

## PERBANDINGAN OPTIMIZER ADAM, RMSPROP, DAN SGD DALAM PREDIKSI KEPRIBADIAN BERBASIS OPENFACE

Riska Octavia Arianto, Salamun Rohman Nudin

Manajemen Informatika, Universitas Negeri Surabaya  
Jalan Ketintang, Ketintang, Kec. Gayungan Surabaya, Indonesia  
riska.21059@mhs.unesa.ac.id

### ABSTRAK

Identifikasi kepribadian tradisional yang mengandalkan kuesioner subjektif memiliki keterbatasan dalam reliabilitas dan efisiensi. Penelitian ini mengembangkan pendekatan objektif berbasis video wajah dengan memanfaatkan model *deep learning* untuk mengklasifikasikan lima dimensi kepribadian OCEAN (*Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism*). Fokus utama penelitian adalah membandingkan efektivitas tiga *optimizer* utama (Adam, RMSProp, dan SGD) pada arsitektur OpenFace untuk menentukan solusi paling optimal. Metode yang digunakan meliputi: deteksi wajah presisi dengan MTCNN, ekstraksi fitur menggunakan OpenFace, dan analisis temporal melalui LSTM pada dataset 10.000 sampel video. Hasil eksperimen menunjukkan bahwa ketiga *optimizer* mencapai akurasi setara ( $\pm 88.15\%$ ), namun RMSProp unggul dalam efisiensi komputasi dengan waktu pelatihan tercepat (3m 23.3s) dibandingkan Adam (5m 49.8s) dan SGD (6m 23.9s). Temuan krusial mengungkap bahwa pendekatan tanpa *optimizer* justru lebih cepat (3m 41.7s) namun kurang stabil, menunjukkan kompleksitas unik dalam optimasi model berbasis video. Penelitian ini tidak hanya membuktikan validitas pendekatan berbasis video sebagai alternatif kuesioner konvensional, tetapi juga menetapkan RMSProp sebagai pilihan optimal untuk sistem klasifikasi kepribadian *real-time*, dengan rekomendasi khusus untuk peningkatan kualitas dataset guna mengatasi tantangan *frame* tidak ideal.

**Kata kunci :** *big five personality, OpenFace, LSTM, Adam, RMSProp, SGD*

### 1. PENDAHULUAN

Kepribadian merupakan aspek mendasar yang memengaruhi cara individu berpikir, merasakan, dan berinteraksi dengan lingkungan sekitarnya. Memahami kepribadian seseorang sangat krusial untuk mengidentifikasi potensi yang dimilikinya. Menurut teori *Big Five Personality*, karakteristik kepribadian manusia yang beragam dapat dikelompokkan ke dalam lima dimensi utama, yaitu *Openness* (keterbukaan), *Conscientiousness* (kedisiplinan), *Extraversion* (ekstraversi), *Agreeableness* (kesepakatan), dan *Neuroticism* (neotirisme), yang sering disingkat sebagai OCEAN [1]. Temuan John Romel dkk. Pada tahun 2020 juga mengungkap bahwa keterbukaan, kedisiplinan, dan kesepakatan berperan penting dalam kinerja akademik karena pengaruhnya bersifat bertahap dan tidak selalu bergantung pada nilai akhir [2]. Dengan demikian, memprediksi kepribadian dapat membantu seseorang mengenali kekuatan dan kelemahan dirinya, sehingga dapat digunakan sebagai dasar pengembangan diri.

Metode tradisional dalam memprediksi kepribadian, seperti kuesioner, dinilai memakan waktu dan cenderung subjektif [3]. Seiring kemajuan teknologi, pengembangan prediksi kepribadian berbasis video mulai dilakukan untuk mempercepat proses, terutama dalam rekrutmen karyawan, dengan memanfaatkan *Deep Learning*. *Deep learning* merupakan teknik pembelajaran mesin yang bekerja dengan membangun representasi data secara bertingkat melalui beberapa lapisan pemrosesan [4]. Salah satu pendekatan yang banyak digunakan adalah *transfer learning*, di mana model yang telah

dilatih sebelumnya (*pre-trained model*) diaplikasikan pada dataset baru untuk meningkatkan akurasi dan efisiensi prediksi [5].

*Automatic Personality Recognition* (APR) adalah bidang penelitian yang bertujuan mengidentifikasi kepribadian seseorang secara otomatis melalui sinyal perilaku seperti ekspresi wajah, suara, atau bahasa tubuh. Song dkk. telah membuktikan bahwa pendekatan *self-supervised learning* (SSL) mampu mempelajari representasi dinamika wajah yang spesifik individu secara otomatis dari data video tanpa bergantung pada label kepribadian, kemudian meningkatkan akurasi APR melalui *transfer learning* [6]. Berdasarkan temuan tersebut, penelitian ini mengadopsi ekstraksi fitur wajah sebagai dasar untuk memprediksi atribut kepribadian.

Ekstraksi fitur wajah dilakukan dengan menggunakan OpenFace, sebuah alat analisis perilaku wajah yang mendeteksi *landmark* wajah dan mengintegrasikannya dengan algoritma yang ada. OpenFace memanfaatkan *Action Units* (AU) untuk menginterpretasikan ekspresi emosi berdasarkan gerakan otot wajah. Model ini menerima input gambar berukuran 96x96 piksel dengan 3 saluran warna dan menghasilkan vector fitur 128 dimensi [7]. Penelitian sebelumnya menunjukkan bahwa OpenFace mampu mencapai akurasi 93,3% [8]. Untuk menganalisis data temporal dari ekspresi wajah dalam video, fitur yang diekstraksi kemudian diproses menggunakan *Long Short-Term Memory* (LSTM), yang merupakan jenis *recurrent neural network* (RNN) yang efektif menangani data berurutan karena kemampuannya menyimpan informasi jangka panjang [9]. Dengan

LSTM, sistem dapat mengidentifikasi perubahan ekspresi wajah secara dinamis dari satu frame ke frame lainnya, sehingga meningkatkan akurasi prediksi kepribadian.

Penggunaan video dapat menjadi sumber data yang efektif untuk memprediksi kepribadian berdasarkan *Big Five Personality*. Penelitian Suman dkk. membuktikan bahwa klip video pendek dapat digunakan untuk menganalisis karakteristik kepribadian seseorang [10]. Dalam penelitian ini, MTCNN (*Multi-task Cascaded Convolutional Neural Network*) digunakan untuk optimasi deteksi wajah. MTCNN unggul dalam mendeteksi wajah secara *real-time*, tahan terhadap variasi pencahayaan, sudut pengambilan gambar, dan ekspresi alami, serta tidak memerlukan memori besar [11].

Pemilihan *optimizer* dalam pengembangan model klasifikasi menjadi faktor penentu efisiensi dan akurasi. Beberapa *optimizer* populer seperti Adam (*Adaptive Moment Estimation*), RMSProp (*Root Mean Square Propagation*), dan SGD (*Stochastic Gradient Descent*) memiliki karakteristik unik dalam mengoptimalkan model. Adam dikenal dengan adaptivitas *learning rate*-nya [12], RMSProp efektif menangani *vanishing gradient*, sedangkan SGD meskipun sederhana memerlukan tuning yang cermat [13].

Setelah ekstraksi fitur menggunakan OpenFace dan dioptimasi menggunakan *optimizer*, data diproses lebih lanjut dengan metode regresi untuk memprediksi skor kepribadian. Evaluasi performa model dilakukan menggunakan *Mean Absolute Error* (MAE), metrik yang mengukur rata-rata selisih absolut antara prediksi dan nilai aktual. MAE dipilih karena memberikan gambaran kesalahan prediksi yang mudah dipahami dan tidak terpengaruh oleh *outlier* [14]. Penelitian ini bertujuan untuk mengisi celah tersebut dengan membandingkan efektivitas Adam, RMSProp, dan SGD dalam mengklasifikasikan dimensi kepribadian OCEAN menggunakan fitur wajah yang diekstraksi oleh OpenFace, sekaligus mengevaluasi potensi pendekatan ini sebagai alternatif yang lebih objektif dan efisien dibandingkan metode kuesioner konvensional.

## 2. TINJAUAN PUSTAKA

Tinjauan Pustaka merupakan bagian penting dalam penelitian yang berisi analisis kritis terhadap berbagai sumber referensi terkait topik penelitian. Tujuannya adalah untuk mengetahui perkembangan ilmu, menemukan gap (celah penelitian), dan membangun landasan teori yang mendukung penelitian yang sedang dilakukan. Pada penelitian ini, tinjauan Pustaka mencakup *Big Five Personality*, Deep Learning, MTCNN, OpenFace, LSTM, 3 Optimizer (Adam, RMSProp, SGD), Regresi, dan Akurasi.

### 2.1. Big Five Personality

*Big Five Personality* (OCEAN) merupakan model psikologi yang secara komprehensif

mengklasifikasikan kepribadian manusia ke dalam lima dimensi utama yang saling independen. Kelima dimensi tersebut meliputi *Openness* yang mencerminkan tingkat keterbukaan terhadap pengalaman baru, kreativitas, dan imajinasi; *Conscientiousness* yang menunjukkan tingkat kedisiplinan, tanggung jawab, dan keteraturan dalam bertindak; *Extraversion* yang menggambarkan kecenderungan untuk bersosialisasi dan berenergi tinggi; *Agreeableness* yang mengukur keramahan, empati, dan kemampuan kerja sama; serta *Neuroticism* yang menilai kecenderungan mengalami emosi negatif seperti kecemasan dan kerentanan terhadap stress [15]. Model ini dikembangkan berdasarkan penelitian empiris yang ekstensif dan diakui sebagai salah satu pendekatan paling valid dan reliabel dalam psikologi kepribadian, karena mampu menjelaskan variasi perilaku manusia secara luas dan stabil dalam berbagai konteks kehidupan lintas budaya [16]. Kelima dimensi ini tidak hanya memberikan kerangka teoritis yang kuat, tetapi juga memiliki aplikasi praktis yang luas dalam berbagai bidang seperti psikologi klinis, sumber daya manusia, dan pengembangan diri.

### 2.2. Deep Learning

Deep learning merupakan cabang pembelajaran mesin yang menggunakan arsitektur jaringan saraf tiruan (*neural networks*) untuk memproses dan mempelajari data secara hierarkis. Berbeda dengan metode tradisional yang membutuhkan ekstraksi fitur manual, deep learning mampu secara otomatis mengekstrak dan mempelajari fitur-fitur kompleks dari data mentah melalui lapisan-lapisan jaringan yang dalam [16]. Teknologi ini mendukung berbagai jenis arsitektur jaringan seperti *Convolutional Neural Networks* (CNN) untuk pemrosesan gambar [17] dan *Recurrent Neural Networks* (RNN) untuk data berurutan seperti teks atau suara. Perkembangannya yang pesat didorong oleh ketersediaan data skala besar, peningkatan daya komputasi GPU, serta teknik optimasi seperti *backpropagation*, memungkinkannya mencapai kinerja setara manusia dalam berbagai tugas seperti pengenalan wajah, pemrosesan bahasa alami, dan analisis perilaku. Kemampuannya dalam menangani data kompleks menjadikan *deep learning* sebagai solusi efektif untuk berbagai aplikasi industri dan penelitian, termasuk dalam analisis kepribadian berbasis fitur wajah [18].

### 2.3. MTCNN

MTCNN (*Multi-task Cascaded Convolutional Neural Network*) merupakan arsitektur deep learning khusus yang dirancang untuk deteksi wajah secara komprehensif melalui pendekatan *multi-task* terintegrasi. Model ini mengombinasikan tiga fungsi utama sekaligus: deteksi keberadaan wajah, penentuan *bounding box*, dan identifikasi *landmark* wajah (seperti posisi mata, hidung, dan mulut) dalam satu sistem terpadu. MTCNN menggunakan struktur bertingkat yang terdiri dari tiga jaringan kaskade - P-

Net (*Proposal Network*), R-Net (*Refinement Network*), dan O-Net (*Output Network*) yang bekerja secara berurutan untuk menyaring dan memurnikan hasil deteksi secara progresif [18]. Arsitektur ini memungkinkan deteksi wajah yang robust terhadap variasi sudut pandang, ekspresi wajah, dan kondisi pencahayaan, sekaligus mempertahankan efisiensi komputasi untuk aplikasi *real-time*. Keunggulan utamanya terletak pada kemampuan multiskala dan presisi *landmark* yang membuatnya banyak diaplikasikan dalam sistem keamanan biometrik, *augmented reality*, dan analisis wajah canggih, termasuk dalam penelitian analisis kepribadian berbasis fitur wajah [11].

## 2.4. OpenFace

OpenFace adalah *toolkit open-source* yang dikembangkan untuk analisis perilaku wajah, seperti deteksi wajah, pengenalan wajah, dan ekstraksi fitur wajah (*landmark facial*). Dirancang agar mudah digunakan dalam berbagai aplikasi, OpenFace menggunakan algoritma *deep learning* dan model yang dilatih secara mendalam untuk mengenali ekspresi wajah, emosi, dan fitur seperti posisi alis, mata, hidung, dan mulut. OpenFace banyak digunakan dalam penelitian, pendidikan, dan aplikasi industri untuk memahami interaksi manusia dengan lingkungan dan perangkat [7]. Keunggulan OpenFace adalah kemampuannya untuk melakukan analisis perilaku wajah secara *real-time* dan dengan akurasi yang cukup tinggi. Openface masih menggunakan *loss function* yang sama yaitu *Triplet Loss* dalam pembuatan embedding. Model Openface yang digunakan mempunyai input gambar 96x96 piksel dengan 3 *channel*, yang menghasilkan vektor 128 dimensi [5].

## 2.5. LSTM

*Long Short-Term Memory* (LSTM) merupakan pengembangan dari *Recurrent Neural Network* (RNN) yang secara khusus dirancang untuk mengatasi keterbatasan RNN konvensional dalam memproses data sekuensial jangka panjang [9]. Arsitektur unik LSTM diperkenalkan pertama kali oleh Hochreiter dan Schmidhuber pada tahun 1997, dengan ciri khas adanya mekanisme gerbang (*gate*) yang terdiri atas *forget gate*, *input gate*, dan *output gate*. *Forget gate* berfungsi untuk memfilter informasi yang tidak relevan dari memori sebelumnya menggunakan fungsi sigmoid, sementara *input gate* menentukan informasi baru yang akan disimpan melalui kombinasi fungsi sigmoid dan tanh. *Output gate* kemudian mengatur informasi apa yang akan diteruskan ke langkah waktu berikutnya. Mekanisme pintar ini memungkinkan LSTM untuk secara dinamis mempertahankan atau membuang informasi, sehingga mampu menangkap dependensi jangka panjang dalam data sekuensial dengan lebih efektif dibandingkan RNN biasa. Dalam konteks penelitian analisis video untuk klasifikasi kepribadian, kemampuan LSTM dalam memproses

informasi temporal dari rangkaian frame video menjadikannya komponen krusial untuk memahami dinamika ekspresi wajah secara berkesinambungan. Keunggulan inilah yang membuat LSTM banyak diaplikasikan dalam berbagai tugas pemrosesan data sekuensial, mulai dari pengenalan ucapan hingga prediksi deret waktu. Prosesnya dirumuskan sebagai berikut:

$$\begin{aligned} i^t &= \sigma(W_{xi}\tilde{X}_t + W_{hi}h^{t-1} + W_{ci}c^{t-1} + b_i) \\ f^t &= \sigma(W_{xf}\tilde{X}_t + W_{hf}h^{t-1} + W_{cf}c^{t-1} + b_f) \\ c^t &= f^t c^{t-1} + i^t \tanh(W_{xc}\tilde{X}_t + W_{hc}h^{t-1} + b_c) \\ o^t &= \sigma(W_{xo}\tilde{X}_t + W_{ho}h^{t-1} + W_{co}c^{t-1} + b_o) \\ h^t &= o^t \tanh(c^t) \end{aligned} \quad (1)$$

Pada arsitektur *Long Short-Term Memory* (LSTM), fungsi  $\sigma(\cdot)$  mengacu pada fungsi aktivasi sigmoid, yang berperan dalam menentukan seberapa banyak informasi yang dapat diteruskan dalam jaringan. Notasi  $W$  dengan dua subskrip (misalnya  $W_{xi}, W_{hi}$ ) menyatakan bobot koneksi antar komponen dalam sel LSTM, seperti antara input dan memori atau antar sel memori itu sendiri.

## 2.6. ADAM

Adam (*Adaptive Moment Estimation*) adalah algoritma optimisasi yang menggabungkan keunggulan dua metode, yaitu AdaGrad dan RMSProp. Adam menghitung rata-rata eksponensial dari gradien dan kuadrat gradien untuk memperbarui bobot, memungkinkan penyesuaian langkah yang lebih cerdas berdasarkan sifat data. Hal ini menjadikannya efisien dan efektif untuk banyak jenis model, terutama dalam *deep learning*. Kelebihan Adam termasuk kemampuan untuk beradaptasi dengan laju pembelajaran yang berbeda untuk setiap parameter dan mempercepat konvergensi selama pelatihan [12]. Berikut adalah persamaan untuk menghitung algoritma Adam:

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t \\ v_t &= \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \\ \hat{m}_t &= \left( \frac{m_t}{1 - \beta_1^t} \right) \quad \hat{v}_t = \left( \frac{v_t}{1 - \beta_2^t} \right) \\ \theta_{t+1} &= \theta_t - \frac{\eta}{\sqrt{\hat{v}_t - \epsilon}} \hat{m}_t \end{aligned} \quad (2)$$

Dimana  $\beta_1$  adalah laju penyusutan eksponensial. Nilai default  $\beta_1$  dan  $\beta_2$  adalah 0,9 dan 0,999.  $\hat{m}_t$  dan  $\hat{v}_t$  adalah bias koreksi untuk  $m_t$  dan  $v_t$ .

## 2.7. RMSProp

*Root Mean Square Prop* (RMSProp) adalah algoritma yang membantu model belajar lebih cepat dan stabil. Cara kerjanya adalah dengan mengatur seberapa besar perubahan yang dilakukan pada setiap langkah pembelajaran (*learning rate*), dan itu dilakukan secara otomatis untuk setiap parameter model. RMSProp menyimpan catatan tentang

seberapa besar gradien (perubahan) yang terjadi sebelumnya, lalu menghitung rata-ratanya. Dengan begitu, jika gradien terlalu besar, RMSProp akan mengecilkan langkah pembelajaran agar model tidak "loncat-loncat" dan jadi tidak stabil. Sebaliknya, jika gradien kecil, ia akan menyesuaikan agar pembelajaran tetap berjalan. Intinya, RMSProp membantu model belajar dengan langkah yang pas, sehingga lebih cepat mencapai hasil yang baik, terutama saat datanya berubah-ubah [13]. Persamaan RMSProp sebagai berikut:

$$\begin{aligned} v_{t+1} &= \gamma v_t + (1 - \gamma)(\nabla f(W_t))^2 \\ W_{t+1} &= W_t - \frac{\alpha}{v_{t+1}} \nabla f(W_t) \end{aligned} \quad (3)$$

## 2.8. SGD

*Stochastic Gradient Descent* (SGD) adalah algoritma optimisasi yang umum digunakan dalam *machine learning* dan *deep learning*. Dalam metode ini, pembaruan bobot dilakukan secara acak dengan menggunakan hanya sebagian kecil dari data (*mini-batch*) pada setiap iterasi, yang membantu mempercepat proses pelatihan. SGD efektif dalam mengurangi memori yang dibutuhkan dan dapat mempercepat konvergensi, tetapi juga memiliki kelemahan dalam stabilitas dan bisa terjebak dalam lokal minima. Untuk mengatasi masalah ini, sering digunakan teknik seperti momentum, yang memperhitungkan pembaruan sebelumnya untuk menghaluskan pembaruan bobot [13]. Berikut adalah persamaan untuk menghitung SGD:

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla_{\theta} J(\theta_t; x^{(i)}; y^{(i)}) \quad (4)$$

Dimana  $\eta$  adalah laju pembelajaran,  $\nabla_{\theta} J(\theta_t)$  adalah gradien fungsi kerugian terhadap  $J(\theta_t)$ , selain itu  $x^{(i)}$  dan  $y^{(i)}$  merepresentasikan data pelatihan dalam pasangan *input-output*. Jika kurva *loss function* memiliki titik pelana dimana satu dimensi miring ke atas, dan dimensi lainnya miring ke bawah, algoritma ini tidak akan bekerja dengan baik.

## 2.9. Regresi

Regresi adalah metode dalam *machine learning* yang digunakan untuk memprediksi nilai kontinu, seperti harga, suhu, atau skor kepribadian. Dalam proses evaluasi model regresi, salah satu metrik yang umum digunakan adalah MAE (*Mean Absolute Error*). MAE mengukur rata-rata dari selisih absolut antara nilai prediksi model dan nilai sebenarnya (*ground truth*). Semakin kecil nilai MAE, berarti prediksi model semakin akurat. MAE memiliki keunggulan karena sederhana dan tidak terlalu sensitif terhadap outlier, berbeda dengan metrik lain seperti MSE (*Mean Squared Error*) yang menghukum kesalahan besar lebih keras. Oleh karena itu, MAE sering dipilih ketika

kesalahan besar dan kecil ingin dianggap sama penting dalam evaluasi model regresi [14]. Berikut merupakan rumus dari MAE:

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| \quad (5)$$

## 2.10. Akurasi

Akurasi adalah salah satu metrik evaluasi dalam *machine learning* dan statistik yang digunakan untuk mengukur seberapa tepat model melakukan prediksi dibandingkan dengan data sebenarnya. Secara sederhana, akurasi menunjukkan seberapa banyak prediksi model yang benar dari total seluruh prediksi. Dari perhitungan nilai akurasi akan berkisar antara 0% (prediksi salah total) hingga 100% (prediksi sempurna). Akurasi sendiri dapat dihitung dengan rumus berikut. Semakin kecil selisih  $|\hat{y}_i - y_i|$ , semakin tinggi nilai akurasinya [19].

$$Accuracy = 1 - \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (6)$$

## 3. METODE PENELITIAN



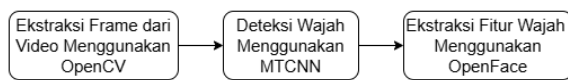
Gambar 1. Alur Penelitian

Metode penelitian adalah kerangka sistematis yang digunakan untuk merancang, melaksanakan, dan menganalisis suatu penelitian. Ini mencakup langkah-langkah, teknik, dan alat yang dipilih untuk menjawab pertanyaan penelitian atau menguji hipotesis secara valid dan reliabel. Pada penelitian ini memiliki langkah atau alur penelitian seperti pada gambar 1, yaitu pengumpulan data, persiapan data, pengembangan model, dan evaluasi.

### 3.1. Pengumpulan Data

Dataset merupakan kumpulan data terstruktur yang disimpan dalam memori (*in-memory*) dengan karakteristik dan fungsi serupa database konvensional. Objek ini merepresentasikan data beserta relasinya dalam format yang mudah diolah komputer untuk menghasilkan informasi [13]. Pengumpulan data yang dilakukan pada penelitian ini menggunakan data sekunder yang diambil dari dataset ChaLearn Looking at People (LAP) yang digunakan dalam kompetisi di CVPR 2017 terdiri dari sekitar 10.000 video. Perbandingan data train, validation, dan testing yang digunakan sebesar 3:1:1. Video diambil dari berbagai konteks sosial dan telah diberi label oleh *Amazon Mechanical Turk* (AMT) dengan variabel ciri-ciri kepribadian *Big Five Personality* (*Openness, Conscientiousness, Extraversion, Agreeableness, dan Neuroticism*). Label lainnya terdiri dari ragam etnis (Asia, Kaukasia, dan Afrika-Amerika), jenis kelamin (pria dan wanita), dan rentang usia (0-6, 7-13, 14-18, 19-24, 25-32, 33-45, 46-60, 61+) [20].

### 3.2. Persiapan Data

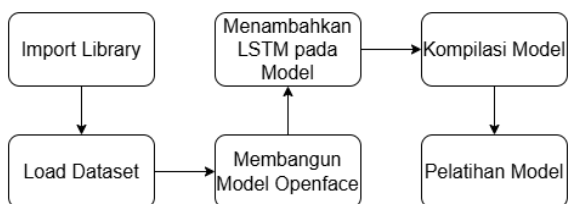


Gambar 2. Alur persiapan data

Proses persiapan data merupakan tahapan kritis setelah pengumpulan data untuk memastikan data siap diproses oleh model. Seperti pada gambar 2, terdapat 3 tahap yaitu sebagai berikut:

- Tahap pertama, mengekstraksi frame dari video menggunakan OpenCV. Proses diawali dengan ekstraksi frame dari video menggunakan OpenCV, di mana setiap frame video berukuran  $1280 \times 720$  piksel atau  $406 \times 720$  piksel diambil sebagai bahan dasar. OpenCV membantu mengurai video menjadi serangkaian gambar statis yang siap untuk diproses lebih lanjut.
- Tahap kedua, mendeteksi wajah menggunakan MTCNN. Selanjutnya, setiap frame yang telah diekstraksi diproses menggunakan MTCNN untuk mendeteksi area wajah. Hasil deteksi berupa bounding box wajah kemudian dipotong dan diresize menjadi ukuran  $96 \times 96$  piksel, sesuai dengan kebutuhan input model OpenFace. Jika jumlah wajah yang terdeteksi kurang dari 10, dilakukan pengulangan atau padding untuk memenuhi jumlah tersebut.
- Tahap ketiga, mengekstraksi fitur wajah menggunakan OpenFace. Pada tahap terakhir, gambar wajah berukuran  $96 \times 96$  piksel diproses menggunakan OpenFace untuk mengekstraksi fitur-fitur wajah yang relevan. Hasil ekstraksi fitur kemudian dipetakan dengan label kepribadian OCEAN berdasarkan data anotasi yang tersedia. Data akhir dikonversi ke dalam format `tf.data.Dataset` TensorFlow, dengan struktur tensor (10, 96, 96, 3) untuk citra wajah dan (5,) untuk label, sehingga mempermudah manajemen batch dan mempercepat proses pelatihan model prediksi kepribadian.

### 3.3. Pengembangan Model



Gambar 3. Alur pengembangan model

Lingkungan pengembangan dalam penelitian ini menggunakan Visual Studio Code sebagai Integrated Development Environment (IDE) utama. Dari sisi perangkat keras, sistem ini memanfaatkan prosesor Intel Core i7 atau AMD Ryzen 7 dengan 8 core untuk menangani komputasi intensif, didukung oleh GPU NVIDIA RTX 3060 berkapasitas 12 GB VRAM guna

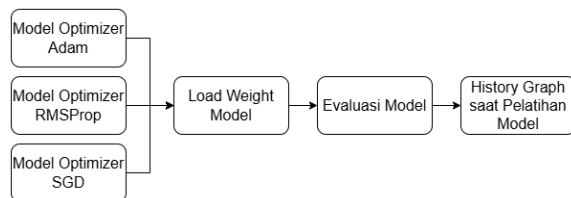
mempercepat proses pelatihan model deep learning. Kapasitas RAM sebesar 32 GB memungkinkan pengolahan dataset berukuran besar secara efisien, sementara SSD 1 TB digunakan untuk memastikan kecepatan akses data yang optimal. Dalam pengembangan model pada gambar 3 terdapat 6 tahap, diantaranya:

- Tahap pertama, import library. Proses pengembangan model diawali dengan mengimpor seluruh pustaka yang diperlukan, termasuk TensorFlow/Keras untuk pembangunan model, OpenCV dan MTCNN untuk pemrosesan gambar, serta library pendukung seperti NumPy, Matplotlib, dan os untuk manajemen data dan visualisasi. Pustaka tambahan seperti OpenFace juga diimpor untuk ekstraksi fitur wajah.
- Tahap kedua, load dataset. Dataset yang telah melalui tahap pre-processing dimuat dari direktori lokal dalam format `tf.data.Dataset`, yang terbagi menjadi `train_ds` (data pelatihan) dan `val_ds` (data validasi). Data ini berisi urutan 10 frame wajah (berukuran  $96 \times 96$  piksel) beserta label kepribadian OCEAN, siap digunakan untuk pelatihan model.
- Tahap ketiga, membangun model OpenFace. Model OpenFace dimuat dengan arsitektur yang mengacu pada implementasi orisinal dari repository GitHub `cmusatyalab`. Model ini berfungsi sebagai ekstraktor fitur utama, mengubah setiap frame wajah menjadi vektor 128-dimensi. OpenFace digunakan dalam mode beku (*freeze*) agar bobotnya tidak berubah selama pelatihan, sehingga fokus pembelajaran hanya pada lapisan berikutnya.
- Tahap keempat, menambahkan LSTM pada model. Setelah ekstraksi fitur, sebuah lapisan LSTM diintegrasikan untuk memproses data sekuensial dari 10 frame berturut-turut. Input LSTM berbentuk (10, 128), memungkinkan model menangkap dinamika temporal ekspresi wajah yang tidak terdeteksi oleh analisis spasial. Lapisan *fully connected* (Dense) kemudian memetakan output LSTM ke lima nilai prediksi OCEAN.
- Tahap kelima, kompilasi model. Model dikompilasi menggunakan tiga optimizer (Adam, RMSProp, dan SGD) untuk keperluan perbandingan kinerja. Fungsi loss yang digunakan adalah Mean Absolute Error (MAE), sesuai dengan sifat masalah regresi dalam prediksi kepribadian. Metrik evaluasi seperti MAE dan RMSE juga ditambahkan untuk memantau performa model.
- Tahap keenam, pelatihan model. Pelatihan dilakukan dengan `batch_size=8` dan `epoch=100`, menggunakan mekanisme callback penting seperti *EarlyStopping* (menghentikan pelatihan jika tidak ada peningkatan validasi loss) dan *ModelCheckpoint* (menyimpan bobot terbaik).

dalam format .t5). Setelah pelatihan selesai, bobot model terbaik dimuat kembali untuk evaluasi akhir, memastikan model siap digunakan dalam prediksi kepribadian berbasis video.

Dengan keenam tahap ini, alur pengembangan model menjadi lebih terstruktur, mulai dari persiapan data hingga pelatihan dan penyimpanan model akhir.

### 3.4. Evaluasi



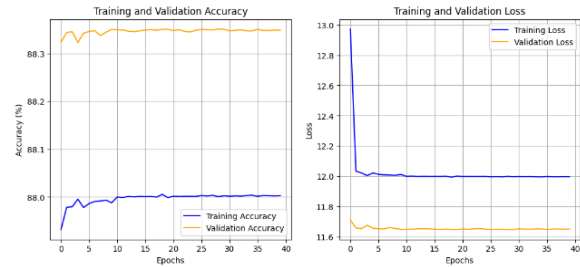
Gambar 4. Alur evaluasi model

Proses evaluasi model dilakukan untuk mengukur performa model dalam memprediksi kepribadian menggunakan tiga subset data: train\_ds, val\_ds, dan test\_ds. Gambar 4 merupakan alur evaluasi model yang memiliki 3 tahap, antara lain:

- Tahap pertama *load weight model*. Proses evaluasi diawali dengan memuat bobot model terbaik dari ketiga optimizer yang telah disimpan dalam format .t5 menggunakan `tf.data.Dataset.load()`. Bobot ini dipilih berdasarkan performa validasi terbaik selama pelatihan (melalui *callback ModelCheckpoint*), memastikan model yang dievaluasi merupakan versi optimal dengan generalisasi tertinggi.
- Tahap kedua, evaluasi model. Kinerja model diukur dengan menghitung loss (MAE) dan akurasi pada data latih, validasi, dan uji. Selain itu, dilakukan analisis per-dimensi kepribadian (OCEAN) untuk mengidentifikasi kelemahan prediksi pada trait tertentu. Hasil evaluasi kuantitatif ini menjadi dasar penyempurnaan model.
- Tahap ketiga, *history graph* pelatihan. Proses dilengkapi dengan visualisasi history graph yang menampilkan tren akurasi dan loss selama pelatihan. Grafik ini membantu menganalisis *overfitting* (jika loss validasi meningkat), *underfitting* (jika loss stagnan), atau konvergensi optimal. Sehingga memberikan insight komprehensif tentang dinamika pembelajaran model.

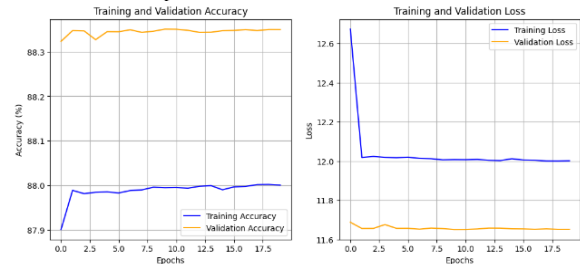
Dengan ketiga tahap ini, evaluasi model mencakup aspek teknis (bobot), statistik (metric), dan visual (grafik), sehingga memudahkan interpretasi hasil secara holistic. Pada gambar 5-7 merupakan hasil *history graph* dari ketiga optimizer:

#### a. Adam



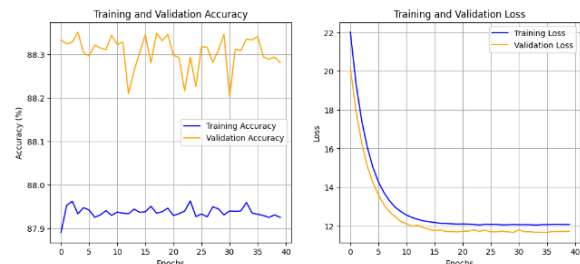
Gambar 5. Grafik optimizer Adam

#### b. RMSProp



Gambar 6. Grafik optimizer RMSProp

#### c. SGD



Gambar 7. Grafik optimizer SGD

## 4. HASIL DAN PEMBAHASAN

Sebagai bagian integral dari proses validasi model, penelitian ini melakukan evaluasi komprehensif terhadap efektivitas berbagai algoritma optimasi pada arsitektur OpenFace menggunakan dataset skala besar yang terdiri dari 10.000 sampel video. Hasil pengujian yang dirangkum dalam Tabel 1 menunjukkan perbandingan mendalam antara beberapa optimizer berdasarkan dua kriteria utama: akurasi prediksi dan efisiensi komputasi (durasi pelatihan).

Table 1. Perbandingan Optimizer

| Optimizer    | Training | Validation | Testing | Time     |
|--------------|----------|------------|---------|----------|
| Adam         | 88,006   | 88,350     | 88,148  | 5m 49.8s |
| RMSProp      | 88,005   | 88,350     | 88,150  | 3m 23.3s |
| SGD          | 88,001   | 88,351     | 88,145  | 6m 23.9s |
| No Optimizer | 88,003   | 88,351     | 88,149  | 3m 41.7s |

Berdasarkan hasil evaluasi yang disajikan dalam tabel, dapat diamati bahwa keempat pendekatan optimasi (Adam, RMSProp, SGD, dan tanpa optimizer) menunjukkan performa yang sangat kompetitif dalam hal akurasi, dengan selisih yang minimal di antara mereka. Pada dataset testing, seluruh metode mencapai

akurasi sekitar 88.14-88.15%, menunjukkan konsistensi yang tinggi dalam kemampuan prediksi model.

Dari segi efisiensi komputasi, RMSProp menonjol sebagai optimizer tercepat dengan waktu pelatihan hanya 3 menit 23.3 detik, mengungguli bahkan pendekatan tanpa optimizer (3 menit 41.7 detik). Hal ini cukup menarik karena biasanya penggunaan optimizer diharapkan akan menambah waktu komputasi. Adam dan SGD memerlukan waktu lebih lama, masing-masing 5 menit 49.8 detik dan 6 menit 23.9 detik.

Terdapat pembahasan mendalam yang mengungkapkan beberapa temuan kunci:

- a. Stabilitas Model: Konsistensi akurasi yang tinggi (88.14-88.15%) pada data testing menunjukkan bahwa arsitektur OpenFace memiliki stabilitas yang baik terlepas dari metode optimasi yang digunakan.
- b. Efisiensi RMSProp: Keunggulan RMSProp dalam hal kecepatan (15% lebih cepat dibanding tanpa optimizer) mungkin disebabkan oleh kemampuannya dalam menyesuaikan learning rate secara adaptif untuk setiap parameter.
- c. Performa Tanpa Optimizer: Hasil yang setara tanpa menggunakan optimizer menimbulkan pertanyaan menarik tentang kompleksitas model dan kemungkinan bahwa gradien sudah cukup baik tanpa perlu teknik optimasi tambahan.
- d. Implikasi Praktis: Untuk implementasi sistem prediksi kepribadian berbasis video, RMSProp muncul sebagai pilihan optimal dengan pertimbangan *balance* antara akurasi dan efisiensi komputasi.

Temuan ini memberikan wawasan berharga bahwa dalam kasus arsitektur OpenFace untuk prediksi kepribadian, pemilihan *optimizer* tidak secara signifikan mempengaruhi akurasi akhir, tetapi berdampak nyata pada efisiensi komputasi. Hal ini membuka peluang untuk optimasi lebih lanjut dalam hal kecepatan proses tanpa mengorbankan kualitas prediksi.

## 5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan sistem identifikasi kepribadian berbasis video wajah menggunakan pendekatan *deep learning*, dengan MTCNN untuk deteksi wajah dan OpenFace untuk ekstraksi fitur wajah (96×96 piksel) menjadi vektor 128 dimensi, yang kemudian diproses oleh LSTM untuk analisis temporal dari 10 frame video. Penggunaan optimizer RMSProp terbukti optimal dengan akurasi prediksi OCEAN mencapai 88% dan nilai *loss* stabil (11.6–11.9), menunjukkan kemampuan generalisasi model yang baik tanpa *overfitting*. Namun, beberapa *frame* video dengan gangguan (gambar putih, *fake face*, atau subjek non-utama) dapat menurunkan akurasi, sehingga disarankan untuk melakukan kurasi data lebih ketat

dan mengeksplorasi model *deep learning* alternatif guna meningkatkan performa sistem.

## DAFTAR PUSTAKA

- [1] W. Tinwala, "Big Five Personality Detection Using Deep Convolutional Neural Networks," no. September, 2021, doi: 10.20944/preprints202109.0199.v1.
- [2] R. John, R. John, and Z.-R. Rao, "THE BIG FIVE PERSONALITY TRAITS AND ACADEMIC PERFORMANCE," *J. Law Soc. Stud.*, vol. 2, no. 1, pp. 10–19, Jun. 2020, doi: 10.52279/jlss.02.01.1019.
- [3] S. R. Nudin, R. Gernowo, M. Somantri, and A. Wibowo, "Multi Task Classification Using Deep Learning Approaches for Big Five Personality Traits Prediction: A Review," *Proc. - 11th Int. Conf. Inf. Technol. Comput. Electr. Eng. ICITACEE 2024*, pp. 37–42, 2024, doi: 10.1109/ICITACEE62763.2024.10762804.
- [4] H. Yaya and W. Teguh, *DASAR-DASAR DEEP LEARNING DAN IMPLEMENTASINYA*, 1st ed. Yogyakarta: Gava Media, 2021.
- [5] D. Chang, Y. Yin, Z. Li, M. Tran, and M. Soleymani, "LibreFace: An Open-Source Toolkit for Deep Facial Expression Analysis," *Proc. - 2024 IEEE Winter Conf. Appl. Comput. Vision, WACV 2024*, pp. 8190–8200, 2024, doi: 10.1109/WACV57701.2024.00802.
- [6] S. Song, Shashank Jaiswal, E. Sanchez, G. Tzimiropoulos, L. Shen, and M. Valstar, "Self-Supervised Learning of Person-Specific Facial Dynamics for Automatic Personality Recognition," *IEEE Trans Affect Comput*, 2021, doi: 10.1109/TAFFC.2021.3064601.
- [7] M. Bhagya Sri and D. Sridevi, "Real-Time Multi-Face Recognition Using Support Vector Machine Through OpenCV," *Int. J. Eng. Sci. Adv. Technol.*, vol. 24, 2024, [Online]. Available: www.ijesat.com
- [8] F. Cahyono, "PENGENALAN WAJAH MENGGUNAKAN MODEL FACENET UNTUK PRESENSI PEGAWAI," Institut Teknologi Sepuluh Nopember, 2020.
- [9] Zhancheng Ren, Q. Shen, Xiaolei Diao, and H. Xu, "A sentiment-aware deep learning approach for personality detection from text," *Inf. Process. Manag.*, vol. 58, no. 3, doi: 10.1016/j.ipm.2021.102532.
- [10] C. Suman, S. Saha, A. Gupta, S. K. Pandey, and P. Bhattacharyya, "A multi-modal personality prediction system," *Knowledge-Based Syst.*, vol. 236, p. 107715, 2022, doi: 10.1016/j.knosys.2021.107715.
- [11] H. Ku and W. Dong, "Face Recognition Based on MTCNN and Convolutional Neural Network," *Front. Signal Process.*, vol. 4, no. 1, Jan. 2020, doi: 10.22606/fsp.2020.41006.
- [12] L. Muflikhah *et al.*, "Single nucleotide polymorphism based on hypertension potential

- risk prediction using LSTM with Adam optimizer,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 33, no. 2, pp. 1126–1139, Feb. 2024, doi: 10.11591/ijeecs.v33.i2.pp1126-1139.
- [13] L. Lingutla *et al.*, “Performance Analysis of Cycle GAN in Photo to Portrait Transfiguration using Deep Learning Optimizers,” *IEEE Access*, vol. 11, pp. 136541–136551, 2023, doi: 10.1109/ACCESS.2023.3337430.
- [14] M. Ramezani, M.R. Feizi-Derakhshi, and M.A. Balafar, “Knowledge Graph-Enabled Text-Based Automatic Personality Prediction,” *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/3732351.
- [15] M. A. M. A. C. A. D. M. H. C. M. Moreno-Sotelo, “Estimation of Personality Traits from Portrait Pictures Using the Five-Factor Model,” *IEEE Access*, vol. 8, pp. 201649–201665, 2020, doi: 10.1109/ACCESS.2020.3034639.
- [16] J. Park *et al.*, “Comparison of the inter-item correlations of the Big Five Inventory-10 (BFI-10) between Western and non-Western contexts,” *Personal. Individ. Differ. (Pers Individ Dif)*, vol. 196, 2022, doi: 10.1016/j.paid.2022.111751.
- [17] R. Mahajan, R. Mahajan, E. Sharma, and V. Mansotra, “Are we tweeting our real selves?” Personality Prediction of Indian Twitter Users using Deep Learning Ensemble Model,” *Comput. Human Behav.*, vol. 128, 2022, doi: 10.1016/j.chb.2021.107101.
- [18] C. Bhatt, I. Kumar, V. Vijayakumar, K. U. Singh, and A. Kumar, “The state of the art of deep learning models in medical science and their challenges,” *Multimed. Syst.*, vol. 27, no. 4, pp. 599–613, 2021, doi: 10.1007/s00530-020-00694-1.
- [19] A. Ouarka, T. Ait, B. Youssef, E. Saady, and M. El Hajji, “A deep multimodal fusion method for personality traits prediction,” *Multimed. Tools Appl.*, no. 0123456789, 2024, doi: 10.1007/s11042-024-20356-y.
- [20] “Dataset First Impressions V2 (CVPR’17),” Chalearn Looking at People. Accessed: Nov. 07, 2024. [Online]. Available: <https://chalearnlap.cvc.uab.cat/dataset/24/description/>