

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

SISTEM VERIFIKASI WAJAH DEEP LEARNING MOBILENETV2 DAN CONVOLUTIONAL NEURAL NETWORK BERBASIS SIAMESE MODEL

ABD RAHMAN



**PROGRAM STUDI MAGISTER STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul “Sistem Verifikasi Wajah *Deep Learning* MobileNetV2 dan *Convolutional Neural Network* Berbasis Siamese Model” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2025

Abd Rahman
G1501231055

RINGKASAN

ABD RAHMAN. Sistem Verifikasi Wajah Deep Learning MobileNetV2 dan Convolutional Neural Network Berbasis Siamese Model. Dibimbing oleh AGUS MOHAMAD SOLEH dan ERFIANI.

Teknologi keamanan informasi berkembang pesat seiring meningkatnya kebutuhan perlindungan data dan identitas pengguna. Sistem autentikasi konvensional yang mengandalkan kata sandi atau PIN memiliki kelemahan mendasar karena rentan terhadap pencurian, peretasan, dan kelalaian pengguna. Teknologi biometrik menawarkan alternatif yang lebih aman dengan memanfaatkan ciri fisik atau perilaku unik seperti sidik jari, iris mata, dan wajah. Metode verifikasi wajah mendapatkan perhatian khusus karena sifatnya yang non-invasif, kemudahan penggunaan, dan potensi penerapan yang luas. Peningkatan penggunaan perangkat *mobile* dan *desktop* dalam kehidupan sehari-hari menimbulkan tantangan ketika verifikasi wajah harus dijalankan pada perangkat dengan sumber daya komputasi terbatas. Tantangan tersebut meliputi kebutuhan akurasi tinggi, efisiensi pemrosesan, dan stabilitas performa dalam kondisi nyata. Penelitian ini bertujuan mengevaluasi kinerja enam model *embedding* *Convolutional Neural Network* (CNN), *L2-normalized CNN*, *MobileNetV2*, *MobileNetV2 Adjusted*, *pre-trained MobileNetV2*, dan *fine-tuned MobileNetV2* yang diintegrasikan dalam arsitektur *Siamese Neural Network* (SNN). Penelitian ini juga bertujuan menilai pengaruh teknik optimasi seperti *pre-trained*, *fine-tuning*, dan *L2-normalization* terhadap performa model, serta menguji efektivitas implementasi model terbaik pada skenario *real-time*.

Penelitian ini merancang dan menguji arsitektur *Siamese Neural Network* (SNN) dengan enam model *embedding* yang memiliki karakteristik berbeda. Model CNN terdiri dari empat blok konvolusi berlapis *max pooling* dan satu lapisan *dense* berukuran 4.096 unit, menghasilkan total 178.421.056 parameter *trainable*. Model *L2-normalized CNN* menambahkan *batch normalization* dan *dropout* pada setiap blok konvolusi serta *L2 normalization layer* pada keluaran *dense* 512 unit, dengan total 23.361.600 parameter. Model *MobileNetV2* menggunakan arsitektur *lightweight* dengan tiga lapisan *dense* (4.096, 1.024, 512 unit) dan *L2 normalization*, berjumlah 12.230.208 parameter. Model *MobileNetV2 Adjusted* mengintegrasikan augmentasi langsung di arsitektur (*rescaling*, *random flip*, *random rotation*), menghasilkan *embedding* berdimensi 128 dengan total 3.870.144 parameter. Model *pre-trained MobileNetV2* memanfaatkan bobot *ImageNet* tanpa *fine-tuning* (9.969.664 parameter *trainable* dari total 12.230.208), sedangkan *fine-tuned MobileNetV2* mengaktifkan pelatihan ulang pada lapisan atas untuk menyesuaikan karakteristik data primer (11.831.104 parameter *trainable* dari total 12.230.208). Seluruh model *embedding* menghasilkan vektor representasi berdimensi rendah yang dibandingkan menggunakan metrik jarak *euclidean distance* dan dioptimalkan dengan *Contrastive Loss*. Proses pelatihan dilakukan selama 20 *epoch* menggunakan *Adam optimizer* dengan *early stopping*, sedangkan evaluasi performa mencakup metrik akurasi, presisi, *recall*, *F1-score*, AUC, serta pengujian *real-time* pada CPU dan GPU.

Penelitian ini memanfaatkan kombinasi data primer, sekunder, dan empiris untuk melatih dan menguji model. Data primer terdiri dari 2.520 citra wajah delapan



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan sumber :

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah

b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

IPB University

individu (empat pria dan empat wanita) yang diperoleh menggunakan tiga jenis *webcam*, yaitu Acer Nitro AN515-54, Asus ROG GC21, dan *Smart Monitor* Samsung M8, masing-masing menghasilkan 105 citra per individu. Data sekunder berasal dari *Labeled Faces in the Wild* (LFW) yang digunakan sebagai label negatif, sedangkan *dataset ImageNet* digunakan untuk mendukung pelatihan model *pre-trained*. Data empiris berupa *embedding* yang dihasilkan keenam model setelah proses ekstraksi fitur. Tahap *preprocessing* meliputi augmentasi citra, normalisasi skala piksel ke rentang $[0,1]$, serta penyesuaian ukuran menjadi 160×160 piksel. *Dataset* dibentuk menjadi pasangan positif-negatif yang seimbang, menghasilkan total 43.680 pasangan yang dibagi menjadi 70% untuk pelatihan (30.576 pasangan) dan 30% untuk pengujian (13.104 pasangan). Distribusi data yang seimbang ini memastikan proses pelatihan model bebas dari bias kelas dan *valid* untuk evaluasi performa.

Hasil evaluasi menunjukkan bahwa *pre-trained* MobileNetV2 dan CNN mencapai akurasi tertinggi masing-masing sebesar 100% dan 99,99%, dengan *precision*, *recall*, *F1-score*, dan AUC yang sangat tinggi. Model MobileNetV2 *Adjusted* menonjol karena mampu mempertahankan akurasi tinggi ($AUC > 0,99$) sekaligus menawarkan efisiensi komputasi, sehingga direkomendasikan sebagai model terbaik untuk aplikasi *real-time*. Model MobileNetV2 standar dan L2-normalized CNN menunjukkan performa lebih rendah dan cenderung tidak stabil, sedangkan *fine-tuned* MobileNetV2 memperlihatkan akurasi tinggi namun berisiko *overfitting* pada beberapa ulangan. Pengujian *real-time* pada CPU menunjukkan bahwa MobileNetV2 *Adjusted* menghasilkan *average distance* rendah pada wajah cocok, *latency* 152-202 ms, dan *throughput* 4,82-6,35 gambar per detik. Pengujian pada GPU juga memberikan hasil verifikasi yang baik, namun dengan *latency* lebih tinggi dan *throughput* lebih rendah dibanding CPU. Perbandingan CPU dan GPU menegaskan bahwa CPU lebih optimal untuk model ringan.

Penelitian ini menyimpulkan bahwa MobileNetV2 *Adjusted* merupakan model *embedding* terbaik untuk verifikasi wajah berbasis arsitektur *Siamese Neural Network* pada perangkat dengan sumber daya terbatas. Model MobileNetV2 *Adjusted* mampu memadukan akurasi tinggi dengan efisiensi komputasi yang mendukung implementasi *real-time*, khususnya pada CPU. Teknik *pre-trained* terbukti memberikan stabilitas hasil, sedangkan *fine-tuning* meningkatkan akurasi tetapi memerlukan mitigasi risiko *overfitting*. Evaluasi *real-time* memperlihatkan bahwa CPU memberikan kinerja lebih optimal dibanding GPU untuk model ringan pada *batch* kecil. Penelitian selanjutnya disarankan untuk mengoptimalkan pemrosesan GPU melalui *batch inference* serta membandingkan MobileNetV2 *Adjusted* dengan model *lightweight* lainnya seperti EfficientNet, ShuffleNet, atau MobileNetV3 guna memperoleh pemahaman lebih mendalam tentang trade-off antara performa dan efisiensi pada sistem verifikasi wajah.

Kata kunci: Convolutional Neural Network, MobileNetV2, real-time, Siamese Neural Network, Verifikasi wajah.

SUMMARY

ABD RAHMAN. Face Verification System Deep Learning MobileNetV2 and Convolutional Neural Network Based on Siamese Model. Supervised by AGUS MOHAMAD SOLEH and ERFIANI.

Information security technology has evolved rapidly along with the increasing demand for data and user identity protection. Conventional authentication systems that rely on passwords or PINs have fundamental weaknesses because they are vulnerable to theft, hacking, and user negligence. Biometric technology offers a more secure alternative by utilizing unique physical or behavioral characteristics such as fingerprints, iris patterns, and facial features. Facial verification methods have attracted particular attention due to their non-invasive nature, ease of use, and broad application potential. The growing use of mobile and desktop devices in daily life has created challenges when facial verification must be executed on devices with limited computational resources. These challenges include the need for high accuracy, efficient processing, and stable performance under real-world conditions. This study aims to evaluate the performance of six embedding models: Convolutional Neural Network (CNN), L2-normalized CNN, MobileNetV2, MobileNetV2 Adjusted, pre-trained MobileNetV2, and fine-tuned MobileNetV2, integrated into a Siamese Neural Network (SNN) architecture. This study also aims to assess the impact of optimization techniques such as pre-training, fine-tuning, and L2-normalization on model performance, as well as to test the effectiveness of the best-performing model in real-time scenarios.

This study designs and tests a Siamese Neural Network (SNN) architecture with six embedding models that have different characteristics. The CNN model consists of four convolutional blocks with max pooling layers and one dense layer of 4,096 units, resulting in a total of 178,421,056 trainable parameters. The L2-normalized CNN model adds batch normalization and dropout to each convolutional block and includes an L2 normalization layer on the 512-unit dense output, with a total of 23,361,600 parameters. The MobileNetV2 model employs a lightweight architecture with three dense layers (4,096, 1,024, and 512 units) and L2 normalization, totaling 12,230,208 parameters. The MobileNetV2 Adjusted model integrates direct augmentation in the architecture (rescaling, random flip, and random rotation) to produce 128-dimensional embeddings with a total of 3,870,144 parameters. The pre-trained MobileNetV2 model utilizes ImageNet weights without fine-tuning, having 9,969,664 trainable parameters from a total of 12,230,208 parameters. The fine-tuned MobileNetV2 model enables retraining on the upper layers to adapt to the characteristics of the primary data, with 11,831,104 trainable parameters from a total of 12,230,208 parameters. All embedding models generate low-dimensional representation vectors that are compared using the Euclidean distance metric and optimized with Contrastive Loss. The training process is conducted for 20 epochs using the Adam optimizer with early stopping, while performance evaluation includes accuracy, precision, recall, F1-score, AUC, and real-time testing on both CPU and GPU.

This study utilizes a combination of primary, secondary, and empirical data to train and test the models. The primary data consist of 2,520 facial images of eight

individuals (four males and four females) captured using three types of webcams: Acer Nitro AN515-54, Asus ROG GC21, and Samsung Smart Monitor M8, each producing 105 images per individual. The secondary data are obtained from the Labeled Faces in the Wild (LFW) dataset, which is used as negative labels, while the ImageNet dataset is employed to support the training of the pre-trained models. The empirical data comprise the embeddings produced by the six models after the feature extraction process. The preprocessing stage includes image augmentation, pixel scale normalization to the range $[0,1]$, and resizing to 160×160 pixels. The dataset is constructed into balanced positive-negative pairs, resulting in a total of 43,680 pairs, which are split into 70% for training (30,576 pairs) and 30% for testing (13,104 pairs). This balanced data distribution ensures that the model training process is free from class bias and valid for performance evaluation.

The evaluation results show that the pre-trained MobileNetV2 and CNN models achieved the highest accuracy of 100% and 99.99%, respectively, with very high precision, recall, F1-score, and AUC values. The MobileNetV2 Adjusted model stands out for its ability to maintain high accuracy ($AUC > 0.99$) while offering computational efficiency, making it the recommended model for real-time applications. The standard MobileNetV2 and L2-normalized CNN models demonstrated lower and less stable performance, while the fine-tuned MobileNetV2 model exhibited high accuracy but showed a risk of overfitting in several repetitions. The real-time testing on a CPU showed that the MobileNetV2 Adjusted model produced a low average distance for matching faces, latency ranging from 152 to 202 ms, and throughput of 4.82 to 6.35 images per second. The testing on a GPU also yielded good verification results but with higher latency and lower throughput compared to the CPU. The comparison between CPU and GPU confirms that the CPU is more optimal for lightweight models.

This study concludes that the MobileNetV2 Adjusted model is the best embedding model for face verification based on the Siamese Neural Network architecture on devices with limited resources. The MobileNetV2 Adjusted model combines high accuracy with computational efficiency that supports real-time implementation, particularly on CPUs. The pre-trained technique has been shown to provide result stability, while fine-tuning increases accuracy but requires mitigation of overfitting risks. The real-time evaluation shows that the CPU delivers more optimal performance than the GPU for lightweight models with small batch sizes. Future research is recommended to optimize GPU processing through batch inference and to compare the MobileNetV2 Adjusted model with other lightweight models such as EfficientNet, ShuffleNet, or MobileNetV3 to gain deeper insights into the trade-off between performance and efficiency in face verification systems.

Keywords: Convolutional Neural Network, face verification, MobileNetV2, real-time, Siamese Neural Network.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

SISTEM VERIFIKASI WAJAH DEEP LEARNING MOBILENETV2 DAN CONVOLUTIONAL NEURAL NETWORK BERBASIS SIAMESE MODEL

ABD RAHMAN

Tesis
sebagai salah satu syarat untuk memperoleh gelar
Magister pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI MAGISTER STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Penguji Luar Komisi Pembimbing pada Ujian Tesis:
Dr Bagus Sartono, S.Si., M.Si.



Judul Tesis : Sistem Verifikasi Wajah *Deep Learning* MobileNetV2 dan
Convolutional Neural Network Berbasis *Siamese* Model

Nama : Abd Rahman

NIM : G1501231055

@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Disetujui oleh

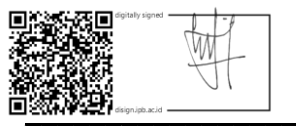
Pembimbing 1:

Dr. Agus Mohamad Soleh S.Si., M.T



Pembimbing 2:

Dr. Ir. Erfiani M.Si.



Diketahui oleh

Ketua Program Studi:

Dr. Agus Mohamad Soleh, S.Si., M.T.

NIP 19750315 199903 1 004



Dekan Sekolah Sains Data, Matematika dan Informatika :

Prof. Dr. Ir. Agus Buono, M.Si., N.Kom.

NIP 19660702 199302 1 001



Tanggal Ujian:
08 Agustus 2025

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Agustus 2024 sampai bulan Mei 2025 dengan judul “Sisitem Verifikasi Identitas Wajah *Deep Learning* MobileNetV2 dan *Convolutional Neural Network* Berbasis *Siamese Model*”.

Terima kasih penulis ucapkan kepada para pembimbing, Dr. Agus Mohamad Soleh S.Si., M.T dan Dr. Ir. Erfiani M.Si. atas bimbingan dan nasihatnya sejak awal penelitian ini. Ucapan terima kasih juga penulis sampaikan kepada pembimbing akademik (untuk program S-2), moderator seminar, dan penguji luar komisi pembimbing. Di samping itu, penghargaan penulis sampaikan kepada seluruh dosen dan administrasi program studi statistika dan sains data IPB University yang telah banyak memberikan ilmu, kenyamanan dan kemudahan dalam proses pembelajaran selama menempuh jenjang magister.

Penulis mengucapkan terimakasih kepada Bapak H Muhammad Yusuf dan Ibu Hj Hasnawati yang telah berjuang untuk membiayai anak-anaknya hingga pendidikan magister, dan seluruh anggota keluarga atas dukungan, doa, dan cinta yang tidak terhingga. Penulis juga mengucapkan terima kasih kepada Tim Outlier yang telah banyak membantu selama penelitian, kepada *volunteer* Meisya, Windy, Uswatun, Rifda, Yudith, Kevin, Reza, Caca, Rahmi, Dwi, Risha, Monica, Wawan, Adib yang telah membantu selama pengumpulan data primer, teman-teman Statistika dan Sains Data Angkatan 2023 dan teman-teman di LSI Pojok Statistik yang selalu mendukung dan memberi semangat selama pengerjaan Tesis.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Juli 2025

Abd Rahman

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



DAFTAR ISI

DAFTAR TABEL	xi
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xi
DAFTAR ISTILAH	xi
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Manfaat	3
1.5 Ruang Lingkup	3
1.6 Hipotesis	4
II TINJAUAN PUSTAKA	5
2.1 Identifikasi dan verifikasi	5
2.2 Citra Digital	5
2.3 Convolutinal Neural Network	6
2.4 MobileNetV2	8
2.5 Siamese Neural Network	8
2.6 Evaluasi Model	9
III METODE	11
3.1 Model	11
3.2 Data	11
3.3 Prosedur Analisis Data	12
3.4 Analisis data	16
IV HASIL DAN PEMBAHASAN	20
4.1 Ringkasan Dataset Setelah Preprocessing	20
4.2 Pemodelan Berbasis Siamese	21
4.3 Evaluasi Overfitting Model	22
4.4 Evaluasi Performa Model	26
4.5 Analisis Perbandingan Model Embedding Siamese Model	27
4.6 Hasil Evaluasi Real Time	32
4.7 Pengujian Konsistensi Kinerja Model pada Skenario Real-Time	35
V SIMPULAN DAN SARAN	38
5.1 Simpulan	38
5.2 Saran	38
DAFTAR PUSTAKA	39
LAMPIRAN	42
RIWAYAT HIDUP	60

DAFTAR TABEL

1	Struktur Data Gambar	20
2	Distribusi Pasangan Gambar	21
3	Ringkasan Nilai <i>Loss</i> dan Akurasi Model <i>Siamese</i>	22
4	Standar Deviasi Performa Model <i>Embedding</i>	26
5	Nilai Evaluasi Rata-Rata Setiap Model <i>Embedding</i>	27
6	Kompleksitas Arsitektur <i>Embedding Siamese Neural Network</i>	28
7	Hasil Uji MANOVA	29
8	Hasil Uji ANOVA pada Masing-Masing Metrik Kinerja	29
9	Hasil Uji <i>Post-hoc Tukey HSD</i> pada Masing-Masing Metrik Kinerja	30
10	Hasil Evaluasi <i>Real Time</i> CPU	33
11	Hasil Evaluasi <i>Real Time</i> GPU	34
12	Uji Konsistensi Model Skenario <i>Real-Time</i>	36
13	Uji Konsistensi Model <i>Real-Time</i> pada Variasi Latar Belakang	37

DAFTAR GAMBAR

1	Tampilan sederhana <i>face recognition</i>	5
2	Ilustrasi citra biner	6
3	Arsitektur <i>convolutional neural network</i>	7
4	Evolusi <i>separable convolutional block</i>	8
5	<i>Classification network</i>	9
6	Evaluasi performa model <i>confusion matrix</i> dan ROC	10
7	<i>Flowchart</i> Pembentukan <i>Dataset</i>	18
8	<i>Flowchart</i> Proses Analisis	19
9	<i>Loss</i> dan <i>accuracy curve</i> (a) CNN, (b) L2 Normalization CNN	23
10	<i>Loss</i> dan <i>accuracy curve</i> (a) MobileNetV2, (b) MobileNetV2 <i>Adjust</i>	24
11	<i>Loss</i> dan <i>accuracy curve</i> (a) <i>pre-trained</i> , (b) <i>fine-tuned</i> MobileNetV2	25
12	Distribusi Kinerja per model (a) <i>Accuracy</i> , (b) <i>Precision</i> , (c) <i>Recall</i> , (d) <i>F1-Score</i> , (e) <i>Loss</i>	31
13	(a) Skor Efisiensi Model, (b) Radar Efisiensi Model	31
14	GPU vs CPU (a) <i>box-plot latency</i> , (b) <i>Average latency</i> , (c) <i>Average Throughput</i>	35

DAFTAR LAMPIRAN

1	Visualisasi <i>Embedding</i> Model	43
2	Integrasi <i>Siamese</i> Model	49
3	Data Primer (Citra Wajah)	48
4	Data sekunder (LWF dan <i>ImageNet</i>)	49
5	Data Empiris <i>Embedding</i>	50

6	Syntax Optimizer dan Loss	54
7	Uji Statistik	55
8	Hasil Rata-Rata Metrik Kinerja Model per Epoch 5 Kali Looping	57
9	Evaluasi Efisiensi Model	58
10	Threshold Optimal	59

DAFTAR ISTILAH

- 1 Lapisan Konvolusi: Lapisan dalam jaringan saraf tiruan yang berfungsi mengekstraksi fitur dari citra dengan mengaplikasikan operasi konvolusi menggunakan filter atau kernel.
- 2 Lapisan Pooling: Lapisan dalam jaringan saraf tiruan yang berfungsi mengurangi dimensi fitur dengan mempertahankan informasi penting, biasanya menggunakan metode *max pooling* atau *average pooling*.
- 3 Lapisan Fully Connected: Lapisan dalam jaringan saraf tiruan di mana setiap neuron terhubung dengan semua neuron pada lapisan sebelumnya, berfungsi menggabungkan fitur untuk menghasilkan prediksi akhir.
- 4 Thermal Throttling: Mekanisme penurunan kinerja perangkat keras secara otomatis untuk mencegah panas berlebih dengan mengurangi kecepatan prosesor atau komponen terkait.
- 5 Inverted Residuals: Blok arsitektur jaringan saraf tiruan yang menggunakan lapisan *bottleneck* sebagai input dan output, dengan lapisan ekspansi di tengah, untuk meningkatkan efisiensi komputasi dan akurasi.
- 6 Linear Bottleneck: Lapisan *bottleneck* dalam arsitektur jaringan saraf tiruan yang menggunakan aktivasi linear pada output untuk mempertahankan informasi penting dan mengurangi hilangnya representasi fitur.
- 7 Embedding: Representasi data dalam bentuk vektor berdimensi rendah yang memuat informasi penting dari data asli, digunakan untuk mempermudah pemrosesan dan analisis oleh model pembelajaran mesin.
- 8 Weight Sharing: Teknik dalam jaringan saraf tiruan di mana beberapa koneksi atau lapisan menggunakan bobot yang sama untuk mengurangi jumlah parameter dan meningkatkan efisiensi komputasi.
- 9 Scale Invariant Feature Transform (SIFT): Algoritma deteksi dan deskripsi fitur pada citra yang mampu mengenali objek secara konsisten meskipun terjadi perubahan skala, rotasi, atau pencahayaan.
- 10 Feature Map: Representasi dua dimensi dari hasil ekstraksi fitur pada lapisan konvolusi, yang memuat respon aktivasi filter terhadap bagian-bagian tertentu dari input.
- 11 Konvolusi Depthwise: Operasi konvolusi yang menerapkan satu filter untuk setiap saluran input secara terpisah, digunakan untuk mengurangi jumlah parameter dan mempercepat pemrosesan dalam jaringan saraf tiruan.
- 12 Precision: Metrik evaluasi yang mengukur proporsi prediksi positif yang benar dibandingkan dengan seluruh prediksi positif yang dihasilkan model.
- 13 Recall: Metrik evaluasi yang mengukur proporsi prediksi positif yang benar dibandingkan dengan seluruh data positif yang sebenarnya ada dalam dataset.

- 14 F1-score: Metrik evaluasi yang merupakan rata-rata harmonis dari *precision* dan *recall*, digunakan untuk menilai keseimbangan antara keduanya dalam kinerja model.
- 15 Image Sequence: Kumpulan citra yang diambil secara berurutan dalam rentang waktu tertentu, digunakan untuk merepresentasikan pergerakan atau perubahan objek dalam analisis visual.
- 16 Open Source Computer: Perangkat keras atau perangkat lunak komputer yang desain, kode sumber, atau spesifikasinya tersedia secara bebas untuk digunakan, dimodifikasi, dan didistribusikan oleh publik.
- 17 CUDA Toolkit: Paket pengembangan dari NVIDIA yang menyediakan pustaka, *compiler*, dan alat pemrograman untuk memanfaatkan kemampuan pemrosesan paralel GPU dalam mempercepat komputasi.
- 18 cuDNN: Pustaka percepatan komputasi dari NVIDIA yang dioptimalkan untuk operasi *deep learning* pada GPU, menyediakan implementasi efisien untuk operasi seperti konvolusi, normalisasi, dan aktivasi.
- 19 Perspective Transform: Proses mengubah sudut pandang citra sehingga tampak diambil dari perspektif berbeda, digunakan untuk meningkatkan variasi data dalam pelatihan model.
- 20 Occlusion: Teknik augmentasi yang menutupi sebagian area objek dalam citra untuk mensimulasikan kondisi objek terhalang sebagian, sehingga model lebih robust terhadap gangguan visua
- 21 Sharpening: Proses meningkatkan ketajaman tepi dan detail pada citra dengan mempertegas kontras lokal, digunakan untuk memperjelas fitur penting.
- 22 Batch: Sekumpulan data yang diproses secara bersamaan dalam satu iterasi pelatihan model pembelajaran mesin untuk mempercepat komputasi dan mengoptimalkan pemakaian sumber daya.
- 23 Pipeline: Rangkaian tahapan pemrosesan data atau eksekusi model yang dilakukan secara berurutan dan terstruktur untuk menyelesaikan suatu tugas secara efisien.
- 24 Depthwise Separable Convolutions: Jenis operasi konvolusi yang memisahkan proses konvolusi menjadi dua tahap, yaitu *depthwise convolution* untuk memproses setiap saluran input secara terpisah, dan *pointwise convolution* untuk menggabungkan hasilnya, sehingga mengurangi jumlah parameter dan mempercepat komputasi.
- 25 Loss: Nilai yang mengukur seberapa besar perbedaan antara prediksi model dan nilai sebenarnya, digunakan sebagai acuan untuk memperbarui bobot selama pelatihan.
- 26 Optimizer: Algoritma dalam pembelajaran mesin yang digunakan untuk menyesuaikan bobot model berdasarkan nilai *loss* agar kinerja model meningkat selama proses pelatihan.
- 27 Learning Rate: Parameter yang menentukan besarnya langkah pembaruan bobot model pada setiap iterasi pelatihan, yang memengaruhi kecepatan dan stabilitas proses pembelajaran.
- 28 GradientTape: Fasilitas dalam TensorFlow yang merekam operasi selama eksekusi untuk menghitung gradien secara otomatis, biasanya digunakan dalam proses pelatihan model pembelajaran mesin.

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Forward Pass: Proses propagasi data input melalui lapisan-lapisan model untuk menghasilkan output atau prediksi sebelum dilakukan perhitungan *loss* dan pembaruan bobot.

Checkpoint: Fitur penyimpanan kondisi model, termasuk bobot dan konfigurasi, pada titik tertentu selama pelatihan sehingga proses dapat dilanjutkan atau dipulihkan tanpa mengulang dari awal.

Gradien Loss: Turunan pertama dari fungsi *loss* terhadap bobot model, yang menunjukkan arah dan besarnya perubahan bobot yang diperlukan untuk meminimalkan *loss* selama pelatihan.

@tf.function: Dekorator pada TensorFlow yang mengubah fungsi Python menjadi grafik TensorFlow (*computational graph*) untuk meningkatkan efisiensi dan kecepatan eksekusi.

train_step: Fungsi yang mendefinisikan satu langkah pelatihan model, mencakup proses *forward pass*, perhitungan *loss*, perhitungan gradien, dan pembaruan bobot menggunakan *optimizer*.

Early Stop: Teknik penghentian pelatihan model secara otomatis ketika kinerja pada data validasi tidak menunjukkan perbaikan selama sejumlah iterasi tertentu, untuk mencegah *overfitting*.

Trainable Parameters: Parameter dalam model pembelajaran mesin yang nilainya diperbarui selama proses pelatihan melalui optimisasi untuk meminimalkan *loss*.

Pillai's Trace: Statistik uji dalam analisis multivariat yang mengukur proporsi variansi yang dijelaskan oleh model, dikenal cukup robust terhadap pelanggaran asumsi.

Wilks' Lambda: Statistik uji dalam analisis multivariat yang membandingkan variansi dalam kelompok dan antar kelompok, dengan nilai lebih kecil menunjukkan perbedaan yang lebih besar antar kelompok.

Hotelling–Lawley Trace: Statistik uji dalam analisis multivariat yang mengukur rasio variansi antar kelompok terhadap variansi dalam kelompok, sensitif terhadap perbedaan besar pada mean antar kelompok.

Roy's Greatest Root: Statistik uji dalam analisis multivariat yang menggunakan nilai eigen terbesar untuk mengukur perbedaan antar kelompok, paling sensitif ketika perbedaan terfokus pada satu dimensi utama.

Num DF (Numerator Degrees of Freedom): Derajat kebebasan yang terkait dengan variansi antar kelompok atau efek yang diuji dalam analisis statistik.

Den DF (Denominator Degrees of Freedom): Derajat kebebasan yang terkait dengan variansi dalam kelompok atau galat (error) dalam analisis statistik

p-adj (Adjusted p-value): Nilai *p* yang telah disesuaikan untuk mengoreksi pengujian ganda, digunakan untuk mengurangi kemungkinan kesalahan tipe I.

Mean Diff (Mean Difference): Selisih rata-rata antara dua kelompok yang dibandingkan dalam analisis statistik.

Latency: Waktu yang diperlukan oleh sistem atau model untuk memproses input hingga menghasilkan output, biasanya diukur dalam milidetik pada aplikasi *real-time*.

Throughput: Jumlah data atau tugas yang dapat diproses oleh sistem atau model dalam satuan waktu tertentu, digunakan untuk mengukur kapasitas pemrosesan.