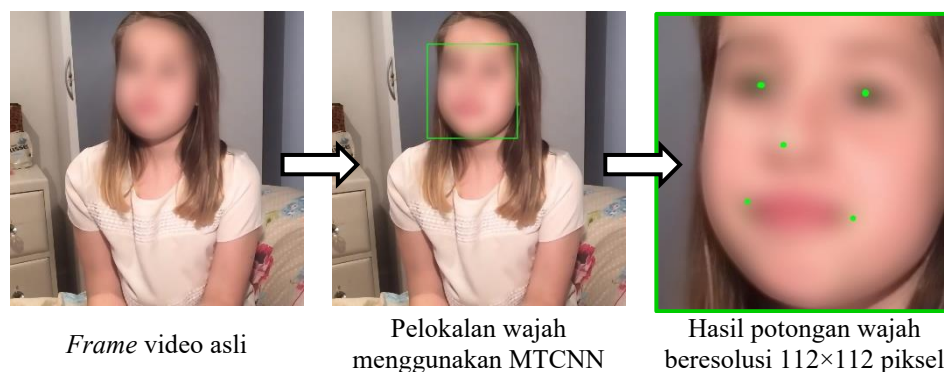
Tahap prapemrosesan spasio-temporal (data preparation) merupakan fase krusial guna memastikan data video mentah dapat ditransformasikan menjadi representasi matriks yang siap diproses secara optimal oleh arsitektur *Deep Learning*. Proses ini mencakup serangkaian tahapan sekuensial:

a. Ekstraksi Frame Video

Tahap awal ini bertujuan untuk mengambil cuplikan gambar statis (frame) dari video input menggunakan pustaka OpenCV [17]. Untuk mendapatkan representasi visual yang komprehensif tanpa membebani komputasi, sistem melakukan down-sampling dengan mengekstrak tepat 10 frame secara konstan berdasarkan interval durasi yang telah ditentukan. Apabila durasi video terlalu singkat, sistem akan secara otomatis melakukan zero-padding (penambahan matriks bernilai nol) guna mempertahankan konsistensi dimensi matriks data.

b. Deteksi dan Pemotongan Wajah

Setelah ekstraksi frame selesai, langkah berikutnya adalah menghilangkan gangguan latar belakang (background noise). Sistem memanfaatkan algoritma MTCNN (Multi-task Cascaded Convolutional Networks) [18] untuk melokalisasi posisi wajah pada setiap frame. Area di dalam koordinat bounding box kemudian dipotong (crop) dan diskalakan ulang (resize) menjadi ukuran standar 112×112 piksel menyesuaikan spesifikasi dimensi input arsitektur *backbone* PolyFace [11] sebagaimana ditunjukkan pada Gambar 2.



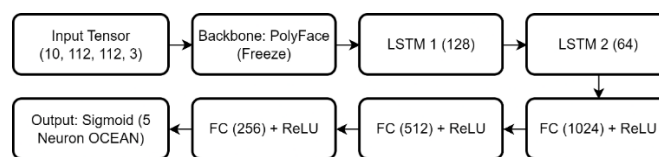
Gambar 2 Tahapan prapemrosesan visual menggunakan MTCNN

- c. Integrasi Anotasi Label OCEAN
Bersamaan dengan pemrosesan visual, sistem juga memuat data ground truth berupa anotasi label kepribadian OCEAN berformat ekstensi .pkl (Pickle). Data direpresentasikan dalam bentuk array yang memuat lima nilai probabilitas untuk masing-masing subjek.
- d. Penyusunan Format Dataset TensorFlow
Guna mengoptimalkan memori komputasi selama fase pelatihan, data dikompilasi memanfaatkan struktur `tf.data.Dataset` yang menghasilkan tensor spasio-temporal berdimensi (10, 112, 112, 3) untuk gambar dan (5,) untuk label.

2.4 Ekstraksi Fitur Hibrida

Fase ini difokuskan pada perancangan arsitektur jaringan saraf tiruan spasio-temporal menggunakan bahasa pemrograman Python dengan kerangka kerja utama PyTorch untuk pemodelan arsitektur serta TensorFlow untuk manajemen pemrosesan data. Tahapannya meliputi:

- a. Pemuatan Pustaka (Import Library)
Dimulai dengan mendeklarasikan pustaka relevan. PyTorch difungsikan sebagai fondasi arsitektur utama karena fleksibilitasnya dalam komputasi grafis [19], sementara TensorFlow dipanggil khusus untuk membaca dataset spasio-temporal secara terdistribusi [20].
- b. Pemuatan Dataset (Load Dataset)
Kelas iterator `TFSnapshotToTorch` dibangun untuk mengekstrak data dari `tf.data.Dataset`, mengubahnya menjadi matriks Numpy, dan memuatnya menggunakan `DataLoader` dengan ukuran batch 4.
- c. Pemuatan *Backbone* PolyFace
Sistem memanggil arsitektur PolyFace sebagai backbone pengekstraksi fitur spasial tunggal. Parameter bobot jaringan PolyFace ini dibekukan (frozen weights) dengan menonaktifkan komputasi gradien selama fase pelatihan. Mekanisme pembekuan ini memblokir rantai propagasi balik (backpropagation) menuju layer ekstraksi spasial. Dampaknya terhadap propagasi gradien pada layer LSTM di atasnya sangat krusial: LSTM kini menerima representasi fitur spasial yang statis dan deterministik pada setiap epoch. Hal ini mengisolasi proses pembaruan gradien agar murni difokuskan pada pembelajaran dependensi temporal di dalam memori LSTM tanpa terganggu oleh pergeseran bobot konvolusional di bawahnya. Selain memangkas jumlah parameter yang dilatih (trainable parameters) secara drastis, pemotongan rantai gradien ini efektif mencegah terjadinya ketidakstabilan propagasi (vanishing atau exploding gradient) yang umum terjadi saat melatih arsitektur CNN-RNN secara bersamaan (end-to-end). Dengan demikian, klaim arsitektur yang lightweight dan stabil secara matematis dapat tervalidasi.
- d. Integrasi Arsitektur Spasio-Temporal
Fitur visual dari PolyFace diteruskan ke lapisan LSTM (Long Short-Term Memory) [21] ganda (stacked) dengan ukuran hidden state 128 dan 64. Keluaran time-step terakhir dilewatkan ke lapisan Fully Connected (1024, 512, 256) menggunakan aktivasi ReLU, Dropout (0.2 dan 0.5) guna mencegah terjadinya overfitting pada jaringan saraf yang kompleks [22], dan diakhiri lima neuron output beraktivasi Sigmoid seperti yang diilustrasikan pada Gambar 3.



Gambar 3 Arsitektur jaringan saraf tiruan spasio-temporal PolyFace-LSTM

- e. Kompilasi Hyperparameter
Pada tahap ini skenario komparasi diterapkan. Konfigurasi parameter pelatihan diatur sedemikian rupa untuk memastikan laju konvergensi yang stabil. Proses pembaruan bobot (*weights*) dikelola secara terpisah menggunakan tiga *optimizer* berbeda, yakni Adam, SGD, dan AdaGrad. *Mean Squared Error* (MSE) ditetapkan sebagai fungsi objektif utama (*loss function*) yang harus diminimalkan. Secara keseluruhan, ringkasan *hyperparameter* arsitektur dan parameter pelatihan model yang digunakan ditunjukkan pada Tabel 1.

Tabel 1 Ringkasan Hyperparameter Arsitektur dan Pelatihan Model

Komponen / Parameter	Konfigurasi
<i>Backbone</i> Fitur Spasial	PolyFace (Frozen Weights)
Arsitektur Sekuensial	Stacked LSTM (128, 64)
Lapisan <i>Fully Connected</i>	1024, 512, 256
Fungsi Aktivasi	ReLU, Sigmoid (Output Layer)
Dropout Rate	0.2 dan 0.5
Ukuran Batch (<i>Batch Size</i>)	4
Maksimum Iterasi (Epoch)	100
<i>Optimizer</i> (Skenario Komparasi)	Adam, SGD, AdaGrad

Komponen / Parameter	Konfigurasi
Fungsi Kerugian (<i>Loss Function</i>)	Mean Squared Error (MSE)
Early Stopping	True
Patience	10 Epoch

f. Siklus Pelatihan (Training Loop)

Proses pelatihan iteratif dieksekusi dengan batas maksimum 100 epoch. Mekanisme Model Checkpoint menyimpan bobot terbaik secara otomatis, didukung algoritma Early Stopping dengan patience 10 epoch.

2.5 Pelatihan dan Komparasi Optimizer

Tahap pelatihan dan komparasi optimizer merupakan fase krusial di mana parameter bobot jaringan diperbarui secara matematis selama proses propagasi balik (backpropagation) menggunakan algoritma optimasi yang dievaluasi. Pada tahap ini, eksperimen pelatihan dikelola secara terpisah menggunakan tiga algoritma optimasi, yaitu Adam (Adaptive Moment Estimation) yang unggul dalam perhitungan bias-koreksi [23], AdaGrad (Adaptive Gradient Algorithm) yang secara dinamis menyesuaikan laju pembelajaran berdasarkan akumulasi gradien historis [24], serta SGD (Stochastic Gradient Descent). Ketiga algoritma ini dievaluasi secara komparatif untuk meminimalkan fungsi kerugian *Mean Squared Error* (MSE), yang bertugas menghitung rata-rata kuadrat selisih antara nilai probabilitas prediksi kelima dimensi kepribadian (OCEAN) dengan label anotasi asli (*ground truth*).

Selain penerapan algoritma optimasi, tahap ini juga mengatur mekanisme pengendalian siklus pelatihan (training loop) yang dieksekusi dengan batas maksimum 100 epoch. Guna melindungi model dari risiko overfitting, sistem mengimplementasikan mekanisme Early Stopping dengan batas toleransi (patience) sebanyak 10 epoch, yang akan menghentikan pelatihan secara otomatis jika metrik kerugian validasi tidak menunjukkan perbaikan. Bersamaan dengan itu, fitur Model Checkpoint digunakan untuk memantau dan menyimpan berkas parameter bobot jaringan terbaik (.pth) yang dihasilkan selama siklus tersebut berlangsung.

2.6 Evaluasi Kinerja Regresi

Tahap ini bertujuan mengukur tingkat keandalan arsitektur PolyFaceOceanModel dalam skenario regresi *multi-output*:

- Perhitungan Fungsi Kerugian dan Kesalahan Prediksi: Evaluasi dasar dilakukan menggunakan fungsi kerugian utama Mean Squared Error (MSE) sebagai objektif pelatihan, dan metrik Mean Absolute Error (MAE) yang secara komparatif lebih tangguh terhadap data pencilan (outliers) [25] untuk mengukur deviasi absolut setelah model dialihkan ke mode evaluasi (`model.eval()`). Nilai MSE dan MAE dihitung menggunakan Persamaan (1) dan Persamaan (2).

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

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

Berdasarkan fungsi komputasi pada persamaan (1) dan (2) tersebut, variabel n merepresentasikan jumlah total data pengujian secara keseluruhan. Sementara itu, variabel y_i menunjukkan nilai probabilitas asli (*ground truth*) dari label dimensi kepribadian OCEAN, dan variabel $\hat{y}_i$ merujuk pada nilai probabilitas hasil prediksi yang secara dinamis dihasilkan oleh arsitektur model selama fase evaluasi.

- Akurasi Spesifik Dimensi OCEAN Sistem mengeksekusi fungsi kustom menggunakan parameter `multioutput='raw_values'` dari pustaka Scikit-Learn [26], sebuah pustaka standar industri untuk evaluasi pembelajaran mesin, untuk memisahkan hasil evaluasi kelima dimensi secara independen.
- Eksekusi Pengujian Metrik dieksekusi berurutan pada partisi *Training*, *Validation*, dan *Testing* untuk menghasilkan perbandingan kinerja *optimizer* secara utuh.

2.7 Analisis Perbandingan Optimasi

Tahap analisis perbandingan merupakan langkah akhir dari kerangka metodologi yang bertujuan untuk membedah kinerja algoritma Adam, SGD, dan AdaGrad secara komparatif berdasarkan tingkat pencapaian akurasi prediksi, kestabilan laju konvergensi kurva kerugian, serta perolehan nilai Mean Absolute Error (MAE) terendah. Analisis ini meninjau metrik secara agregat sekaligus membandingkan ketangguhan masing-masing algoritma dalam menghindari jebakan minimum lokal (*local minima*) saat memproses kompleksitas fitur matriks spasio-temporal wajah. Hasil rumusan komparasi empiris ini akan menjadi landasan objektif untuk menentukan algoritma optimasi yang paling superior dalam menyelesaikan kasus regresi multi-output dimensi OCEAN, di mana model dengan parameter bobot terbaik tersebut selanjutnya ditetapkan sebagai basis arsitektur final untuk diimplementasikan sebagai instrumen pengukuran psikologi kepribadian yang otomatis dan objektif.

3. HASIL DAN PEMBAHASAN

Bagian ini memaparkan secara komprehensif temuan yang diperoleh dari rangkaian eksperimen pengembangan arsitektur spasio-temporal hibrida, pengujian ketahanan (robustness) model terhadap berbagai skenario input video,

serta hasil implementasi akhir ke dalam antarmuka berbasis web. Pemaparan disusun secara sistematis untuk menjawab seluruh rumusan masalah terkait keandalan prediksi kepribadian (OCEAN) menggunakan kombinasi metode MTCNN, PolyFace, dan LSTM.

3.1 Hasil Preprocessing Data

Tahap prapemrosesan berhasil mentransformasi data video mentah First Impressions menjadi tensor spasio-temporal yang siap latih. Algoritma MTCNN efektif mengeliminasi background noise dengan melokalisasi dan memotong area wajah ke resolusi standar 112×112 piksel di berbagai variasi pose dan pencahayaan, didukung mekanisme zero-padding untuk menjaga konsistensi dimensi temporal sebagaimana ditunjukkan pada Gambar 2.

Dari sisi pemrosesan label, data anotasi probabilitas OCEAN beserta variabel tambahan lainnya yang tersimpan di dalam format *pickled dictionaries* (.pkl) telah berhasil diekstrak dan dipetakan agar sejajar dengan urutan data visual. Melalui integrasi menggunakan pustaka `tf.data.Dataset`, sistem berhasil mengompilasi pasangan data tersebut menjadi format tensor akhir. Matriks fitur visual (input) memiliki dimensi (10, 112, 112, 3) yang merepresentasikan 10 frame berurutan, resolusi 112×112 piksel, dengan 3 saluran warna (*channels*) RGB. Sementara itu, vektor target (output) memiliki dimensi (5,) yang merepresentasikan skor kontinuitas lima sifat kepribadian Big Five.

Total keseluruhan dataset berjumlah 10.000 data spasio-temporal. Sesuai dengan skenario pengujian, dataset tersebut kemudian dipartisi secara terstruktur untuk mencegah terjadinya bias evaluasi model. Distribusi pembagian dataset akhir untuk tahapan pelatihan dan pengujian ditunjukkan pada Tabel 2.

Tabel 2 Distribusi Partisi Dataset Spasio-Temporal

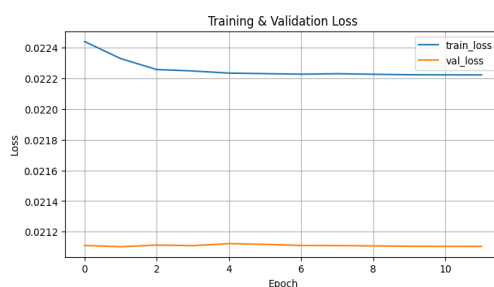
Partisi Data	Jumlah Tensor
Data Latih (Training Set)	6,000
Data Validasi (Validation Set)	2,000
Data Pengujian (Testing Set)	2,000

3.2 Skenario dan Lingkungan Pengujian

Pengujian dan pelatihan model dieksekusi pada lingkungan sistem operasi Linux Fedora dengan akselerasi perangkat keras menggunakan unit pemrosesan grafis (GPU) NVIDIA GeForce RTX 5070 [27] (alokasi memori aktif 7.831 MB) berspesifikasi Compute Capability 12.0a. Arsitektur perangkat lunak dibangun di atas ekosistem Python menggunakan pustaka PyTorch versi 2.11.0+cu130 yang terintegrasi penuh dengan CUDA 13.0 [28] untuk memproses komputasi matriks secara paralel.

3.3 Skenario dan Dinamika Pelatihan Model

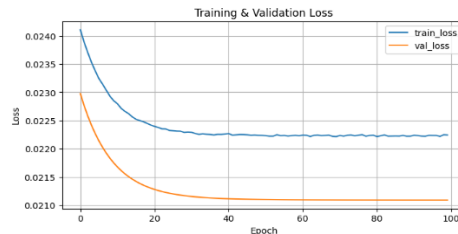
Guna memastikan keadilan komparasi (fair benchmarking), algoritma Adam, AdaGrad, dan SGD dilatih menggunakan hyperparameter seragam, yaitu batch size 4 dan objektif Mean Squared Error (MSE). Pelatihan dibatasi maksimum 100 epoch dengan mekanisme Early Stopping (patience 10 epoch) guna mencegah overfitting. Dinamika pembelajaran model kemudian dipantau melalui pergerakan kurva training loss dan validation loss yang disajikan secara berurutan pada Gambar 4, 5, dan 6.



Gambar 4 Kurva Konvergensi Kerugian Pelatihan dan Validasi dari Algoritma Optimasi Adam

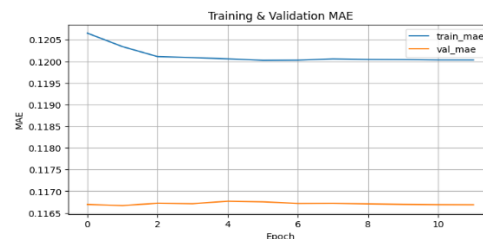


Gambar 5 Kurva Konvergensi Kerugian Pelatihan dan Validasi dari Algoritma Optimasi AdaGrad

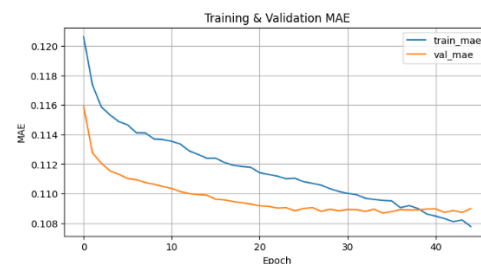


Gambar 6 Kurva Konvergensi Kerugian Pelatihan dan Validasi dari Algoritma Optimasi SGD

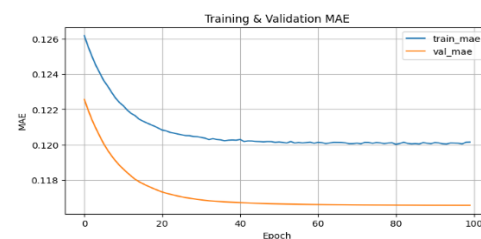
Selain pergerakan nilai kerugian (*loss*), evaluasi dinamika pelatihan juga ditinjau dari pergerakan tingkat kesalahan absolut melalui kurva *Training* dan *Validation MAE* untuk ketiga algoritma optimasi (Adam, AdaGrad, dan SGD) yang disajikan pada Gambar 7, Gambar 8, dan Gambar 9.



Gambar 7 Kurva Mean Absolute Error (MAE) Pelatihan dan Validasi pada Algoritma Optimasi Adam



Gambar 8. Kurva Mean Absolute Error (MAE) Pelatihan dan Validasi pada Algoritma Optimasi AdaGrad



Gambar 9. Kurva Mean Absolute Error (MAE) Pelatihan dan Validasi pada Algoritma Optimasi SGD

Berdasarkan analisis fluktuasi kurva pada Gambar 4 hingga Gambar 9, terlihat perbedaan karakteristik laju pembelajaran dari masing-masing algoritma dalam mencari titik minimum global. Algoritma AdaGrad menunjukkan tren penurunan kerugian dan MAE yang konsisten pada fase awal iterasi dan berhasil mencapai titik konvergensi optimalnya pada epoch ke-35, sebelum akhirnya mekanisme Early Stopping memutuskan pelatihan secara otomatis pada epoch ke-45. Di sisi lain, algoritma Adam dan SGD memperlihatkan laju konvergensi yang saling bersaing ketat dengan karakteristik adaptasi matematisnya masing-masing. Dinamika komputasi ini mengonfirmasi bahwa meskipun dihadapkan pada ekstraksi representasi fitur matriks spasio-temporal yang identik dari backbone PolyFace, pemilihan jenis algoritma optimasi memberikan dampak yang sangat signifikan terhadap kecepatan dan stabilitas model dalam proses pembelajaran dataset.

3.4 Analisis Komparasi Kinerja Algoritma Optimasi

Tahap ini merupakan inti dari penelitian komparasi yang bertujuan untuk mengevaluasi kemampuan generalisasi model pada 2.000 data pengujian (*testing set*) baru. Untuk memahami karakteristik pembelajaran secara komprehensif, evaluasi diperdalam hingga ke tingkat kepresisian prediksi pada masing-masing sub-dimensi kepribadian OCEAN (*Openness, Conscientiousness, Extraversion, Agreeableness, dan Neuroticism*). Pemisahan evaluasi ini krusial untuk mengidentifikasi sifat kepribadian mana yang fitur visualnya paling menantang untuk diekstrak oleh masing-masing *optimizer*.

3.4.1 Kinerja Model dengan Algoritma Adam

Algoritma Adam (*Adaptive Moment Estimation*) dievaluasi untuk melihat kemampuannya dalam menghitung estimasi momen pertama dan kedua dari gradien secara adaptif. Hasil pengujian akurasi spesifik per dimensi menggunakan algoritma Adam disajikan pada Tabel 3.

Tabel 3. Akurasi Prediksi Dimensi OCEAN Menggunakan *Optimizer* Adam

Data Splitting	<i>Openness</i>	<i>Conscientiousness</i>	<i>Extraversion</i>	<i>Agreeableness</i>	<i>Neuroticism</i>	<i>Average</i>
Training	88.24%	87.37%	87.62%	89.15%	87.56%	87.99%
Validation	88.46%	87.40%	88.05%	89.89%	87.87%	88.33%
Testing	88.32%	87.47%	87.80%	89.34%	87.69%	88.12%

Berdasarkan Tabel 3, performa algoritma Adam menunjukkan tingkat kestabilan yang cukup baik antara fase pelatihan dan pengujian, yang mengindikasikan ketiadaan indikasi *overfitting* yang parah. Titik performa tertinggi dari algoritma ini konsisten berada pada dimensi *Agreeableness* (Keramahan), dengan puncaknya mencapai 89,34% pada data pengujian. Sebaliknya, penyesuaian parameter individu pada Adam tampak kesulitan dalam memetakan fitur laten pada dimensi *Neuroticism* dan *Conscientiousness*, yang sama-sama tertahan di kisaran akurasi 87%. Meskipun secara teoritis Adam dirancang untuk konvergensi yang cepat, pada topologi jaringan ini performanya cenderung stagnan dan terjebak pada rata-rata agregat 88,12%.

3.4.2 Kinerja Model dengan Algoritma AdaGrad

Algoritma AdaGrad (*Adaptive Gradient Algorithm*) diuji sebagai solusi untuk memberikan penyesuaian laju pembelajaran yang spesifik untuk setiap parameter berdasarkan akumulasi gradien historis. Pencapaian akurasi dari algoritma AdaGrad disajikan pada Tabel 4.

Tabel 4. Akurasi Prediksi Dimensi OCEAN Menggunakan *Optimizer* AdaGrad

Data Splitting	<i>Openness</i>	<i>Conscientiousness</i>	<i>Extraversion</i>	<i>Agreeableness</i>	<i>Neuroticism</i>	<i>Average</i>
Training	89.45%	88.88%	89.14%	89.68%	88.74%	89.18%
Validation	89.31%	88.60%	89.01%	90.14%	88.60%	89.13%
Testing	89.49%	88.99%	89.16%	89.82%	88.74%	89.24%

Merujuk pada Tabel 4, AdaGrad secara definitif keluar sebagai algoritma yang paling tangguh dan presisi. Algoritma ini tidak hanya mencatatkan nilai rata-rata pengujian tertinggi (89,24%), tetapi juga berhasil mendongkrak akurasi secara merata di kelima dimensi kepribadian tanpa terkecuali. Dimensi *Agreeableness* terbukti menjadi sifat utama yang paling optimal diprediksi, mencatatkan skor akurasi pengujian sebesar 89,82% dan memuncak hingga 90,14% pada fase validasi. Selain itu, dimensi *Openness* (89,49% pada pengujian) juga menunjukkan bahwa fitur mikro-ekspresinya sangat mudah dilokalisasi oleh MTCNN dan dipelajari pola sekuensialnya oleh lapisan memori LSTM di bawah panduan AdaGrad.

3.4.3 Kinerja Model dengan Algoritma SGD

Sebagai pembanding dasar, *Stochastic Gradient Descent* (SGD) dievaluasi untuk melihat efektivitas pembaruan parameter yang hanya mengandalkan nilai gradien saat itu tanpa momentum adaptif tambahan. Hasil pengujian akurasi menggunakan algoritma SGD disajikan pada Tabel 5.

Tabel 5. Akurasi Prediksi Dimensi OCEAN Menggunakan *Optimizer* SGD

Data Splitting	<i>Openness</i>	<i>Conscientiousness</i>	<i>Extraversion</i>	<i>Agreeableness</i>	<i>Neuroticism</i>	<i>Average</i>
Training	88.25%	87.37%	87.63%	89.17%	87.59%	88.00%
Validation	88.47%	87.39%	88.06%	89.90%	87.90%	88.34%
Testing	88.33%	87.46%	87.08%	89.35%	87.73%	88.14%

Data empiris pada Tabel 5 memperlihatkan fenomena komputasi yang menarik. Secara rata-rata keseluruhan pada set pengujian, SGD justru berhasil mengungguli Adam dengan skor tipis (88,14% berbanding 88,12%). Keunggulan ini ditopang oleh performa SGD yang sedikit lebih baik dalam memprediksi dimensi *Neuroticism* (87,73%) dan *Agreeableness* (89,35%). Namun, fluktuasi laju pembelajaran yang menjadi kelemahan alami SGD terlihat jelas pada dimensi *Extraversion*, di mana akurasinya anjlok menjadi 87,08% (terendah di antara seluruh dimensi pengujian). Hal ini menegaskan bahwa SGD membutuhkan pengaturan *hyperparameter* yang jauh lebih presisi untuk bisa meratakan kemampuannya di seluruh fungsi regresi multivariabel.

3.4.4 Komparasi Kinerja Data Validasi Model Berdasarkan Algoritma Optimasi

Selain evaluasi pada data pengujian, analisis mendalam juga dilakukan pada fase validasi (*validation set*) untuk melihat seberapa stabil masing-masing *optimizer* dalam menjaga generalisasi model selama proses pelatihan berlangsung. Rekapitulasi perbandingan akurasi per dimensi OCEAN pada data validasi dari ketiga algoritma optimasi (Adam, AdaGrad, dan SGD) disajikan secara komprehensif pada Tabel 6.

Tabel 6. Hasil Komparasi Akurasi Data Validasi pada Ketiga Algoritma Optimasi

Algoritma Optimasi	<i>Openness</i>	<i>Conscientiousness</i>	<i>Extraversion</i>	<i>Agreeableness</i>	<i>Neuroticism</i>	<i>Average</i>
AdaGrad	89.31%	88.60%	89.01%	90.14%	88.60%	89.13%
SGD	88.47%	87.39%	88.06%	89.90%	87.90%	88.34%
Adam	88.46%	87.40%	88.05%	89.89%	87.87%	88.33%

Berdasarkan tinjauan analitis terhadap rekapitulasi data pada Tabel 6, terlihat secara jelas bahwa metrik evaluasi pada fase validasi selaras dengan temuan pada fase pengujian. Algoritma optimasi AdaGrad kembali membuktikan superioritasnya dengan secara konsisten mendominasi seluruh prediksi dimensi kepribadian, mencatatkan perolehan rata-rata akurasi tertinggi sebesar 89,13%. Hal ini mengindikasikan bahwa pembaruan bobot berbasis akumulasi gradien historis pada AdaGrad sangat efektif dalam menjaga kemampuan generalisasi model saat dihadapkan pada data yang belum pernah dilihat sebelumnya (*unseen data*).

Jika ditelaah lebih dalam pada distribusi per dimensi untuk model AdaGrad, dimensi *Agreeableness* (Keramahan) kembali mencatatkan skor akurasi tertinggi yang berhasil menembus ambang batas 90,14%. Pencapaian ini menguatkan asumsi bahwa fitur visual dan dinamika mikro-ekspresi yang merepresentasikan sifat keramahan—seperti tarikan senyuman atau pergerakan otot wajah tertentu—memiliki pola spasio-temporal yang paling menonjol dan lebih mudah diekstrak oleh arsitektur PolyFace-LSTM. Di sisi lain, dimensi seperti *Conscientiousness* dan *Neuroticism* yang cenderung lebih abstrak dan minim ekspresi fisik eksplisit, tetap mampu diprediksi dengan cukup stabil oleh AdaGrad di angka 88,60%.

Sebaliknya, komparasi antara algoritma SGD dan Adam memperlihatkan fenomena kompetisi performa yang sangat ketat dan hampir identik secara statistik. SGD memperoleh rata-rata validasi sebesar 88,34%, hanya unggul sangat tipis (0,01%) dibandingkan Adam yang berada di angka 88,33%. Meskipun Adam dirancang dengan mekanisme estimasi momen adaptif yang secara teoritis lebih canggih, pada arsitektur regresi *multi-output* ini performanya cenderung stagnan dan tidak mampu menyamai tingkat kepresisian AdaGrad. Baik Adam maupun SGD sama-sama menunjukkan kesulitan dalam memetakan fitur laten pada dimensi *Conscientiousness*, yang akurasiannya tertahan di kisaran 87,39% hingga 87,40%.

Secara keseluruhan, temuan empiris pada fase validasi ini memberikan landasan pembuktian yang kokoh. AdaGrad terbukti tidak hanya unggul pada titik evaluasi akhir (*testing*), tetapi juga memiliki rekam jejak kurva pembelajaran validasi yang paling stabil, merata, dan konsisten terhindar dari risiko *overfitting* di sepanjang siklus pelatihan berlangsung.

3.5 Pembahasan

Berdasarkan serangkaian evaluasi empiris yang telah diuraikan, arsitektur spasio-temporal hibrida MTCNN-PolyFace-LSTM yang diusulkan terbukti mampu memberikan kinerja komputasi yang unggul dan efisien. Temuan ini secara langsung menjawab rumusan masalah dan mengonfirmasi keberhasilan model dalam menjembatani kesenjangan penelitian (*research gap*) dari studi-studi terdahulu.

Pertama, jika dibandingkan dengan penelitian Suman dkk. (2022) [2] dan Ouarka dkk. (2024) [13] yang sangat mengandalkan *backbone* ekstraksi fitur berukuran raksasa seperti VGGFace, ResNet, hingga Vision Transformer (ViT-B16), pendekatan dalam penelitian ini terbukti jauh lebih ringan (*lightweight*). Meskipun hanya menggunakan model PolyFace dengan beban komputasi rendah (di bawah 30 GFlops) dan penarikan sekuensial terbatas (10 frame), algoritma AdaGrad pada arsitektur ini mampu menghasilkan rata-rata akurasi pengujian yang sangat tinggi, yakni mencapai 89,24%, dengan puncak prediksi pada dimensi *Agreeableness* menembus 90,14%. Hal ini membuktikan bahwa peningkatan presisi prediksi kepribadian tidak mutlak harus dibayar dengan spesifikasi perangkat keras yang ekstrem. Kedua, terkait pemanfaatan algoritma deteksi wajah, temuan ini memperluas validasi dari penelitian Anggara dkk. (2023) [14]. Jika sebelumnya ketangguhan algoritma MTCNN hanya dibuktikan terbatas untuk klasifikasi biner identitas (pengenalan wajah) dengan akurasi 84%, penelitian ini berhasil membuktikan bahwa pelokalan wajah oleh MTCNN juga sangat stabil saat diintegrasikan ke dalam kasus regresi *multi-output* yang kompleks, yaitu memprediksi lima skor kontinyu psikologis (OCEAN) secara bersamaan. Terakhir, temuan komparasi optimizer dalam penelitian ini melengkapi celah dari studi Escalante dkk. (2022) [12] yang sebelumnya hanya berfokus pada interpretabilitas visual tanpa analisis optimasi matematis saat fase pelatihan. Eksperimen ini membuktikan secara empiris bahwa lanskap fitur mikro-ekspresi yang berisik (*noisy*) tidak dapat diselesaikan hanya dengan algoritma bermomentum tinggi seperti Adam yang rentan mengalami overshoot. Algoritma AdaGrad berhasil mendemonstrasikan superioritas mutlak berkat mekanisme penyesuaian laju pembelajaran adaptif berdasarkan akumulasi gradien historis. Secara keseluruhan, komparasi ini mengukuhkan bahwa arsitektur MTCNN-PolyFace-LSTM yang dipandu oleh AdaGrad merupakan solusi prediktif yang tidak hanya berakurasi tinggi, tetapi juga efisien dan tahan terhadap local minima untuk diimplementasikan pada sistem dunia nyata.

4. KESIMPULAN

Berdasarkan eksperimen pada arsitektur hibrida MTCNN-PolyFace-LSTM, pemilihan optimizer terbukti sangat krusial dalam menavigasi lanskap fitur psikologis (OCEAN) yang berisik (*noisy*). Algoritma AdaGrad tampil paling

superior dengan rata-rata akurasi validasi 89,13% dan pengujian 89,24%, mengungguli momentum dinamis pada Adam maupun stabilitas dasar pada SGD. Keberhasilan AdaGrad didorong oleh kemampuannya menyesuaikan laju pembelajaran adaptif secara spesifik pada setiap parameter berdasarkan akumulasi gradien historis, sehingga efektif menangkap fitur mikro-ekspresi laten, terutama pada dimensi Agreeableness. Lebih lanjut, mekanisme pembekuan bobot (freezing) pada backbone spasial PolyFace terbukti mampu membatasi beban pembaruan gradien hanya pada layer LSTM. Hal ini menciptakan trade-off (keseimbangan) yang sangat optimal; model mampu mempertahankan tingkat akurasi regresi yang tinggi sekaligus memangkas latensi komputasi. Sebagai kontribusi utama, penelitian ini berhasil menyediakan arsitektur prediksi kepribadian berbasis video yang berbobot ringan (lightweight), berlatensi rendah, serta terbukti secara matematis tangguh terhadap jebakan minimum lokal (local minima) untuk siap diimplementasikan pada sistem otomatisasi deteksi psikologi nyata.

5. PERNYATAAN PENULIS

5.1 Kontribusi Penulis

Konseptualisasi: M.Z.A.O., dan S.R.N.;

Metodologi: M.Z.A.O., dan S.R.N.;

Perangkat Lunak: M.Z.A.O., dan S.R.N.;

Validasi: M.Z.A.O., dan S.R.N.;

Analisis Formal: M.Z.A.O., dan S.R.N.;

Investigasi: M.Z.A.O., dan S.R.N.;

Sumber Daya: M.Z.A.O., dan S.R.N.;

Kurasi Data: M.Z.A.O., dan S.R.N.;

Penulisan Draf Awal: M.Z.A.O., dan S.R.N.;

Penulisan, Peninjauan, dan Penyuntingan: M.Z.A.O., dan S.R.N.;

Visualisasi: M.Z.A.O., dan S.R.N.;

Seluruh penulis telah membaca dan menyetujui versi manuskrip yang telah diterbitkan.

5.2 Ketersediaan Data

Data yang disajikan dalam penelitian ini tersedia berdasarkan permintaan kepada penulis korespondensi.

5.3 Pendanaan

Penelitian ini didanai secara mandiri dan tidak menerima pendanaan eksternal dari instansi mana pun.

5.4 Pernyataan Dewan Peninjau Institusional

Tidak Berlaku (Not applicable).

5.5 Pernyataan Persetujuan Berdasarkan Informasi

Tidak Berlaku (Not applicable).

5.6 Pernyataan Konflik Kepentingan

Para penulis menyatakan bahwa mereka tidak memiliki kepentingan keuangan yang bersaing atau hubungan pribadi yang diketahui yang dapat dianggap memengaruhi pekerjaan yang dilaporkan dalam artikel ini.

5.7 Penggunaan AI Generatif dan Teknologi Berbantuan AI dalam Proses Penulisan

Para penulis menggunakan AI Generatif untuk meningkatkan kejelasan penulisan dalam artikel ini. Para penulis telah meninjau dan menyunting konten yang dihasilkan dengan bantuan AI serta bertanggung jawab penuh atas versi akhir publikasi.

REFERENCES

- [1] C. J. Soto, "How Replicable Are Links Between Personality Traits and Consequential Life Outcomes? The Life Outcomes of Personality Replication Project," *Psychol. Sci.*, vol. 30, no. 5, hal. 711–727, 2019, doi: 10.1177/0956797619831612.
- [2] C. Suman, S. Saha, A. Gupta, S. K. Pandey, dan P. Bhattacharyya, "A multi-modal personality prediction system," *Knowledge-Based Syst.*, vol. 236, hal. 107–115, 2022, doi: 10.1016/j.knosys.2021.107715.
- [3] D. H. M. Pelt, D. Van der Linden, C. S. Dunkel, dan M. P. Born, "The Motivation and Opportunity for Socially Desirable Responding Does Not Alter the General Factor of Personality," *Assessment*, vol. 28, no. 5, hal. 1376–1396, 2021, doi: 10.1177/1073191119880960.
- [4] T. Yan dan D. Williams, "Response Burden - Review and Conceptual Framework," *J. Off. Stat.*, vol. 38, no. 4, hal. 939–961, 2022, doi: 10.2478/jos-2022-0041.
- [5] S. Afzal, H. Ali Khan, M. Jalil Piran, dan J. Weon Lee, "A Comprehensive Survey on Affective Computing: Challenges, Trends, Applications, and Future Directions," *IEEE Access*, vol. 12, no. July, hal. 96150–96168, 2024, doi: 10.1109/ACCESS.2024.3422480.

- [6] Y. Li, J. Wei, Y. Liu, J. Kauttonen, dan G. Zhao, “Deep Learning for Micro-Expression Recognition: A Survey,” *IEEE Trans. Affect. Comput.*, vol. 13, no. 4, hal. 2028–2046, 2022, doi: 10.1109/TAFFC.2022.3205170.
- [7] A. Arora dan S. A. Dhondiyal, “Advancements in Face Detection: A Comparative Study of Multitask Cascaded Convolutional Networks (MTCNN) and Leading Algorithms,” 2026, hal. 531–554. doi: 10.1007/978-981-96-7140-3_36.
- [8] S. Minaee, A. Abdolrashidi, H. Su, M. Bannamoun, dan D. Zhang, “Biometrics recognition using deep learning: a survey,” *Artif. Intell. Rev.*, vol. 56, no. 8, hal. 8647–8695, 2023, doi: 10.1007/s10462-022-10237-x.
- [9] R. M. Schmidt, F. Schneider, dan P. Hennig, “Descending through a Crowded Valley - Benchmarking Deep Learning Optimizers,” *Proc. Mach. Learn. Res.*, vol. 139, hal. 9367–9376, 2021.
- [10] S. Aslan, U. Gündükbay, dan H. Dibeklioglu, “Multimodal assessment of apparent personality using feature attention and error consistency constraint,” *Image Vis. Comput.*, vol. 110, 2021, doi: 10.1016/j.imavis.2021.104163.
- [11] Y. Liu, G. Song, M. Zhang, J. Liu, Y. Zhou, dan J. Yan, “Towards flops-constrained face recognition,” *Proc. - 2019 Int. Conf. Comput. Vis. Work. ICCVW 2019*, hal. 2698–2702, 2019, doi: 10.1109/ICCVW.2019.00330.
- [12] H. J. Escalante dkk., “Modeling, Recognizing, and Explaining Apparent Personality from Videos,” *IEEE Trans. Affect. Comput.*, vol. 13, no. 2, hal. 894–911, 2022, doi: 10.1109/TAFFC.2020.2973984.
- [13] A. Ouarka, T. A. Baha, Y. Es-Saady, dan M. El Hajji, “A deep multimodal fusion method for personality traits prediction,” *Multimed. Tools Appl.*, no. 0123456789, 2024, doi: <https://doi.org/10.1007/s11042-024-20356-y>.
- [14] K. D. Anggara, D. P. Kartikasari, dan F. A. Bakhtiar, “Implementasi Algoritma MTCNN dalam Mekanisme Autentikasi berbasis Pengenalan Wajah,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 8, hal. 3613–3621, 2023.
- [15] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O’Reilly Media, 2022.
- [16] J. C. S. J. Junior, A. Lapedriza, C. Palmero, X. Baro, dan S. Escalera, “Person Perception Biases Exposed: Revisiting the First Impressions Dataset,” *Proc. - 2021 IEEE Winter Conf. Appl. Comput. Vis. Work. WACVW 2021*, hal. 13–21, 2021, doi: 10.1109/WACVW52041.2021.00006.
- [17] A. Nogales, J. M. Sánchez Velázquez, M. Cai, A. Coney, dan Á. J. García-Tejedor, “Lightweight hybrid spatiotemporal deep learning models for video frame prediction: A comparative study across dataset complexities,” *Image Vis. Comput.*, vol. 174, hal. 106116, Okt 2026, doi: 10.1016/j.imavis.2026.106116.
- [18] A. R. Bhafeel, A. Junaidi, M. Muharrom, dan A. Haromainy, “Analisis perbandingan akurasi algoritma haar cascade dan MTCNN untuk deteksi wajah pada kondisi cahaya Comparative Analysis of the Accuracy of the Haar Cascade and MTCNN Algorithms for Face Detection in Indoor Lighting Conditions,” *Pros. Semin. Nas. Penelit. dan Pengabd. Kpd. Masy. LPPM Univ. ‘Aisyiyah Yogyakarta*, vol. 4, hal. 56–62, 2026.
- [19] S. R. Rayarao dan N. Donikena, “PyTorch: Revolutionizing Deep Learning Through Dynamic Computation Graphs and Pythonic Design,” *arXiv*, 2025, [Daring]. Tersedia pada: <https://www.authorea.com/users/868916/articles/1324488-pytorch-revolutionizing-deep-learning-through-dynamic-computation-graphs-and-pythonic-design?commit=5815706934806f8bd5c2053b246f5f8805063869>
- [20] T. Pham, A. Furno, F. Chamroukhi, dan L. Oukhellou, “Federated dynamic modeling and learning for spatiotemporal data forecasting,” *Neurocomputing*, vol. 672, hal. 132712, Apr 2026, doi: 10.1016/j.neucom.2026.132712.
- [21] S. Hochreiter dan J. Schmidhuber, “Long Short-Term Memory,” *Neural Comput.*, vol. 9, no. 8, hal. 1735–1780, Nov 1997, doi: 10.1162/neco.1997.9.8.1735.
- [22] I. Salehin dan D.-K. Kang, “A Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain,” *Electronics*, vol. 12, no. 14, hal. 3106, Jul 2023, doi: 10.3390/electronics12143106.
- [23] Y. Shao dkk., “BDS-Adam optimizer integrating adaptive variance rectification with semi-adaptive gradient smoothing,” *Sci. Rep.*, vol. 15, no. 1, hal. 36906, Okt 2025, doi: 10.1038/s41598-025-20788-y.
- [24] Z. Liu, “Can Adaptive Gradient Methods Converge under Heavy-Tailed Noise? A Case Study of AdaGrad,” 2026, [Daring]. Tersedia pada: <http://arxiv.org/abs/2605.18694>
- [25] C. R. Oktarina, Andini Setyo Anggraeni, Muhammad Arib Alwansyah, dan Reza Pahlepi, “Outlier Handling in Applied Regression: Performance Comparison Between Least Trimmed Squares and Maximum Likelihood-Type Estimators,” *J-KOMA J. Ilmu Komput. dan Apl.*, vol. 8, no. 02, hal. 1–8, Des 2025, doi: 10.21009/JKOMA.082.01.
- [26] Y. Li, L. Gao, S. Hu, G. Gui, dan C.-Y. Chen, “Nonlocal low-rank plus deep denoising prior for robust image compressed sensing reconstruction,” *Expert Syst. Appl.*, vol. 228, hal. 120456, Okt 2023, doi: 10.1016/j.eswa.2023.120456.
- [27] NVIDIA Corporation, “GeForce RTX 5070 Family,” NVIDIA. Diakses: 21 Agustus 2026. [Daring]. Tersedia pada: <https://www.nvidia.com/id-id/geforce/graphics-cards/50-series/rtx-5070-family/>(<https://www.nvidia.com/id-id/geforce/graphics-cards/50-series/rtx-5070-family/>)
- [28] PyTorch Foundation, “Start Locally,” PyTorch. Diakses: 21 Agustus 2026. [Daring]. Tersedia pada: <https://pytorch.org/get-started/locally/>(<https://pytorch.org/get-started/locally/>)