

Implementasi MTCNN dan Transfer Learning Model DeepFace untuk Prediksi Kepribadian Berbasis Video

Implementation of MTCNN and DeepFace Transfer Learning Model for Video-Based Personality Prediction

Shandy Ilham Alamsyah^{*1}, Salamun Rohman Nudin²

Universitas Negeri Surabaya

*E-mail : shandy.21015@mhs.unesa.ac.id^{*1}, salamunrohman@unesa.ac.id²*

**Corresponding author*

Received 16 June 2025; Revised 19 July 2025; Accepted 27 July 2025

Abstrak - Kepribadian adalah aspek penting yang mempengaruhi pilihan hidup, karir, kinerja, kesehatan, dan juga preferensi atau keinginan seseorang. Model Big-Five Personality adalah yang paling umum, namun pengukurannya masih secara konvensional melalui kuesioner, hal ini memiliki beberapa keterbatasan seperti adanya potensi manipulasi jawaban oleh responden sehingga mempengaruhi hasil dari pengukuran kepribadian tersebut. Untuk mengatasi keterbatasan tersebut, penelitian ini bertujuan untuk mengembangkan sistem untuk melakukan pengukuran atau prediksi kepribadian menggunakan Deep Learning untuk mendeteksi kepribadian berdasarkan ekspresi wajah dalam sebuah video perkenalan. Model yang dikembangkan mencapai akurasi 90.04% dengan loss terendah 9.95%, menunjukkan kemampuannya dalam memprediksi kepribadian secara konsisten. Sistem ini dibangun dengan framework Flask dan mampu menghasilkan prediksi kepribadian seseorang. Dengan demikian penggunaan Deep Learning berpotensi menjadi alat yang efektif dalam pengembangan teknologi di bidang psikologi, menjadikannya alat yang transformatif untuk mengukur kepribadian seseorang dengan lebih efektif di masa depan.

Keywords - Big-Five Personality, Deep Learning, MTCNN, DeepFace, deteksi kepribadian

Abstract - *Personality is an important aspect that influences a person's life choices, career, performance, health, and also preferences or desires. The Big-Five Personality model is the most common, but its measurement is still conventionally through questionnaires, this has several limitations such as the potential for manipulation of answers by respondents so that it affects the results of the personality measurement. To overcome these limitations, this research aims to develop a system to measure or predict personality using Deep Learning to detect personality based on facial expressions in an introduction video. The developed model achieved an accuracy of 90.04% with the lowest loss of 9.95%, demonstrating its ability to predict personality consistently. The system is built with the Flask framework and is able to generate predictions of a person's personality. Thus, the use of Deep Learning has the potential to be an effective tool in the development of technology in the field of psychology, making it a transformative tool for more effective personality assessment in the future.*

Keywords - Big-Five Personality, Deep Learning, MTCNN, DeepFace, Personality Detector

1. PENDAHULUAN

Kepribadian bisa didefinisikan sebagai kombinasi dari perilaku, karakteristik, emosi, dan pola pikir seseorang [1]. Memahami kepribadian seseorang bisa menjadi aspek yang cukup penting dalam berbagai bidang, mulai dari psikologi, forensik, politik bahkan sampai ke interaksi manusia dan komputer. Kepribadian menjadi faktor yang dapat menentukan bagaimana seseorang memilih pilihan hidup atau karirnya, kinerja, kesehatan, dan juga preferensi atau keinginan

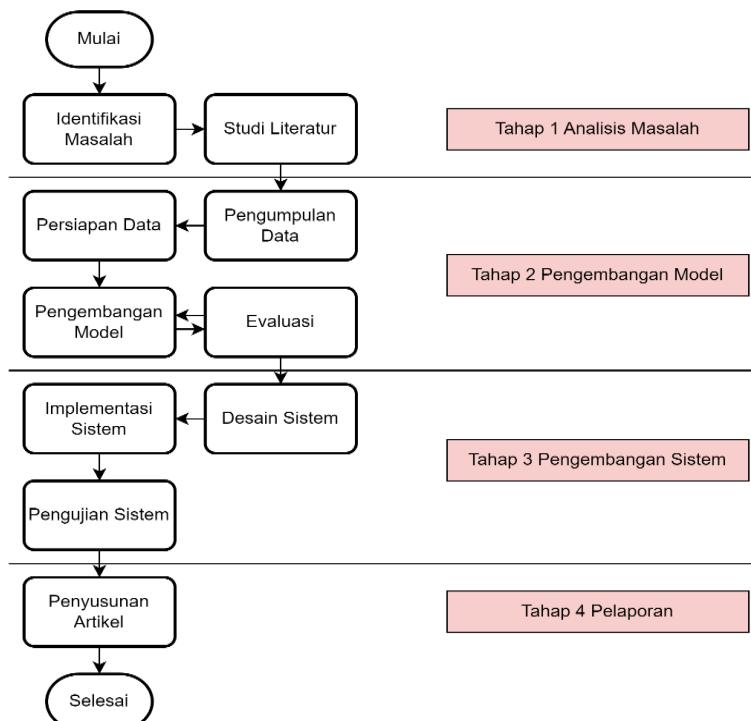
seseorang [2]. Big-Five Personality merupakan model yang terdiri dari 5 dimensi utama yaitu Openness (keterbukaan terhadap hal baru), Conscientiousness (menggambarkan keteraturan dan tanggung jawab), Extraversion (kecenderungan bersosialisasi), Agreeableness (interaksi sosial dalam suatu hubungan), dan Neuroticism (stabilitas emosional) atau yang sering dikenal dengan istilah OCEAN [3]. Big-Five Personality telah menjadi salah satu kerangka atau model kepribadian yang paling banyak diadaptasi dalam literatur serta bisa diterima dan diakui banyak kalangan psikolog dan peneliti di seluruh dunia, dikarenakan konsisten dan akan tetap relevan dalam kehidupan seseorang [4], [5]. Saat ini pengukuran kepribadian sering kali dilakukan melalui kuesioner, misalkan BFI ataupun NEO-PI. Namun, sistem ini memiliki beberapa keterbatasan yang dapat mempengaruhi hasil dari prediksi atau pengukuran kepribadian, misal faktor seperti kejujuran dan kondisi psikologis responden. Dengan hanya mengisi kuesioner, responden bisa saja tidak memberikan jawaban yang sesuai dengan kondisi yang sebenarnya [6].

Kemajuan teknologi seperti saat ini, khususnya dengan berkembangnya deep learning dalam computer vision berupa face recognition, dapat memunculkan peluang untuk mengembangkan sebuah metode baru yang memungkinkan komputer untuk memahami dan menganalisis gambar atau video serupa dengan manusia [7]. Banyak penelitian yang mengaitkan hubungan antara wajah dengan kepribadian yang dimiliki seseorang, dikarenakan beberapa alasan teoritis. Pertama, karena korelasi genetik karakteristik kraniofasial, jadi selain membentuk wajah, gen berperan dalam pembentukan sifat kepribadian seseorang. Kedua, beberapa bukti menunjukkan bahwa hormon seperti testosteron, estrogen, prenata, dan postnatal, memang mempengaruhi perilaku dan fitur wajah. Ketiga, persepsi fitur wajah, sama halnya dengan persepsi ‘kepandaian’ seseorang yang dapat digunakan untuk prediksi pencapaian pendidikan, prasangka atau persepsi fitur wajah seseorang bisa mengarah pada pengembangan karakteristik seseorang yang maladaptif (misal ‘Quasimodo Complex’). Keempat, kepribadian seseorang bisa dikaitkan dengan perilaku ekspresif secara emosional, yang mengarah pada perkembangan otot wajah dan kerutan [8]. Penggunaan video bisa dijadikan sebagai sumber data untuk menganalisis dan memprediksi kepribadian seseorang, terutama dalam konteks Big-Five Personality ini. Karena terdapat sebuah penelitian sebelumnya [1], dimana penelitian tersebut menggunakan potongan video pendek yang digunakan untuk menganalisis kepribadian seseorang dalam video tersebut. Penelitian tersebut mengembangkan sistem multimodal yang menggabungkan fitur-fitur yang dihasilkan oleh modalitas yang berbeda, dengan mengekstrak ambient dari video menggunakan ResNet, fitur wajah menggunakan MTCNN, fitur audio menggunakan VGGish CNN, dan fitur teks menggunakan n-gram melalui CNN. Semua fitur tersebut akan digabungkan menggunakan mekanisme berbasis attention serta konsolidasi menggunakan fully connected layer dan sigmoid. Dan akurasi yang dicapai menggunakan sistem tersebut sebesar 91,43% menggunakan dataset Chalearn-17. Walaupun pendekatan multimodal seperti tersebut terbukti sangat efektif, namun kompleksitas dan tuntutan komputasi yang tinggi menjadi kendala praktis dalam penerapannya. Hal ini lantas memunculkan pertanyaan: mungkinkah sebuah pendekatan yang lebih sederhana, dengan hanya fokus pada satu modalitas seperti analisis visual, tetapi mampu memberikan prediksi yang andal? Oleh karena itu, penelitian ini dirancang untuk menjawab pertanyaan tersebut dengan mengisolasi dan mengevaluasi efektivitas pendekatan unimodal (hanya visual), menggunakan MTCNN dan DeepFace untuk memprediksi kepribadian Big-Five. Tujuannya adalah untuk melihat apakah model yang lebih ramping ini dapat menjadi alternatif terhadap arsitektur multimodal yang kompleks.

Penelitian ini akan menggunakan MTCNN (Multi-task Cascaded Convolutional Networks) dan DeepFace yang digunakan untuk menentukan dan menganalisis fitur wajah dari video, dimana MTCNN akan mendeteksi wajah dalam video [9], dan DeepFace akan membantu mengekstraksi fitur-fitur wajah yang terdeteksi [10]. Untuk model regresi dalam memprediksi dimensi Big-Five Personality berdasarkan fitur wajah yang telah diekstraksi akan menggunakan mean absolute error (MAE) sebagai metrik evaluasi. Algoritma SGD (Stochastic Gradient Descent) akan digunakan untuk mengoptimisasi atau meningkatkan akurasi dari model tersebut. Dengan metode di atas diharapkan penelitian ini akan memberikan hasil yang efektif, dapat memberikan kontribusi dalam pengembangan ilmu terkait kecerdasan buatan, khususnya yang memanfaatkan model DeepFace, dan bisa menjadi dasar untuk penelitian lanjutan. Hal ini juga membuka peluang pengaplikasian praktis, misalkan rekrutmen tenaga kerja atau analisis psikologi. Batasan penelitian ini meliputi : analisis kepribadian penelitian ini hanya berfokus pada Big-Five Personality, evaluasi difokuskan pada hasil regresi yang diperoleh, dan dataset yang digunakan menggunakan dataset sekunder yang diambil dari Chalearn CVPR'17, yang kemungkinan memiliki demografi yang tidak seimbang, terutama terkait usia dan etnis. Terkait hal tersebut, penelitian ini tidak melakukan evaluasi spesifik terhadap potensi bias dataset, sehingga generalisasi hasil ke populasi yang lebih beragam mungkin terbatas.

2. METODE PENELITIAN

Metodologi penelitian ini melalui beberapa tahapan yang diadaptasi dari metodologi IBM Data Science [11], yang menghasilkan alur seperti pada Gambar 1 :



Gambar 1 Metodologi Penelitian

2.1 Tahap 1 Analisis Masalah

Penelitian diawali dengan mengidentifikasi masalah, memilih topik, merumuskan masalah, dan menetapkan tujuan penelitian. Tahapan ini didukung oleh studi literatur komprehensif terkait prediksi Big Five Personality berbasis video menggunakan transfer learning DeepFace. Studi literatur ini bertujuan untuk memperdalam pemahaman konsep teori, metode, dan temuan penelitian yang relevan, seperti kecenderungan Big Five Personality [3], penggunaan model DeepFace untuk ekstraksi fitur wajah [12], serta aplikasi deep learning dalam mengidentifikasi kepribadian [13], [14]. Observasi awal juga dilakukan untuk memvalidasi kelayakan prediksi Big Five Personality melalui analisis fitur wajah dari sebuah video.

2.2 Tahap 2 Pengembangan Model

Tahap ini mengadopsi pendekatan IBM Data Science Methodology dalam pengolahan data dan pengembangan solusi, meliputi pengumpulan data, persiapan data, pengembangan model, dan evaluasi.

a. Pengumpulan Data

Penelitian ini menggunakan dataset sekunder bernama "First Impressions V2" yang diperkenalkan pada Konferensi Computer Vision and Pattern Recognition (CVPR) 2017. Dataset ini tersedia secara publik melalui situs web ChaLearn Looking at People (LAP) dan berisi 10.000 klip video pendek dengan durasi sekitar 15 detik. Setiap video telah dianotasi oleh para ahli dengan skor kepribadian Big Five (OCEAN) bernilai antara 0 dan 1, yang berfungsi sebagai ground truth. Untuk kebutuhan pelatihan dan pengujian, dataset dibagi dengan rasio 3:1:1 untuk data latih, validasi, dan uji.

b. Persiapan Data

Data video mentah memerlukan beberapa tahapan persiapan untuk mengekstrak informasi yang relevan. Proses ini memanfaatkan MTCNN untuk deteksi wajah. Multi-task Cascaded Convolutional Networks (MTCNN) adalah sebuah arsitektur deep learning yang digunakan untuk mendeteksi wajah secara akurat dan cepat. Metode ini bekerja melalui tiga jaringan yang tersusun secara kaskade (P-Net, R-Net, dan O-Net) yang secara progresif menyaring dan memperbaiki lokasi wajah pada sebuah gambar [9].

Dalam implementasinya, untuk setiap video diekstrak 10 frame gambar menggunakan OpenCV [15], kemudian wajah pada setiap frame dideteksi menggunakan MTCNN dan ukurannya disesuaikan menjadi 152x152 piksel. Untuk menjaga konsistensi data, video yang terlalu pendek akan mengalami duplikasi frame, dan jika wajah tidak terdeteksi pada sebuah frame, sistem akan menggantinya dengan gambar wajah valid dari frame sebelumnya. Akhirnya, data yang telah bersih ini disusun ke dalam format tf.data.Dataset, di mana setiap sampel terdiri dari tensor gambar (10, 152, 152, 3) dan tensor label (5,).

c. Pengembangan Model

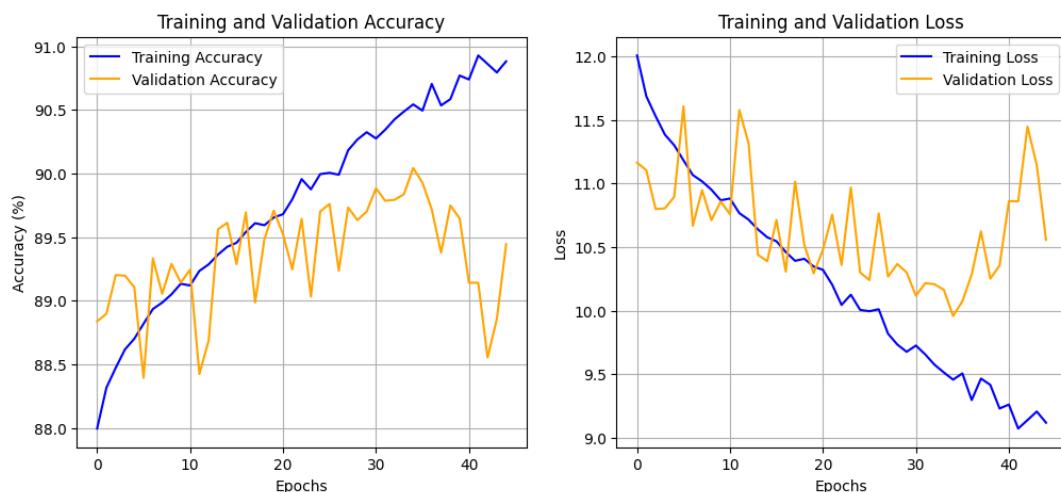
Model prediksi dibangun dengan pendekatan transfer learning menggunakan DeepFace. DeepFace adalah sebuah arsitektur Convolutional Neural Network (CNN) yang dikembangkan oleh Facebook AI, yang terdiri dari 9 lapisan untuk tugas pengenalan wajah.

Model ini dirancang untuk memproses gambar wajah dan mengubahnya menjadi representasi fitur numerik (dikenal sebagai embedding) yang detail. Proses ini dilakukan melalui serangkaian lapisan konvolusi untuk mendeteksi pola visual wajah (seperti tepi dan bentuk) hingga lapisan fully connected yang menginterpretasikan fitur-fitur tersebut menjadi sebuah vektor [10].

Dalam penelitian ini, pendekatan transfer learning dimanfaatkan dengan menggunakan model DeepFace yang telah dilatih sebelumnya (pre-trained) sebagai ekstraktor fitur tetap. Hal ini dilakukan dengan membekukan bobotnya (trainable = False). Model kemudian dilatih menggunakan optimizer SGD dengan parameter batch_size=8 dan epochs=100, serta menerapkan callbacks EarlyStopping dan ModelCheckpoint untuk mengontrol proses pelatihan dan mencegah overfitting.

d. Evaluasi

Kinerja model dievaluasi menggunakan metrik Mean Absolute Error (MAE). MAE merupakan metrik yang paling sesuai untuk penelitian ini karena output model bersifat kontinu (regresi). Metrik ini bekerja dengan mengukur rata-rata selisih absolut antara nilai prediksi dengan nilai anotasi sebenarnya, di mana nilai MAE yang lebih rendah menunjukkan performa model yang lebih akurat. Gambar 2 menunjukkan grafik performa pelatihan model yang telah dilakukan pada penelitian ini :



Gambar 2 Grafik Performa Pelatihan Model

2.3 Tahap 3 Pengembangan Sistem

Tahap ini berfokus pada pengembangan sistem menggunakan framework Flask. Desain sistem dimulai dengan analisis kebutuhan perangkat keras dan perangkat lunak, serta analisis input (video) dan output (prediksi kepribadian). Desain alur kerja sistem dan desain user interface menjadi acuan pengembangan. Implementasi sistem melibatkan penyusunan kode menggunakan Flask untuk membuat website yang memungkinkan pengguna mengunggah video, yang kemudian dianalisis oleh model dan hasilnya ditampilkan. Pengujian sistem dilakukan untuk memastikan semua fitur berfungsi dengan baik sesuai desain.

2.4 Tahap 4 Pelaporan

Tahap terakhir dari penelitian ini adalah penyusunan laporan, yang merupakan dokumentasi komprehensif dari semua tahapan yang telah dilalui dan hasil yang diperoleh. Pada fase ini, seluruh temuan dari analisis masalah, proses pengembangan model, termasuk detail pengumpulan dan persiapan data, arsitektur model, serta metodologi pelatihan dan evaluasi, akan didokumentasikan secara rinci.

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil dari 2 tahapan utama penelitian ini, yaitu pengembangan model dan implementasi sistem beserta pembahasannya. Hasil dari kedua tahapan tersebut adalah dasar penilaian pengembangan sistem secara keseluruhan. Kemudian terdapat beberapa pengujian yang dilakukan untuk evaluasi langsung terhadap model yang telah diimplementasikan ke dalam website, berikut adalah pembahasan mengenai pengujian yang dilakukan pada penelitian ini.

3.1 Hasil Pengembangan Model

Bagian ini berisi tentang hasil dari pengembangan model yang dilakukan pada bab sebelumnya, pengembangan model ini menggunakan beberapa optimizer untuk pelatihan model sebagai perbandingan optimizer terbaik, dan hasilnya ditunjukkan pada tabel 1:

Tabel 1 Hasil Pengembangan Model

Optimizer	Hasil		
	Train	Val	Test
Tanpa Optimizer	91,983	89,947	89,726
RectifiedAdam	87,980	88,324	88,128
Adam	91,536	89,985	89,765
SGD	91,363	90,042	89,829
RMSProp	91,218	89,771	89,617

Berdasarkan Tabel 1, hasil pengembangan model disajikan berdasarkan pelatihan menggunakan berbagai *optimizer* untuk membandingkan kinerjanya. Terlihat bahwa tanpa penggunaan *optimizer*, model mencapai akurasi *train* tertinggi sebesar 91,983%, namun akurasi *validation* dan *test* sedikit lebih rendah, mengindikasikan potensi overfitting, meskipun dalam batas wajar. SGD menunjukkan keseimbangan yang optimal antara *train* dengan akurasi *validation* dan *test*. Akurasi *validation* dan *test* SGD merupakan yang tertinggi di antara semua *optimizer* yang diuji, menegaskan kemampuannya untuk menggeneralisasi data baru dengan baik. *Optimizer* lain seperti Adam dan RMSProp juga menunjukkan kinerja yang baik pada *train*, namun akurasi *validation* dan *test*nya cenderung di bawah SGD. Sementara itu RectifiedAdam menunjukkan akurasi yang paling rendah di semua fase, menunjukkan arsitektur ini mungkin kurang cocok untuk konfigurasi model yang digunakan dalam penelitian ini. Secara keseluruhan, pemilihan optimier memiliki dampak terhadap kemampuan generalisasi model, dan SGD terbukti menjadi pilihan paling efektif dan stabil.

3.2 Pengujian Black Box

Tahapan selanjutnya adalah pengujian sistem, dimana sistem yang telah dikembangkan akan diuji dengan beberapa skenario untuk melihat apakah fitur-fiturnya dapat berjalan dengan baik atau tidak. Pada tahap ini pengujian dilakukan dengan menggunakan Black Box Testing, dimana sistem dijalankan pada device dengan spesifikasi CPU Intel Core i5-1235U, GPU Nvidia MX550, dan RAM 16 GB. Adapun skenario pengujian yang akan diujikan dan statusnya pada tabel 2 :

Tabel 2 Pengujian Black Box

No.	Skenario Pengujian	Status
1.	Tab upload dan rekam video bisa digunakan	Berhasil
2.	Panel pilih file ketika upload bisa digunakan	Berhasil
3.	Tombol buka kamera dapat mengakses kamera	Berhasil
4.	Input video dapat dilakukan	Berhasil
5.	Indikator loading saat melakukan prediksi	Berhasil
6.	Hasil prediksi dapat ditampilkan	Berhasil
7.	File CSV dan PDF bisa diunduh	Berhasil

3.3 Skenario Pengujian Anotasi

Pada pengujian ini dilakukan analisis kepribadian dengan menggunakan video dari dataset yang dilengkapi dengan anotasi dari anotator. Pengujian ini sendiri bertujuan untuk menilai seberapa dekat nilai hasil prediksi model terhadap label kepribadian yang sebenarnya (yang ada di file anotasi), dan menjadi acuan sebelum melakukan pengujian dengan kondisi yang lebih kompleks atau tanpa anotasi. Sebagai contoh digunakan video dengan nama ZhPQrqxo2mw.003.mp4, yang bisa dirunutkan dengan Tabel 3 berikut :

Tabel 3 Tabel Perbandingan Anotasi dan Prediksi

Trait	Anotasi	Prediksi	Selisih
Openness	0.7556	0.6995	0.0561
Conscientiousness	0.7767	0.6602	0.1165
Extraversion	0.5794	0.5683	0.0111
Agreeableness	0.7473	0.6257	0.1216
Neuroticism	0.6875	0.5841	0.1034

Tabel 3 menunjukkan bahwa model berhasil memprediksi kepribadian dengan selisih yang masih relatif kecil dari nilai anotasinya. Nilai extraversion memiliki selisih yang paling kecil dibandingkan dengan yang lainnya, yaitu sebesar 0.0111 atau sekitar 1.11%. walaupun ada beberapa trait menunjukkan selisih yang lebih besar seperti agreeableness yaitu hingga 0.1216 atau 12,16%, namun rata-rata traitnya secara umum masih diambil normal, yakni hanya 0.0817 atau sekitar 8.17%, nilai ini masih dibawah 0.15 atau 15% yang masih dapat diterima dalam konteks prediksi berbasis visual tidak langsung.

3.4 Skenario Pengujian Ukuran

Pada pengujian ini dilakukan dengan mengevaluasi model terhadap variasi ukuran resolusi video, dan dalam skenario ini dilakukan dengan video landscape berukuran 1280 x 720, dan video potrait berukuran 408 x 720. Pengujian ini dilakukan untuk memastikan bahwa performa model stabil ketika digunakan dalam berbagai kondisi input video yang umum dijumpai.

Sebagai bahan mendemonstrasikan pengujian ini, telah disiapkan video dari dataset yang memiliki resolusi 1280 x 720 dan resolusi 408 x 720. Dari pengujian tersebut menunjukkan bahwa tidak ada masalah saat proses ekstraksi wajah, baik pada resolusi 1280 x 720, maupun resolusi 408 x 720. Hal ini menjelaskan bahwa fungsi deteksi wajah pada sistem bekerja secara baik dan konsisten, walaupun terdapat perbedaan resolusi. Adapun hasil prediksinya yang ada di website yang dirunutkan dengan Tabel 4 berikut :

Tabel 4 Hasil Prediksi Video Uji Skenario Ukuran

Nama Video	Trait	Prediksi	Anotasi	Selisih
9KAqOrdiZ4I.003 (408x720)	Openness	0.3998	0.2777	0.1221
	Conscientiousness	0.2573	0.1941	0.0632
	Extraversion	0.2538	0.2897	0.0359
	Agreeableness	0.3113	0.2197	0.0916
	Neuroticism	0.2346	0.2083	0.0263
9cj8rrj1wSk.001 (1280x720)	Openness	0.7099	0.6111	0.0988
	Conscientiousness	0.6945	0.7087	0.0142
	Extraversion	0.5781	0.6355	0.0574
	Agreeableness	0.6385	0.7363	0.0978
	Neuroticism	0.6001	0.6667	0.0666

Hasil tersebut menunjukkan bahwa model mampu memberikan prediksi yang relatif akurat, karena rata-rata selisihnya hanya 0,0678 untuk video 9KAqOrdiZ4I.003 dan 0,0669 untuk video 9cj8rrj1wSk.001, meskipun dengan resolusi video yang berbeda. Selisih rata-rata untuk masing-masing trait juga menunjukkan bahwa sistem cukup robust dalam menangani input video dengan beberapa variasi yang berbeda.

3.5 Skenario Pengujian Pencahayaan

Pada pengujian ini dilakukan untuk mengevaluasi model terhadap variasi video dengan kondisi pencahayaan yang berbeda, karena di situasi nyata, pencahayaan video bisa sangat bervariasi, tergantung kondisi lingkungan maupun kualitas perangkat kamera yang digunakan.

Sebagai bahan mendemonstrasikan pengujian ini, telah disiapkan video dari dataset yang memiliki perbedaan pencahayaan yang cukup signifikan, dimana satu video dengan pencahayaan yang minim, dan satu lagi dengan pencahayaan sangat terang (bahkan cenderung overexposed). Dari pengujian tersebut menunjukkan bahwa tidak ada masalah saat proses ekstraksi wajah, baik pada pencahayaan terang maupun pencahayaan rendah. Hal ini menjelaskan bahwa fungsi deteksi wajah pada sistem bekerja secara baik dan konsisten, walaupun terdapat perbedaan pencahayaan. Adapun hasil prediksinya yang ada di website yang dirumutkan dengan Tabel 5 berikut :

Tabel 5 Hasil prediksi video uji skenario ukuran

Nama Video	Trait	Prediksi	Anotasi	Selisih
9J-KIPMQmqk.002 (Terang)	Openness	0.4420	0.6889	0.2469
	Conscientiousness	0.3183	0.7961	0.4778
	Extraversion	0.2898	0.4579	0.1681
	Agreeableness	0.3601	0.6044	0.2443
	Neuroticism	0.2831	0.6875	0.4044
1Lv72Si4GnY.000 (Gelap)	Openness	0.4456	0.4111	0.0345
	Conscientiousness	0.3233	0.1068	0.2165
	Extraversion	0.2954	0.4673	0.1719
	Agreeableness	0.3644	0.2637	0.1007
	Neuroticism	0.2878	0.3229	0.0351

Berdasarkan tabel 5, dapat disimpulkan bahwa sistem cenderung lebih ideal dengan pencahayaan gelap, dengan selisih yang hanya 0.11 atau sekitar 11%, atau masih dalam batas wajar untuk selisih prediksi. Pengujian ini menunjukkan meskipun pencahayaan terang umumnya dianggap lebih ideal, namun jika kualitas visualnya menunjukkan pencahayaan berlebih (overexposure) dapat mengganggu proses ekstraksi fitur wajah yang akan berdampak pada akurasi prediksi, bisa ditunjukkan dengan video 9J-KIPMQmqk.002 yang selisih antara anotasi dan prediksinya mencapai 0.30 atau sekitar 30% yang bisa dinilai sangat tinggi. Hal ini bisa disebabkan karena hilangnya detail wajah karena cahaya langsung yang berkontras tinggi, sehingga fitur-fitur pentingnya menjadi kurang terdeteksi oleh sistem.

Akan tetapi kedua kondisi anomali ini masih tidak se-optimal ketika pencahayaan normal pada pengujian sebelumnya, yang bisa dilihat hasilnya pada tabel 4, dimana selisihnya hanya

sekitar 6%. Dengan demikian, pengujian ini menunjukkan bahwa pencahayaan merupakan faktor yang sangat penting yang dapat mempengaruhi performa deteksi sistem.

4. KESIMPULAN

Setelah melakukan penelitian ini, dapat disimpulkan bahwa Deep Learning bisa menjadi metode utama untuk menganalisis objek berbasis gambar yang diperoleh dari dataset video. Penggunaan model DeepFace sebagai pengembangan struktur deep learning dalam memprediksi kepribadian menghasilkan akurasi model sebesar 90.04% dan nilai loss terendah sebesar 9.95%. Hasil ini menunjukkan kemampuan model dalam memberikan hasil yang konsisten dan dapat digunakan untuk menganalisis kepribadian secara visual. Dengan demikian, penggunaan DeepFace dapat menjadi alat efektif untuk mengembangkan teknologi dibidang psikologi. Selain itu, pembangunan sistem prediksi kepribadian berbasis video ini berhasil dilakukan dengan menggunakan framework Flask. Proses ini mencakup integrasi model ke dalam sistem dilakukan melalui konversi ke format t5, yang menghasilkan sistem yang mampu menghasilkan output berupa prediksi kepribadian dari seseorang.

Berdasarkan temuan selama penelitian, terdapat beberapa saran untuk pengembangan lebih lanjut. Pertama, disarankan agar dilakukan kurasi data video untuk meminimalkan file dengan kerusakan mencolok (seperti frame putih atau wajah palsu) guna meningkatkan kualitas dataset. Kedua, penelitian selanjutnya disarankan untuk secara eksplisit mengevaluasi performa model pada berbagai kelompok demografis (usia, gender, dan etnis) untuk mengidentifikasi dan memitigasi potensi bias, sehingga dapat meningkatkan validitas dan keadilan model. Ketiga, menggunakan model deep learning lain sebagai perbandingan dapat membuka peluang untuk menemukan arsitektur yang lebih optimal. Terakhir, perlu dilakukan optimasi lebih lanjut untuk meningkatkan akurasi prediksi pada video dengan kondisi pencahayaan yang sangat terang, di mana performa model saat ini masih menunjukkan penurunan signifikan.

DAFTAR PUSTAKA

- [1] C. Suman, S. Saha, A. Gupta, S. K. Pandey, and P. Bhattacharyya, “A multi-modal personality prediction system,” *Knowledge-Based Systems*, vol. 236, p. 107715, Jan. 2022, doi: 10.1016/j.knosys.2021.107715.
- [2] Y. Mehta, N. Majumder, A. Gelbukh, and E. Cambria, “Recent trends in deep learning based personality detection,” *Artif Intell Rev*, vol. 53, no. 4, pp. 2313–2339, Apr. 2020, doi: 10.1007/s10462-019-09770-z.
- [3] W. Tinwala and S. Rauniyar, “Big Five Personality Detection Using Deep Convolutional Neural Networks,” Sep. 13, 2021. doi: 10.20944/preprints202109.0199.v1.
- [4] R. R. McCrae and P. T. Costa, “A Five-Factor theory of personality,” in *Handbook of personality: Theory and research*, 2nd ed., Guilford Press, 1999, pp. 139–153.
- [5] R. R. McCrae and P. T. Costa, “Understanding persons: From Stern’s personalistics to Five-Factor Theory,” *Personality and Individual Differences*, vol. 169, p. 109816, Feb. 2021, doi: 10.1016/j.paid.2020.109816.
- [6] W. Yeye, C. Deyuan, L. Baobin, W. Xiaoyang, L. Xiaoqian, and Z. Tingshao, “Predicting Personality based on Self-Introduction Video,” *IFAC-PapersOnLine*, vol. 53, no. 5, pp. 452–457, 2020, doi: 10.1016/j.ifacol.2021.04.217.
- [7] D. Bhatt *et al.*, “CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope,” *Electronics*, vol. 10, no. 20, p. 2470, Oct. 2021, doi: 10.3390/electronics10202470.

- [8] A. Kachur, E. Osin, D. Davydov, K. Shutilov, and A. Novokshonov, “Assessing the Big Five personality traits using real-life static facial images,” *Sci Rep*, vol. 10, no. 1, p. 8487, May 2020, doi: 10.1038/s41598-020-65358-6.
- [9] N. Zhang, J. Luo, and W. Gao, “Research on Face Detection Technology Based on MTCNN,” in *2020 International Conference on Computer Network, Electronic and Automation (ICCNEA)*, Xi'an, China: IEEE, Sep. 2020, pp. 154–158. doi: 10.1109/ICCNEA50255.2020.00040.
- [10] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, “DeepFace: Closing the Gap to Human-Level Performance in Face Verification,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA: IEEE, Jun. 2014, pp. 1701–1708. doi: 10.1109/CVPR.2014.220.
- [11] P. Heck and G. Schouten, “Turning Software Engineers into AI Engineers,” Jan. 04, 2021, *arXiv*: arXiv:2011.01590. Accessed: Oct. 20, 2024. [Online]. Available: <http://arxiv.org/abs/2011.01590>
- [12] M. Wang and W. Deng, “Deep face recognition: A survey,” *Neurocomputing*, vol. 429, pp. 215–244, Mar. 2021, doi: 10.1016/j.neucom.2020.10.081.
- [13] C. Janiesch, P. Zschech, and K. Heinrich, “Machine learning and deep learning,” *Electron Markets*, vol. 31, no. 3, pp. 685–695, Sep. 2021, doi: 10.1007/s12525-021-00475-2.
- [14] L. Alzubaidi *et al.*, “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” *J Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.
- [15] J. Howse and J. Minichino, *Learning OpenCV 4 computer vision with Python 3: get to grips with tools, techniques, and algorithms for computer vision and machine learning*, Third edition. Birmingham, UK: Packt Publishing, 2020.