

Deteksi Suhu Melalui Citra Termal Wajah Menggunakan *Deep Learning*

Dion Devara Aryasatya, Nanik Suciati, dan Ary Mazharuddin Shiddiqi
Departemen Teknik Informatika, Institut Teknologi Sepuluh Nopember (ITS)
e-mail: nanik@if.its.ac.id

Abstrak—Dalam masa pandemi kasus penularan virus CORONA masih tetap bertambah dari hari ke hari. Salah satu gejala yang umum dialami oleh pasien COVID-19 adalah demam. Hal yang umum dilakukan untuk mengukur suhu di masa pandemi adalah menggunakan termometer non kontak. *Deep Learning* bisa digunakan untuk mendeteksi wajah dan membantu mendeteksi suhu maksimal wajah dari gambar termal. Tujuan Penelitian ini adalah membuat aplikasi pendeteksi suhu pada citra termal menggunakan pendekatan *Deep Learning*. Dalam Penelitian ini dilatih sebuah model deteksi SSD-MobileNet untuk mendeteksi area wajah dari citra termal. Setelah terdeteksi, data suhu diekstrak dari area wajah tersebut. Dalam pelaksanaan penelitian ini digunakan *dataset* citra termal Tuft Face Database, IRDatabase, dan citra termal yang diambil menggunakan Flir One. Dari hasil uji coba didapatkan hasil *mean average precision* deteksi wajah sebesar 0,95 dengan *threshold* dari evaluasi model untuk IoU 0,75 sebesar 0,95 dan *mean absolute error* deteksi suhu sebesar 1,51.

Kata Kunci—Deteksi Suhu, *Deep Learning*, SSD-MobileNet, Citra Termal.

I. PENDAHULUAN

DALAM masa pandemi kasus penularan virus CORONA masih tetap bertambah dari hari ke hari. Diperlukan usaha yang maksimal agar penyebaran virus CORONA bisa ditekan. Orang yang sehat harus dipisahkan dengan orang yang sakit dan orang yang mengalami gejala. Hal tersebut perlu dilakukan agar virus tidak semakin menyebar. Salah satu gejala yang umum dialami oleh pasien COVID-19 adalah demam [1]. Hal yang umum dilakukan untuk mengukur suhu di masa pandemi adalah menggunakan termometer non kontak. *Deep Learning* bisa digunakan untuk mendeteksi wajah dan suhu maksimal wajah dari gambar termal. Dari data yang didapatkan akan digunakan untuk menentukan apakah subjek terdeteksi demam atau tidak.

Deep Learning merupakan sebuah bentuk pembelajaran mesin yang memungkinkan sebuah sistem komputer untuk mengumpulkan berbagai pengetahuan tanpa perlu adanya operator manusia untuk menentukan secara rinci semua pengetahuan yang dibutuhkan oleh sistem tersebut. Hirarki konsep memungkinkan komputer untuk mempelajari konsep-konsep yang rumit dengan membangun konsep-konsep tersebut dari konsep yang lebih sederhana. Jika konsep-konsep yang dibangun digambarkan dalam bentuk grafik, gambaran grafiknya dalam terdiri dari banyak lapisan [2].

Penelitian ini adalah bagian dari ANDROID-MTID yang merupakan aplikasi Android yang menggabungkan sistem deteksi masker, sistem deteksi suhu dan sistem pengenalan wajah bermasker maupun tak bermasker. Sistem deteksi suhu akan diimplementasikan dengan memanfaatkan data yang didapatkan dari kamera termal. *Deep Learning* akan digunakan untuk mendeteksi wajah dari data. Setelah wajah terdeteksi maka sistem akan mengekstrak suhu wajah subjek.



Gambar 1. Kamera termal.

Jika nilai suhu wajah subjek melebihi batas maka sistem akan memberikan peringatan.

II. TINJAUAN PUSTAKA

A. *Hardware Tools*

Kamera termal adalah sebuah perangkat non-kontak yang mendeteksi energi termal (panas) dan mengubahnya menjadi gambar visual. Resolusi yang dihasilkan oleh kamera termal lebih kecil jika dibandingkan dengan kamera RGB karena detektor termal perlu mendeteksi energi yang memiliki panjang gelombang yang jauh lebih besar dari cahaya tampak, sehingga setiap elemen sensor harus jauh lebih besar. Panas yang dideteksi oleh kamera termal bisa diukur dengan sangat tepat sehingga memungkinkan untuk diaplikasikan untuk berbagai macam hal seperti operasi militer, mobil otonom, dan penelitian ilmiah. Foto dari kamera termal yang digunakan bisa dilihat pada Gambar 1.

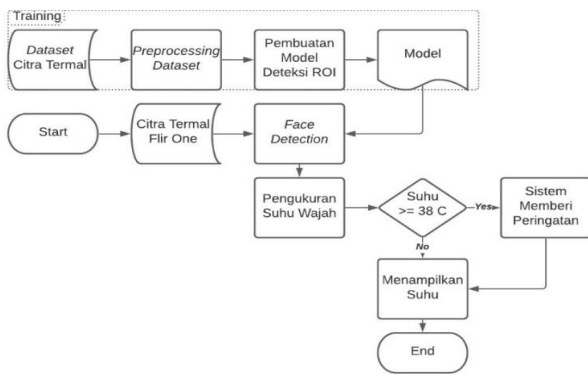
B. *Software Tools*

1) *Google Colab*

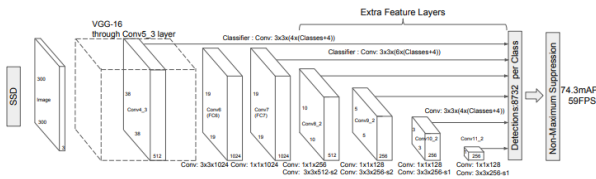
Google Colab adalah sebuah produk dari Google Research yang menyediakan layanan Jupyter Notebook yang dihosting sehingga tidak perlu melakukan instalasi software secara lokal. Colab memungkinkan pengguna untuk menulis dan mengeksekusi kode python arbitrer melalui browser, dan sangat cocok untuk pembelajaran mesin, analisis data, dan pendidikan. Google memberikan akses gratis ke sumber daya komputasi termasuk GPU dan TPU tetapi dengan batasan tertentu.

2) *Tensorflow*

Tensorflow adalah platform *open source* end-to-end untuk *machine learning*. Platform ini memiliki ekosistem alat, pustaka, dan sumber daya komunitas yang komprehensif dan fleksibel yang memungkinkan peneliti untuk mendorong ML mutakhir dan pengembang bisa dengan mudah membangun



Gambar 2. Alur kerja sistem.



Gambar 3. Arsitektur SSD.



Gambar 4. Tuft face dataset TD_IR_A.

dan menerapkan aplikasi yang didukung oleh ML. Untuk pengerjaan penelitian ini digunakan Object Detection API dan Detection Model Zoo. Object Detection API sebuah framework open source yang dibangun di atas TensorFlow yang memudahkan pembuatan, pelatihan, dan penerapan model deteksi objek. Detection Model Zoo merupakan database model deteksi objek yang telah dilatih sebelumnya menggunakan dataset seperti COCO, Kitti, Open Images, AVA v2.1, iNaturalist, dan Snapshot Serengeti. Model-model ini dapat digunakan untuk inferensi *out-of-the-box* atau untuk melakukan pelatihan ulang menggunakan dataset yang lain.

3) LabelImg

LabelImg adalah alat anotasi gambar grafis yang dibuat menggunakan bahasa pemrograman Python dan Qt untuk antarmuka grafisnya. Anotasi disimpan sebagai file XML dalam format PASCAL VOC, format yang digunakan oleh ImageNet. Selain itu, LabelImg juga mendukung format YOLO dan CreateML.

4) Flir Image Extractor

Flir Image Extractor adalah sebuah library Python yang memungkinkan pengguna untuk mengekstrak citra asli dan nilai sensor termal yang dikonversi ke suhu dari kamera

Tabel 1. Arsitektur mobilenet

Type	Stride	Bentuk Kernel	Ukuran Input
Conv	2	3 × 3 × 3 × 32	224 × 224 × 3
Conv dw	1	3 × 3 × 32	112 × 112 × 32
Conv	1	1 × 1 × 32 × 64	112 × 112 × 32
Conv dw	2	3 × 3 × 64	112 × 112 × 64
Conv	1	1 × 1 × 64 × 128	56 × 56 × 64
Conv dw	1	3 × 3 × 128	56 × 56 × 128
Conv	1	1 × 1 × 128 × 128	56 × 56 × 128
Conv dw	2	3 × 3 × 128	56 × 56 × 128
Conv	1	1 × 1 × 128 × 256	56 × 56 × 128
Conv dw	1	3 × 3 × 256	28 × 28 × 256
Conv	1	1 × 1 × 256 × 256	28 × 28 × 256
Conv dw	2	3 × 3 × 256	28 × 28 × 256
Conv	1	1 × 1 × 256 × 512	14 × 14 × 256
Conv dw	1	3 × 3 × 512	14 × 14 × 512
Conv	1	1 × 1 × 512 × 512	14 × 14 × 512
Conv dw	1	3 × 3 × 512	14 × 14 × 512
Conv	1	1 × 1 × 512 × 512	14 × 14 × 512
Conv dw	1	3 × 3 × 512	14 × 14 × 512
Conv	1	1 × 1 × 512 × 512	14 × 14 × 512
Conv dw	1	3 × 3 × 512	14 × 14 × 512
Conv	1	1 × 1 × 512 × 512	14 × 14 × 512
Conv dw	2	3 × 3 × 512	14 × 14 × 512
Conv	1	1 × 1 × 512 × 1024	7 × 7 × 512
Conv dw	2	3 × 3 × 1024	7 × 7 × 1024
Conv	1	1 × 1 × 1024 × 1024	7 × 7 × 1024
Avg Pool	1	Pool 7 × 7	7 × 7 × 1024
Fully Connected	1	1024 × 1000	1 × 1 × 1024
Softmax	1	Classifier	1 × 1 × 1000

termal buatan Flir. Jenis kamera yang didukung adalah Flir One, Xenmuse XTR, dan AX8.

III. PERANCANGAN

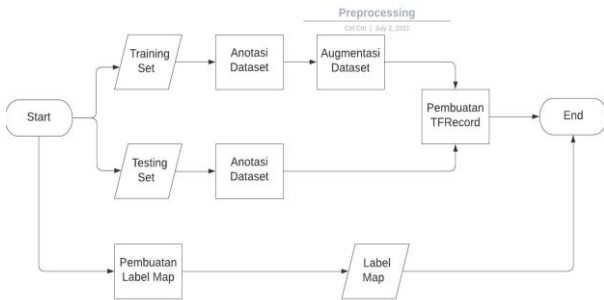
A. Perancangan Arsitektur Sistem

Dalam penelitian ini dibuat sebuah sistem yang bisa mendeteksi suhu wajah pada manusia. Desain sistem ini memakai desain sistem yang dimodifikasi dari referensi [3]. Input yang digunakan sistem berupa citra termal yang diperoleh dari pengambilan gambar menggunakan kamera termal Flir One. Deteksi wajah pada citra input dilakukan menggunakan model SSD-MobileNet. Jika wajah berhasil terdeteksi akan didapatkan koordinat *bounding box* wajah.

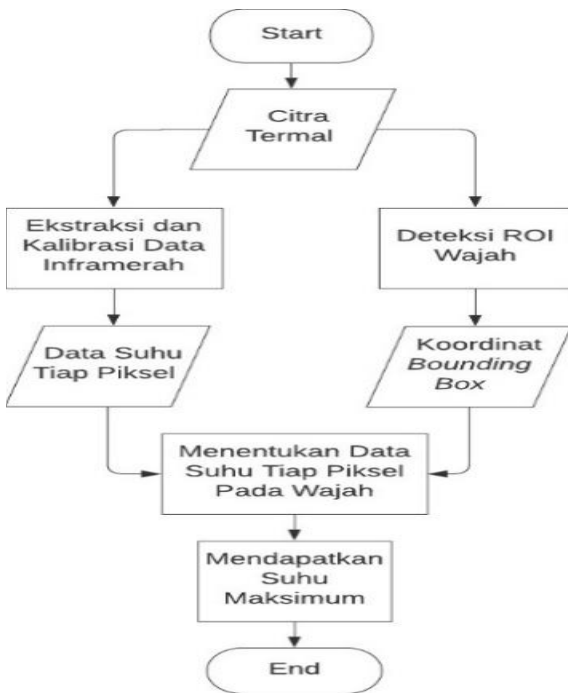
Setelah wajah terdeteksi, data suhu maksimal wajah akan diekstrak. Dari data suhu yang berhasil di ekstrak bisa ditentukan apakah seseorang mengalami demam atau tidak oleh sistem menggunakan Flir Image Extractor. Jika suhu badan lebih dari atau sama dengan 38° Celcius maka sistem akan memberikan peringatan [4]. Setelah itu sistem akan menampilkan suhu wajah yang berhasil diekstrak. Alur kerja sistem dapat dilihat pada Gambar 2.

1) Single Shot Multibox Detector (SSD)

SSD adalah sebuah metode deteksi objek pada gambar menggunakan satu deep neural network. SSD memisahkan ruang keluaran dari bounding box menjadi satu set box default melalui rasio aspek dan skala yang berbeda per lokasi peta fitur. Pada waktu prediksi, network menghasilkan skor untuk keberadaan setiap kategori objek di setiap box default dan menghasilkan penyesuaian pada box agar lebih cocok



Gambar 5. Diagram alir preprocessing.



Gambar 6. Diagram alir pengukuran suhu.

dengan bentuk objek. Selain itu, network menggabungkan prediksi dari beberapa peta fitur dengan resolusi berbeda untuk menangani objek dengan berbagai ukuran secara alami. SSD mudah untuk dilatih dan langsung diintegrasikan ke dalam sistem yang memerlukan komponen deteksi. Gambaran arsitektur SSD bisa dilihat pada Gambar 3.

2) MobileNet

Arsitektur MobileNet adalah sebuah model yang didesain untuk aplikasi perangkat bergerak dan embedded vision. MobileNet didasarkan pada arsitektur ramping yang menggunakan depthwise separable convolutions untuk membangun deep neural network yang ringan. Diperkenalkan dua hyperparameter global sederhana yang secara efisien menukar antara latensi dan akurasi. Hyperparameter ini memungkinkan pembuat model untuk memilih model yang berukuran tepat sesuai dengan batasan masalah. MobileNet bisa digunakan untuk berbagai aplikasi dan kasus penggunaan seperti deteksi objek, klasifikasi fine grain, atribut wajah, dan geo-lokalisasi skala besar [5]. Definisi arsitektur MobileNet bisa dilihat pada Tabel 1.

B. Preprocessing

Dataset yang digunakan untuk melatih model deteksi adalah Tuft Face Dataset TD_IR_A. Sample dari dataset bisa dilihat pada Gambar 4.

Tabel 2. Hasil evaluasi model

Model	Augmentasi	num_steps	decay_steps	mAP		Total Loss
				IoU	IoU	
1	tidak	77396	800720	0,99	0,86	6,02
2	tidak	50000	5000	0,98	0,77	6,20
3	ya	20000	5000	1,00	0,95	3,82

Tabel 3. Hasil uji coba pengukuran suhu

Citra	Subjek	Ground Truth	Suhu Terdeteksi	MAE
1	1	36,50	36,10	
2	1	36,80	34,18	
3	1	36,70	35,55	1,74
4	1	36,60	36,50	
5	1	36,70	34,01	
6	2	36,80	35,97	
7	2	36,80	37,60	
8	2	36,80	37,70	0,99
9	2	36,90	37,54	
10	2	36,60	37,40	
11	3	36,30	36,45	
12	3	35,80	37,07	
13	3	35,70	37,01	1,25
14	3	35,80	37,01	
15	3	35,90	36,94	
16	4	35,60	34,62	
17	4	36,40	34,90	
18	4	35,90	37,94	2,06
19	4	35,70	37,95	
20	4	36,20	37,67	
MAE Rata-rata				1,51

Citra termal yang digunakan memiliki resolusi 80x60. Citra termal diambil dengan menggunakan kamera FLIR Vue Pro mengitari subjek dari 9 posisi berbeda. Total citra adalah 998 gambar. Citra yang disediakan oleh dataset ini belum dilakukan anotasi. Dataset dibagi menjadi menjadi training set dan testing set. Training set berjumlah 890, sedangkan testing set berjumlah 108 gambar [6]. Sebagai gambaran umum Preprocessing Dataset digunakan diagram alir. Diagram alir data dapat dilihat melalui Gambar 5. Langkah-langkah preprocessing adalah sebagai berikut:

1) Anotasi Dataset

Sebelum bisa digunakan sebagai input, perlu dilakukan anotasi atau pelabelan pada dataset agar mesin bisa memahami pola pada citra dengan mudah.

2) Augmentasi Data Training

Deep Learning membutuhkan data yang lebih banyak dibandingkan dengan tipe Machine Learning yang lain. Dengan augmentasi data, data baru secara artifisial bisa dibuat dari data pelatihan yang sudah ada. Akan diimplementasikan horizontal flip, image scaling, dan image rotation pada set training. Setelah augmentasi training set berjumlah 1780 citra dan testing set berjumlah 108 citra.

3) Pembuatan Label Map

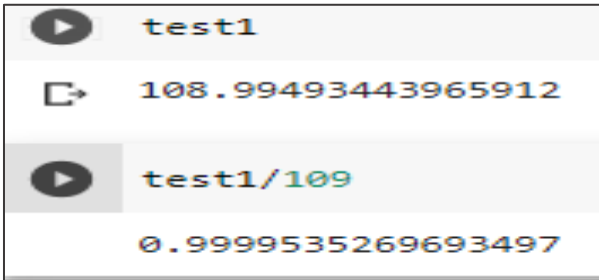
TensorFlow membutuhkan sebuah file label untuk memetakan setiap label yang digunakan ke nilai integer. Label map ini digunakan dalam proses pelatihan model maupun deteksi.

4) Pembuatan File Record Tensorflow

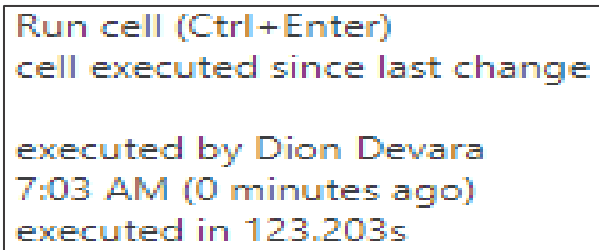
Citra beranotasi yang sudah dibuat dikonversi menjadi format file Tensorflow Records (TFRecord). TFRecord



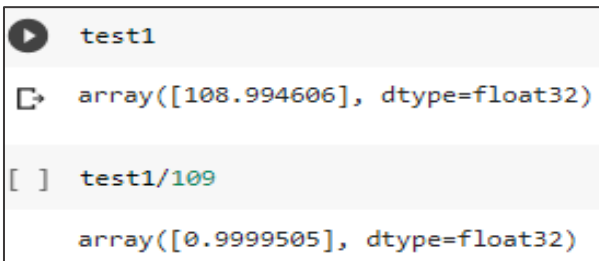
Gambar 7. Visualisasi deteksi model.



Gambar 8. Nilai confidence score model tensorflow.



Gambar 9. Nilai execution time model tensorflow.

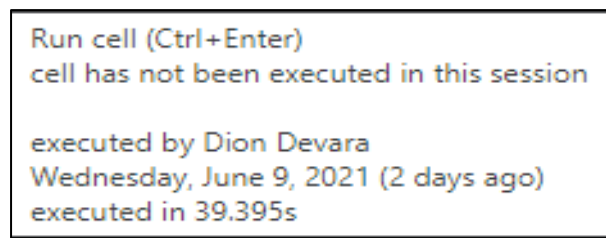


Gambar 10. Nilai confidence score model tfLite.

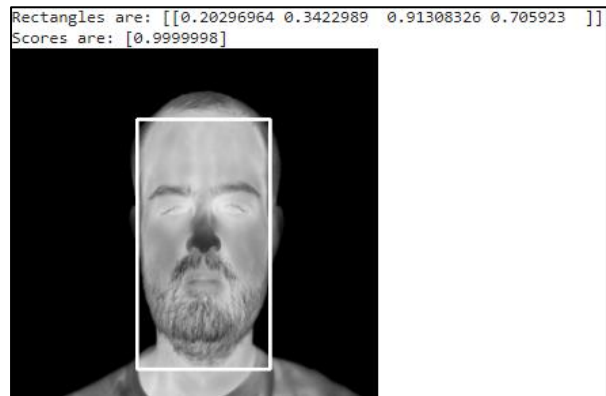
merupakan sebuah format file untuk menumpukan data yang disediakan oleh Tensorflow.

C. Pembuatan Model Deteksi ROI

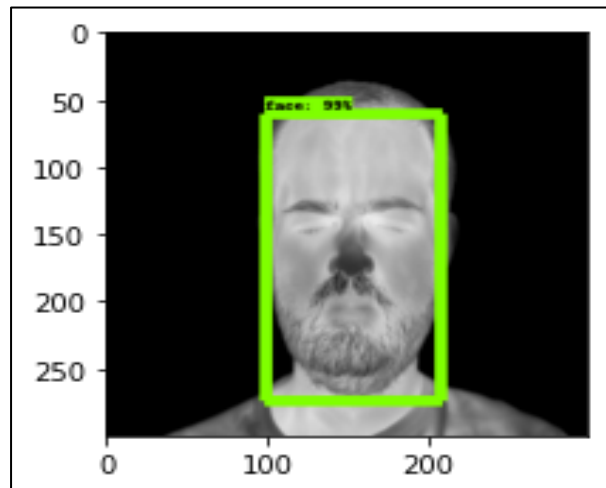
Pada penelitian ini, digunakan deteksi pada bagian wajah agar proses deteksi suhu bisa dilakukan lebih cepat dan mudah tanpa adanya kontak langsung. Model deteksi ROI wajah dibuat dengan melatih model SSD-MobileNet pra-terlatih dari Model Zoo. Digunakan model `ssd_mobilenet_v1_coco` yang dilatih menggunakan dataset COCO. Dataset COCO terdiri dari 330.000 gambar dengan 80 kategori objek dan 91 kategori "stuff". Model tersebut akan dilatih ulang agar bisa mendeteksi wajah pada citra termal menggunakan TensorFlow Object Detection API. Digunakan `file protobuf` untuk mengkonfigurasi proses pelatihan dan evaluasi. Skema untuk pipeline pelatihan dapat ditemukan di `object_detection/protos/pipeline.proto` pada repository github Tensorflow.



Gambar 11. Nilai execution time model tfLite.



Gambar 12. Hasil visualisasi inferensi tfLite.



Gambar 13. Hasil visualisasi inferensi tensorflow.

D. Perancangan Pengukuran suhu

Pengukuran suhu akan diimplementasikan dengan memanfaatkan Flir `Image Extractor` dan koordinat `bounding box` yang didapatkan dari `output` model deteksi, akan dilakukan ekstraksi suhu maksimum dari area wajah yang terdeteksi.

Dari citra termal, didapatkan data energi inframerah yang dideteksi oleh kamera termal. Data inframerah tersebut tersimpan dalam format biner pada resolusi 2^{16} bit dengan 65536 kemungkinan nilai, mulai dari 0. Data tersebut dikalibrasi memanfaatkan sebuah algoritma yang mengintergrasikan hukum Plank, hukum Stephan Boltzmann, serapan atmosfer, serapan kamera termal, emisivitas dan perhitungan jarak sehingga menghasilkan data suhu. Hasil kalibrasi berupa array berisi data suhu tiap piksel pada citra termal. Dari array suhu tersebut, diambil data suhu pada wajah saja dan dipilih nilai suhu maksimum. Diagram alir pengukuran suhu bisa dilihat pada Gambar 6.



Gambar 14. Sampel citra dari empat subjek.

IV. UJI COBA DAN EVALUASI

Dalam uji coba perancangan sistem ini akan dibagi menjadi 3 uji coba, yaitu :

A. Mengevaluasi Model

Uji coba ini dijalankan dengan tujuan untuk menguji performa dari 3 model yang dilatih. Model pertama dilatih dengan konfigurasi *default* dengan perubahan *num_steps*. Pada model kedua diterapkan perubahan pada *decay_step* dengan *num_steps*. Model ketiga diterapkan perubahan *decay_step* dan *num_steps* serta augmentasi pada data *training*. *Output* dari evaluasi berupa nilai mAP dan *total loss*. Hasil evaluasi model bisa dilihat pada Tabel 2 dan visualisasi deteksi model bisa dilihat pada Gambar 6.

Salah satu contoh visualisasi deteksi model bisa dilihat pada Gambar 7. Dari visualisasi tersebut didapatkan perbandingan antara *bounding box* hasil deteksi model di sebelah kiri dan *ground truth* di sebelah kanan.

B. Uji Coba Model Inferensi TFLite

Uji coba ini dijalankan dengan tujuan untuk menguji performa dari model Tensorflow yang dilatih sebelum dan sesudah dikonversi menjadi model TFLite. Tujuan dilakukan konversi model adalah untuk membuat model menjadi lebih ringan. Untuk uji coba ini digunakan 109 gambar citra termal

dari *dataset* IRDatabase [7]. Hasil uji coba bisa dilihat di Gambar 8, Gambar 9, Gambar 10, dan Gambar 11.

Didapatkan nilai rata-rata *Confidence Score* deteksi model Tensorflow adalah 0,9999535 dengan total *Execution time* sebanyak 123,203 detik, sedangkan nilai rata-rata *Confidence Score* deteksi model TFLite adalah 0,9999505 dengan total *Execution time* sebanyak 39,95 detik. Bisa disimpulkan bahwa hasil model TFLite lebih ringan dan tidak memiliki penurunan yang signifikan terhadap performa deteksi. Contoh visualisasi hasil bisa dilihat pada Gambar 12 dan Gambar 13. Dari hasil visualisasi tersebut berhasil didapatkan *bounding box* pada wajah dengan kedua model.

C. Memasukan Uji Coba Pengukuran Suhu

Uji coba ini dijalankan dengan tujuan untuk menguji akurasi hasil ekstraksi suhu dari citra termal. Dalam uji coba ini dilakukan 5 kali pengambilan citra termal dan pengukuran suhu untuk setiap 4 subjek, sehingga didapatkan total 20 citra. *Ground Truth* suhu badan subjek didapatkan menggunakan termometer. Sampel citra dari empat subjek bisa dilihat pada Gambar 14. Untuk mengevaluasi hasil uji coba digunakan *Mean Absolute Error* (MAE). Hasil deteksi bisa dilihat pada Tabel 3. Dari hasil uji coba pengukuran suhu 20 citra termal, didapatkan 4 set pengukuran suhu yang terdiri atas *ground truth* dan suhu terdeteksi. Dari hasil pengukuran bisa dihitung nilai MAE tiap set pengukuran dan rata-rata MAE. Didapat rata-rata nilai MAE sebesar 1,51 derajat celcius.

V. KESIMPULAN

Kesimpulan yang di dapatkan berdasarkan Hasil Uji coba yang telah didapat adalah sebagai berikut: (1) *Dataset Tuft Face Database* dan *IRDatabase* didapatkan dari *repository dataset* Online yang disediakan oleh lembaga pendidikan. Hak akses *repository* diperoleh setelah persetujuan perjanjian lisensi. (2) Area wajah pada dari citra termal berhasil dideteksi menggunakan model deteksi objek SSD-MobileNet terlatih dengan nilai *mean average precision* (mAP) lebih dari 90% dan kecepatan eksekusi model inferensi sebesar 39,95 detik untuk 109 citra. (3) *Library Flir Image Extractor* berhasil digunakan untuk mendeteksi suhu wajah dari data-data yang dikumpulkan dengan rata-rata nilai *Mean Absolute Error* (MAE) sebesar 1,51 derajat celcius. (4) Sistem deteksi suhu bisa dibangun dengan mengintegrasikan model deteksi objek SSD-MobileNet terlatih dan *Flir Image Extractor* sebagai ekstraktor suhu dari citra termal. (5) Kinerja dari sistem deteksi suhu berhasil diukur dari hasil evaluasi performa model dan perhitungan nilai *Mean Absolute Error* (MAE) dari data suhu yang terdeteksi.

DAFTAR PUSTAKA

- [1] WHO, "Coronavirus disease (COVID-19)," *World Health Organization*, 2020. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19>.
- [2] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, 1st ed. Cambridge: MIT Press, 2016.
- [3] J.-W. Lin, M.-H. Lu, and Y.-H. Lin, "A Thermal Camera Based Continuous Body Temperature Measurement System," *Proc. IEEE/CVF Int. Conf. Comput. Vis. Work.*, 2019.
- [4] CDC, "Definitions of Symptoms for Reportable Illnesses," *U.S. Department of Health & Human Services*, 2017. <https://www.cdc.gov/quarantine/air/reporting-deaths-illness/definitions-symptoms-reportable-illnesses.html>.
- [5] A. G. Howard *et al.*, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv Prepr. arXiv1704.04861*, pp. 1–9, 2017.
- [6] K. Panetta *et al.*, "A comprehensive database for benchmarking imaging systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 3, pp. 509–520, 2018.
- [7] M. Kopaczka, R. Kolk, and D. Merhof, "A Fully Annotated Thermal Face Database and Its Application for Thermal Facial Expression Recognition," *IEEE Int. Instrum. Meas. Technol. Conf.*, pp. 1–6, 2018.