

**EVALUASI KEPUASAN PELANGGAN BERDASARKAN EKSPRESI  
WAJAH MENGGUNAKAN *REAL TIME DETECTION TRANSFORMER*  
(RT-DETR)**

SKRIPSI

diajukan untuk memenuhi bagian dari syarat memperoleh gelar sarjana komputer  
pada Program Studi Ilmu Komputer



Oleh:

Shafa Meira Wahyono

2007723

**PROGRAM STUDI ILMU KOMPUTER  
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS PENDIDIKAN INDONESIA**

2025

**EVALUASI KEPUASAN PELANGGAN BERDASARKAN EKSPRESI  
WAJAH MENGGUNAKAN *REAL TIME DETECTION TRANSFORMER*  
(RT-DETR)**

Oleh  
Shafa Meira Wahyono

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat memperoleh gelar  
Sarjana pada Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam

©Shafa Meira Wahyono  
Universitas Pendidikan Indonesia  
Januari 2025

Hak cipta dilindungi undang-undang  
Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian, dengan dicetak  
ulang, difotokopi, atau cara lainnya tanpa izin dari penulis

SHAFa MEIRA WAHYONO

EVALUASI KEPUASAN PELANGGAN BERDASARKAN EKSPRESI  
WAJAH MENGGUNAKAN *REAL TIME DETECTION TRANSFORMER* (RT-  
DETR)

disetujui dan disahkan oleh pembimbing:

Pembimbing I



**Prof. Dr. Munir, M. IT.**

NIP. 196603252001121001

Pembimbing II



**Yaya Wihardi, M. Kom.**

NIP. 198903252015041001

Mengetahui,

Ketua Program Studi Ilmu Komputer



**Dr. Muhammad Nursalman, M.T.**

NIP. 197909292006041002

ii

## PERNYATAAN BEBAS PLAGIARISME

Saya yang bertanda tangan di bawah ini:

Nama : Shafa Meira Wahyono  
NIM : 2007723  
Program Studi : Ilmu Komputer  
Judul Karya : Evaluasi Kepuasan Pelanggan Berdasarkan Ekspresi Wajah Menggunakan *Real Time Detection Transformer* (RT-DETR)

Dengan ini menyatakan bahwa karya tulis ini merupakan hasil kerja saya sendiri. Saya menjamin bahwa seluruh isi karya ini, baik sebagian maupun keseluruhan, bukan merupakan plagiarisme dari karya orang lain, kecuali pada bagian yang telah dinyatakan dan disebutkan sumbernya dengan jelas.

Jika di kemudian hari ditemukan pelanggaran terhadap etika akademik atau unsur plagiarisme, saya bersedia menerima sanksi sesuai peraturan yang berlaku di Universitas Pendidikan Indonesia.

Bandung, 5 Januari 2025

Tanda tangan: \_\_\_\_\_

Shafa Meira Wahyono

## KATA PENGANTAR

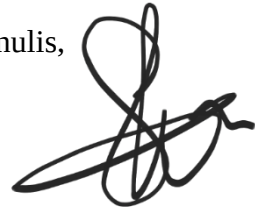
Puji dan syukur penulis panjatkan kepada Allah SWT karena telah memberikan berkah dan rahmat-Nya. Shalawat dan salam semoga terlimpahkan kepada baginda tercinta kita yaitu Nabi Muhammad SAW. Tanpa pertolongan-Nya, tentu penulis tidak akan sanggup untuk menyelesaikan skripsi yang berjudul “Evaluasi Kepuasan Pelanggan Berdasarkan Ekspresi Wajah Menggunakan *Real Time Detection Transformer* (RT-DETR)” dengan baik.

Penulisan skripsi ini memiliki tujuan sebagai salah satu syarat untuk memperoleh gelar Sarjana Ilmu Komputer (S.Kom) pada jenjang studi Strata-1 pada Program Studi Ilmu Komputer di Universitas Pendidikan Indonesia.

Dalam penyusunan skripsi ini, penulis telah berusaha semaksimal kemampuan yang dimiliki. Namun tidak bisa dipungkiri sebagai manusia, penulis pun tidak luput dari kesalahan. Kesalahan tersebut bisa berupa isi, tata bahasa, tanda baca, dan lain-lain. Oleh karena itu, penulis mengharapkan kritik dan saran dari pembaca, agar skripsi ini dapat menjadi lebih baik lagi. Dengan demikian, semoga skripsi ini bisa bermanfaat bagi penulis dan pembaca. Terima kasih

Bandung, 5 Januari 2025

Penulis,



Shafa Meira Wahyono

## UCAPAN TERIMA KASIH

*Alhamdulillah* *rabbi'l'alamin*, puji dan syukur atas kehadiran Allah SWT karena telah memberikan berkah dan rahmat-Nya dalam proses penulisan skripsi ini. Pada proses penulisan skripsi ini, penulis menerima banyak bantuan serta dukungan dari Allah SWT, keluarga, serta banyak pihak lain. Dengan demikian, sudah sepantasnya penulis mengucapkan terima kasih serta penghargaan yang pantas, kepada:

1. Kedua orang tua serta adik penulis yang selalu memberikan doa, dukungan, bantuan, serta semangat kepada penulis dalam menjalankan perkuliahan dan penulisan skripsi ini.
2. Bapak Prof. Dr. Munir, M. IT. selaku dosen pembimbing I yang senantiasa membimbing dan memberi masukan yang bermanfaat untuk penulis.
3. Bapak Yaya Wihardi, M. Kom. selaku dosen pembimbing II yang senantiasa membimbing dan memberi masukan serta saran yang bermanfaat selama penulisan skripsi ini.
4. Ibu Rosa Ariani Sukanto, M.T. selaku dosen pembimbing akademik yang telah memberikan semangat dan motivasi serta saran selama perkuliahan penulis.
5. Ibu Dr. Rani Megasari, M.T., selaku demisioner ketua Departemen Ilmu Komputer Universitas Pendidikan Indonesia.
6. Bapak Dr. Muhammad Nursalman, M.T. selaku ketua Program Studi Ilmu Komputer Universitas Pendidikan Indonesia.
7. Seluruh jajaran dosen serta staff Departemen Pendidikan Ilmu Komputer Universitas Pendidikan Indonesia yang tidak bisa penulis tuliskan satu-persatu yang telah senantiasa membantu, mengarahkan, serta membimbing penulis selama menempuh pendidikan Sarjana.
8. Sahabat penulis, Alfi Amaliandini, yang senantiasa memberikan bantuan, motivasi, dan dukungan kepada penulis dari sekolah menengah pertama hingga pendidikan Sarjana.
9. Teman-teman magang PDDIKTI yang senantiasa memberikan bantuan dan dukungan kepada penulis selama penulisan skripsi serta telah membantu dan memberikan pelajaran yang berharga bagi penulis selama kegiatan magang.

10. Teman-teman Ilmu Komputer angkatan 2020 yang telah berjuang bersama penulis dalam kegiatan perkuliahan maupun luar perkuliahan.
11. Rekan-rekan Badan Eksekutif Mahasiswa periode 2022-2023 yang telah memberikan kepercayaan kepada penulis untuk menjadi Ketua Biro Teknologi dan telah memberikan pengalaman serta pengetahuan baru dalam bidang organisasi.
12. Rekan-rekan kelompok persiapan Duolingo English Test untuk persiapan mengikuti kegiatan IISMA yang telah berbagi semangat serta ilmu yang bermanfaat.
13. Serta seluruh pihak lain yang tidak dapat penulis sebutkan yang telah membantu serta memberikan dukungan dalam perkuliahan serta penulisan skripsi ini hingga selesai.

Akhir kata, penulis berharap agar skripsi ini dapat memberikan manfaat serta menambah ilmu bagi para pembaca. Sekadar tulisan tidak dapat menggambarkan rasa terima kasih penulis kepada seluruh pihak yang telah mendukung. Semoga Tuhan Yang Maha Esa senantiasa memberikan kebahagiaan dan kesehatan untuk sekarang dan seterusnya. *Aamiin ya rabbal alamin.*

# EVALUASI KEPUASAN PELANGGAN BERDASARKAN EKSPRESI WAJAH MENGGUNAKAN *REAL TIME DETECTION TRANSFORMER* (RT- DETR)

Oleh  
Shafa Meira Wahyono – shafameira@upi.edu  
2007723

## ABSTRAK

Ekspresi wajah merupakan indikator paling baik dalam mengetahui perasaan manusia karena hanya teridentifikasi dalam waktu singkat yaitu 0,5 detik. Maka dari itu, ekspresi wajah dapat digunakan sebagai indikator evaluasi kepuasan pelanggan. Namun, karena perubahan ekspresi wajah yang cepat, penggunaan teknologi untuk pengenalan wajah serta klasifikasi ekspresi menjadi lebih ideal. Selain itu, pengimplementasian teknologi di dunia nyata sering kali dihadapi tantangan, seperti keterbatasan perangkat keras. Berdasarkan permasalahan tersebut, dilakukan penelitian untuk membangun model kecerdasan buatan yang mampu mengevaluasi kepuasan pelanggan berdasarkan pengenalan ekspresi wajah dengan memerhatikan kemungkinan keterbatasan sumber daya yang dapat muncul ketika pengimplementasian. Pertama, model dikembangkan untuk mendeteksi wajah menggunakan metode RT-DETR (*Real Time-Detection Transformer*) dengan *backbone* ResNet-18 dan LCNet-0.25 (*Lightweight CPU Convolutional Neural Network*). Kedua, model mengklasifikasikan 2 tipe ekspresi wajah, yaitu 7 ekspresi wajah dan 3 ekspresi wajah, menggunakan metode *Real-Time CNN*. Penelitian yang dilakukan, menunjukkan bahwa metode RT-DETR dengan *backbone* ResNet-18 mendapatkan kinerja terbaik dengan 35.0% AP, 6.64 FPS, dan parameter 20M serta *Real-Time CNN* dengan 3 ekspresi wajah dengan performa *micro-average* F1-score 68.4%. Kombinasi dari model RT-DETR dan model *Real-Time CNN* memiliki kinerja 2.1% AP dan 4.7 FPS.

*Kata Kunci: Detection Transformer, Pengenalan Ekspresi Wajah, Real Time.*



# CUSTOMER SATISFACTION EVALUATION THROUGH FACIAL EXPRESSION USING REAL TIME DETECTION TRANSFORMER (RT-DETR)

Arranged by  
Shafa Meira Wahyono – shafameira@upi.edu  
2007723

## ABSTRACT

*Facial expression is the best indicator to recognize emotion because it is only recognizable in a such short time, which is about 0,5 seconds. Therefore, facial expression is usable as customer satisfaction evaluation indicator. However, as facial expression changes quickly, it would be more ideal to use technology to recognize the faces and classify the expression. Furthermore, technology implementation in real-world was often meet with computation limitation. Based on those problems, this research is done to build Artificial Intelligence model that can evaluate customer satisfaction through facial expression, while keeping attention on the possibility of resource limitation when implementing. Firstly, a model is developed to detect faces using methods, such as RT-DETR (Real Time-Detection Transformer) with backbone ResNet-18 and LCNet-0.25 (Lightweight CPU Convolutional Neural Network. Secondly, a model will classify 2 types of facial expression, such as 7 type of facial expression and 3 type of facial expression, using Real-Time CNN method. RT-DETR with ResNet-18 as backbone achieves best performance with 35.0% AP, 6.64 FPS, and parameter of 20M, while Real-Time CNN with 3 types facial expression achieves micro-average F1-score of 68.4%. The combination of RT-DETR and Real Time-CNN achieves 4.7% AP dan 5.94 FPS.*

*Keywords: Customer Satisfaction, Detection Transformer, Facial Expression Recognition, Real Time*

## DAFTAR ISI

|                                                                       |      |
|-----------------------------------------------------------------------|------|
| PERNYATAAN BEBAS PLAGIARISME .....                                    | iii  |
| KATA PENGANTAR .....                                                  | iv   |
| UCAPAN TERIMA KASIH.....                                              | v    |
| ABSTRAK .....                                                         | vii  |
| ABSTRACT.....                                                         | viii |
| DAFTAR ISI.....                                                       | ix   |
| DAFTAR TABEL .....                                                    | xii  |
| DAFTAR GAMBAR .....                                                   | xiii |
| DAFTAR SINGKATAN .....                                                | xv   |
| BAB I PENDAHULUAN .....                                               | 1    |
| 1.1. Latar Belakang.....                                              | 1    |
| 1.2. Rumusan Masalah .....                                            | 5    |
| 1.3. Tujuan Penelitian.....                                           | 5    |
| 1.4. Manfaat Penelitian.....                                          | 6    |
| 1.5. Ruang Lingkup Penelitian .....                                   | 6    |
| 1.6. Sistematika Penelitian .....                                     | 6    |
| BAB II TINJAUAN PUSTAKA.....                                          | 8    |
| 2.1. Peta Literatur .....                                             | 8    |
| 2.2. Penelitian Terkait.....                                          | 8    |
| 2.3. Kepuasan Pelanggan.....                                          | 10   |
| 2.4. Ekspresi Wajah.....                                              | 11   |
| 2.5. Kecerdasan Buatan .....                                          | 13   |
| 2.6. <i>Deep Learning</i> .....                                       | 14   |
| 2.7. Jaringan Syaraf Tiruan ( <i>Artificial Neural Network</i> )..... | 14   |
| 2.7.1. <i>Convolutional Neural Network</i> .....                      | 15   |
| 2.8. <i>Computer Vision</i> .....                                     | 15   |
| 2.8.1. <i>Bounding Box</i> .....                                      | 17   |
| 2.8.2. <i>Region of Interest (ROI)</i> .....                          | 17   |
| 2.8.3. <i>Intersection over Union (IoU)</i> .....                     | 18   |
| 2.9. Deteksi Wajah.....                                               | 18   |
| 2.10. Klasifikasi Ekspresi .....                                      | 19   |
| 2.11. <i>Facial Expression Recognition (FER)</i> .....                | 20   |

|                                   |                                                                                                |    |
|-----------------------------------|------------------------------------------------------------------------------------------------|----|
| 2.12.                             | Implementasi FER untuk Evaluasi Kepuasan Pelanggan.....                                        | 20 |
| 2.13.                             | DETR ( <i>Detection Transformer</i> ) .....                                                    | 21 |
| 2.13.1.                           | ResNet ( <i>Residual Network</i> ) .....                                                       | 23 |
| 2.13.2.                           | Metode Transformer .....                                                                       | 24 |
| 2.13.2.1.                         | Encoder.....                                                                                   | 24 |
| 2.13.2.2.                         | Decoder .....                                                                                  | 25 |
| 2.13.3.                           | <i>Feed-Forward Networks (FFNs)</i> .....                                                      | 25 |
| 2.14.                             | RT-DETR ( <i>Real Time-Detection Transformer</i> ) .....                                       | 26 |
| 2.14.1.                           | Encoder Hybrid .....                                                                           | 27 |
| 2.14.2.                           | IoU-aware Query Selection .....                                                                | 27 |
| 2.15.                             | L-DETR ( <i>Light-Weight Detector for End-to-End Object Detection With Transformers</i> )..... | 28 |
| 2.15.1.                           | PP-LCNet ( <i>Lightweight CPU Convolutional Neural Network</i> ) .....                         | 29 |
| 2.16.                             | Real-Time CNN ( <i>Convolutional Neural Networks</i> ).....                                    | 30 |
| BAB III METODE PENELITIAN.....    |                                                                                                | 31 |
| 3.1.                              | Desain Penelitian .....                                                                        | 31 |
| 3.1.1.                            | Rumusan Masalah .....                                                                          | 32 |
| 3.1.2.                            | Tinjauan Pustaka .....                                                                         | 32 |
| 3.1.3.                            | Pengumpulan Data.....                                                                          | 32 |
| 3.1.4.                            | Analisa dan Evaluasi .....                                                                     | 32 |
| 3.1.5.                            | Rancangan Implementasi.....                                                                    | 32 |
| 3.1.6.                            | Penarikan Kesimpulan.....                                                                      | 32 |
| 3.2.                              | Set Data .....                                                                                 | 33 |
| 3.2.1.                            | WIDER-face .....                                                                               | 33 |
| 3.2.2.                            | FER-2013 ( <i>Facial Expression Recognition-2013</i> ).....                                    | 34 |
| 3.2.3.                            | IMED ( <i>Indonesian Mixed Emotion Dataset</i> ).....                                          | 35 |
| 3.2.4.                            | Set Data Demonstrasi Pengujian .....                                                           | 37 |
| 3.3.                              | Rancangan Implementasi.....                                                                    | 38 |
| 3.3.1.                            | Model Deteksi Wajah .....                                                                      | 39 |
| 3.3.2.                            | Model Klasifikasi Ekspresi.....                                                                | 40 |
| 3.3.3.                            | Integrasi Model Deteksi Wajah dan Klasifikasi Ekspresi .....                                   | 40 |
| 3.4.                              | Analisa dan Evaluasi .....                                                                     | 41 |
| 3.5.                              | Lingkungan Komputasi .....                                                                     | 42 |
| BAB IV HASIL DAN PEMBAHASAN ..... |                                                                                                | 43 |

|                                |                                                     |    |
|--------------------------------|-----------------------------------------------------|----|
| 4.1                            | Praproses Set Data Pengembangan.....                | 43 |
| 4.1.1                          | WIDER-face .....                                    | 43 |
| 4.1.2                          | FER-2013 dan IMED .....                             | 46 |
| 4.2                            | Model Deteksi Wajah .....                           | 49 |
| 4.2.1                          | RT-DETR (ResNet-18) .....                           | 49 |
| 4.2.2                          | RT-DETR (LCNet-0.25) .....                          | 51 |
| 4.2.3                          | Pemilihan Bobot Terbaik (Deteksi Wajah) .....       | 53 |
| 4.3                            | Model Klasifikasi Ekspresi.....                     | 55 |
| 4.3.1                          | Real Time-CNN (7 Ekspresi) .....                    | 56 |
| 4.3.2                          | Real Time-CNN (3 Ekspresi) .....                    | 58 |
| 4.3.3                          | Pemilihan Bobot Terbaik (Klasifikasi Ekspresi)..... | 60 |
| 4.4                            | Praproses Set Data Pengujian .....                  | 61 |
| 4.5                            | Integrasi Model Deteksi dan Klasifikasi .....       | 63 |
| 4.6                            | Pembahasan .....                                    | 64 |
| BAB V SIMPULAN DAN SARAN ..... |                                                     | 68 |
| 5.1.                           | Simpulan.....                                       | 68 |
| 5.2.                           | Saran .....                                         | 69 |
| DAFTAR PUSTAKA .....           |                                                     | 71 |

## DAFTAR TABEL

|                                                                                 |    |
|---------------------------------------------------------------------------------|----|
| Tabel 4.1 Atribut anotasi <i>ground truth</i> WIDER-face .....                  | 44 |
| Tabel 4.2 Atribut anotasi WIDER-face sesuai format COCO .....                   | 45 |
| Tabel 4.3 Angka label ekspresi set data FER-2013 dan IMED (7 ekspresi).....     | 47 |
| Tabel 4.4 Angka label ekspresi set data FER-2013 dan IMED (3 ekspresi).....     | 48 |
| Tabel 4.5 Metrik evaluasi <i>training</i> dan validasi <i>epoch</i> ke-71 ..... | 50 |
| Tabel 4.6 Metrik evaluasi <i>training</i> dan validasi <i>epoch</i> ke-8 .....  | 52 |
| Tabel 4.7 Hasil pengujian RT-DETR dengan ResNet-18 dan LCNet-0.25.....          | 53 |
| Tabel 4.8 Hasil evaluasi test-set WIDER.....                                    | 55 |
| Tabel 4.9 Evaluasi F1-score model Real Time-CNN (7 ekspresi).....               | 57 |
| Tabel 4.10 Sampel <i>ground truth</i> dan prediksi (7 ekspresi).....            | 58 |
| Tabel 4.11 Evaluasi F1-score model Real Time-CNN (3 ekspresi).....              | 60 |
| Tabel 4.12 Sampel <i>ground truth</i> dan prediksi (3 ekspresi).....            | 60 |
| Tabel 4.13 Evaluasi F1-score model Real Time-CNN .....                          | 61 |
| Tabel 4.14 Hasil evaluasi model integrasi .....                                 | 63 |
| Tabel 4.15 Performa RT-DETR (ResNet-18) pada seluruh set data.....              | 64 |
| Tabel 4.16 Tabel evaluasi parameter dan performa .....                          | 65 |
| Tabel 4.17 Sampel prediksi benar dan salah.....                                 | 66 |

## DAFTAR GAMBAR

|                                                                       |    |
|-----------------------------------------------------------------------|----|
| Gambar 2.1 Peta Literatur .....                                       | 8  |
| Gambar 2.2 Enam ekspresi dasar dan ekspresi netral .....              | 11 |
| Gambar 2.3 Prinsip kerja jaringan syaraf tiruan .....                 | 15 |
| Gambar 2.4 Tulisan tangan dari digit angka (Lecun, 1999) .....        | 16 |
| Gambar 2.5 Deteksi wajah .....                                        | 16 |
| Gambar 2.6 Contoh penggunaan <i>bounding box</i> .....                | 17 |
| Gambar 2.7 Ilustrasi penggunaan ROI .....                             | 17 |
| Gambar 2.8 Ilustrasi Intersection over Union (IoU) .....              | 18 |
| Gambar 2.9 Arsitektur CNN .....                                       | 19 |
| Gambar 2.10 Cara kerja DETR .....                                     | 22 |
| Gambar 2.11 Arsitektur DETR .....                                     | 22 |
| Gambar 2.12 Arsitektur Transformer untuk deteksi objek .....          | 24 |
| Gambar 2.13 Arsitektur RT-DETR .....                                  | 26 |
| Gambar 2.14 Arsitektur L-DETR .....                                   | 28 |
| Gambar 2.15 Arsitektur PP-LCNet .....                                 | 29 |
| Gambar 3.1 Desain penelitian .....                                    | 31 |
| Gambar 3.2 Sampel set data WIDER-face .....                           | 33 |
| Gambar 3.3 Sampel set data FER-2013 .....                             | 34 |
| Gambar 3.4 Bagan jumlah ekspresi FER-2013 .....                       | 35 |
| Gambar 3.5 Sampel set data IMED .....                                 | 35 |
| Gambar 3.6 Bagan jumlah ekspresi IMED .....                           | 36 |
| Gambar 3.7 Sampel set data demonstrasi pengujian .....                | 37 |
| Gambar 3.8 Persentase jenis ekspresi .....                            | 37 |
| Gambar 3.9 Rancangan implementasi .....                               | 38 |
| Gambar 4.1 Sampel anotasi orisinal WIDER-face .....                   | 43 |
| Gambar 4.2 Sampel program format COCO .....                           | 44 |
| Gambar 4.3 Sampel anotasi WIDER-face sesuai format COCO .....         | 45 |
| Gambar 4.4 Praproses IMED .....                                       | 46 |
| Gambar 4.5 Pseudocode praproses IMED .....                            | 46 |
| Gambar 4.6 Sampel set data FER-2013 dan IMED .....                    | 47 |
| Gambar 4.7 Sampel set data FER-2013 dan IMED .....                    | 47 |
| Gambar 4.8 Bagan jumlah ekspresi FER-2013 dan IMED (7 Ekspresi) ..... | 48 |
| Gambar 4.9 Bagan jumlah ekspresi FER-2013 dan IMED (3 Ekspresi) ..... | 48 |
| Gambar 4.10 Sampel foto <i>ground truth</i> .....                     | 49 |
| Gambar 4.11 <i>Training</i> loss ResNet-18 .....                      | 50 |
| Gambar 4.12 <i>Average Precision</i> validasi ResNet-18 .....         | 50 |
| Gambar 4.13 Sampel foto prediksi dengan RT-DETR (ResNet-18) .....     | 51 |
| Gambar 4.14 Konfigurasi backbone LCNet-0.25 .....                     | 51 |
| Gambar 4.15 <i>Training</i> loss LCNet-0.25 .....                     | 52 |
| Gambar 4.16 <i>Average Precision</i> validasi LCNet-0.25 .....        | 52 |
| Gambar 4.17 Sampel foto prediksi dengan RT-DETR LCNet-0.25 .....      | 53 |
| Gambar 4.18 Evaluasi easy-set WIDER .....                             | 54 |
| Gambar 4.19 Evaluasi medium-set WIDER .....                           | 54 |

|                                                                                         |    |
|-----------------------------------------------------------------------------------------|----|
| Gambar 4.20 Evaluasi hard-set WIDER .....                                               | 54 |
| Gambar 4.21 Real Time-CNN (7 ekspresi) <i>training</i> dan <i>validation</i> loss ..... | 56 |
| Gambar 4.22 <i>Confusion matrix</i> Real Time-CNN (7 ekspresi) <i>epoch</i> ke-92 ..... | 57 |
| Gambar 4.23 Real Time-CNN (3 ekspresi) <i>training</i> dan <i>validation</i> loss ..... | 59 |
| Gambar 4.24 <i>Confusion matrix</i> Real Time-CNN (3 ekspresi).....                     | 59 |
| Gambar 4.25 Sampel <i>frame</i> video dengan ROI (garis biru muda) .....                | 61 |
| Gambar 4.26 Sampel <i>ground truth</i> .....                                            | 62 |
| Gambar 4.27 Sampel gambar <i>ground truth</i> dan prediksi dengan anotasi .....         | 64 |

## DAFTAR SINGKATAN

|            |                                                                         |
|------------|-------------------------------------------------------------------------|
| AI         | : <i>Artificial Intelligence</i>                                        |
| ANN        | : <i>Artificial Neural Network</i>                                      |
| AP         | : <i>Average Precision</i>                                              |
| BBOX       | : <i>Bounding Box</i>                                                   |
| CCTV       | : <i>Closed-Circuit Television</i>                                      |
| CNN        | : <i>Convolutional Neural Network</i>                                   |
| COCO       | : <i>Common Objects in Context Dataset</i>                              |
| CPU        | : <i>Central Processing Unit</i>                                        |
| DETR       | : <i>Detection Transformer</i>                                          |
| DINO       | : <i>DETR with Improved DeNoising Anchor</i>                            |
| FACS       | : <i>Facial Action Coding System</i>                                    |
| FER        | : <i>Facial Expression Recognition</i>                                  |
| FER-2013   | : <i>Facial Expression Recognition-2013</i>                             |
| FFN        | : <i>Feed Forward Network</i>                                           |
| FPS        | : <i>Frame per Second</i>                                               |
| GPU        | : <i>Graphics Processing Unit</i>                                       |
| IMED       | : <i>Indonesian Mixed Emotion Dataset</i>                               |
| IoU        | : <i>Intersection over Union</i>                                        |
| JSON       | : <i>JavaScript Object Notation</i>                                     |
| JST        | : <i>Jaringan Syaraf Tiruan</i>                                         |
| KB         | : <i>Kilo Bytes</i>                                                     |
| LCNet-0.25 | : <i>Lightweight CPU Convolutional Neural Network dengan skala 0.25</i> |
| L-DETR     | : <i>Lightweight Detection Transformer</i>                              |
| M          | : <i>Millions</i>                                                       |
| MB         | : <i>Mega Bytes</i>                                                     |
| MKLDNN     | : <i>Math Kernel Library for Deep Neural Networks</i>                   |
| NLP        | : <i>Natural Language Processing</i>                                    |



|               |                                                     |
|---------------|-----------------------------------------------------|
| OCR           | : <i>Optical Character Recognition</i>              |
| ONNX          | : <i>Open Neural Network Exchange</i>               |
| PP-LCNET      | : <i>PaddlePaddle-LCNet</i>                         |
| Real Time-CNN | : <i>Real Time-Convolutional Neural Network</i>     |
| ResNet-18     | : <i>Residual Network</i> dengan 18 layers          |
| ROI           | : <i>Region of Interest</i>                         |
| RT-DETR       | : <i>Real Time Detection Transformer</i>            |
| SMOTE         | : <i>Synthetic Minority Over-sampling Technique</i> |
| YAML          | : <i>Yet Another Markup Language</i>                |
| YOLO          | : <i>You Only Look Once</i>                         |

## DAFTAR PUSTAKA

- (IBM), I. B. M. C. (n.d.). *What is artificial intelligence (AI)?*  
<https://www.ibm.com/topics/artificial-intelligence>
- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan, M., Van Essen, B. C., Awwal, A. A. S., & Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. *Electronics (Switzerland)*, 8(3), 1–67. <https://doi.org/10.3390/electronics8030292>
- Arriaga, O., Valdenegro-Toro, M., & Plöger, P. (2017). Real-time convolutional neural networks for emotion and gender classification. *ArXiv Preprint ArXiv:1710.07557*.
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 77–91.
- Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-end object detection with transformers. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12346 LNCS, 213–229. [https://doi.org/10.1007/978-3-030-58452-8\\_13](https://doi.org/10.1007/978-3-030-58452-8_13)
- Ceccacci, S., Generosi, A., Giraldi, L., & Mengoni, M. (2023). Emotional valence from facial expression as an experience audit tool: An empirical study in the context of opera performance. *Sensors*, 23(5).  
<https://doi.org/10.3390/s23052688>
- Chang, W.-J., Schmelzer, M., Kopp, F., Hsum, C.-H., Su, J.-P., Chen, L.-B., & Chen, M.-C. (2023). *A deep learning facial expression recognition based scoring system for restaurants*. 3–6.
- Chetana, P., Gunasekaran, M., & Tiwari, S. K. (2022). An efficient system for classifying customer facial expressions to obtain feedback with accuracy at unmanned restaurants using improved support vector machine and flattened convolutional neural networks. *14th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics, MACS*

- 2022, 1–8. <https://doi.org/10.1109/MACS56771.2022.10022836>
- Cui, C., Gao, T., Wei, S., Du, Y., Guo, R., Dong, S., Lu, B., Zhou, Y., Lv, X., Liu, Q., Hu, X., Yu, D., & Ma, Y. (2021). *PP-LCNet: A lightweight CPU convolutional neural network*. 1–8. <http://arxiv.org/abs/2109.15099>
- Dai, J. (2020). Real-time and accurate object detection on edge device with TensorFlow Lite. *Journal of Physics: Conference Series*, 1651(1). <https://doi.org/10.1088/1742-6596/1651/1/012114>
- Darwin, C. (1898). The expression of the emotions in man and animals. *The Portable Darwin*, 364–393. <https://cir.nii.ac.jp/crid/1571417124682544000>
- Decaro, C., Montanari, G., Molinariz, R., Gilberti, A., Bagnoli, D., Bianconi, M., & Bellanca, G. (2019). Machine learning approach for prediction of hematic parameters in hemodialysis patients. *IEEE Journal of Translational Engineering in Health and Medicine*, PP, 1. <https://doi.org/10.1109/JTEHM.2019.2938951>
- Eikvil, L. (1993). Optical character recognition. *Citeseer. Ist. Psu. Edu/142042.Html*, 26.
- Ekman, P. (1970). Universal facial expressions of emotion. *California Mental Health Research Digest*, 8(4), 151–158.
- Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and cue to deception. In *Psychiatry: Journal for the Study of Interpersonal Processes* (Vol. 32, Issue 1, pp. 88–106).
- Ekman, P., & Friesen, W. V. (1975). Unmasking the face: A guide to recognizing emotions from facial clues. In *Unmasking the face: A guide to recognizing emotions from facial clues*. Prentice-Hall.
- Ekman, P., & Friesen, W. V. (1978). *Facial action coding system* (Issue v. 1). Consulting Psychologists Press. <https://books.google.co.id/books?id=08l6wgEACAAJ>
- Ekman, P., & Heider, K. G. (1988). The universality of a contempt expression: A replication. *Motivation and Emotion*, 12(3), 303–308. <https://doi.org/10.1007/BF00993116>
- Enholt, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2022). Artificial

- intelligence and business value: A literature review. *Information Systems Frontiers*, 24(5), 1709–1734.
- Goldblum, M., Sourì, H., Ni, R., Shu, M., Prabhu, V., Somepalli, G., Chattopadhyay, P., Ibrahim, M., Bardes, A., Hoffman, J., Chellappa, R., Wilson, A. G., & Goldstein, T. (2023). Battle of the backbones: A large-scale comparison of pretrained models across computer vision tasks. *Advances in Neural Information Processing Systems*, 36(NeurIPS), 1–29.
- González-Rodríguez, M. R., Díaz-Fernández, M. C., & Pacheco Gómez, C. (2020). Facial-expression recognition: An emergent approach to the measurement of tourist satisfaction through emotions. *Telematics and Informatics*, 51(October 2019), 101404.  
<https://doi.org/10.1016/j.tele.2020.101404>
- Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., & Lee, D.-H. (2013). Challenges in representation learning: A report on three machine learning contests. *Neural Information Processing: 20th International Conference, ICONIP 2013, Daegu, Korea, November 3-7, 2013. Proceedings, Part III* 20, 117–124.
- Hamzah, A. A. bin, & Shamsudin, M. F. (2020). Why customer satisfaction is important to business? *Journal of Undergraduate Social & Technology*, 2(1), 2710–6918. [www.jusst.abrn.asia](http://www.jusst.abrn.asia)
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- Karadağ, B. (2018). Measurement of customer satisfaction through emotion analysis in the banking sector. *EasyChair Preprints*, April.  
<https://doi.org/10.29007/x72l>
- Kelleher, J. D. (2019). *Deep learning*. MIT Press.  
[https://books.google.co.id/books?hl=en&lr=&id=b06qDwAAQBAJ&oi=fnd&pg=PP9&dq=what+is+deep+learning&ots=\\_pASRJpWVO&sig=p9oL59JbaWSnwP\\_QR4c0Rfjmgno&redir\\_esc=y#v=onepage&q=what is deep learning&f=false](https://books.google.co.id/books?hl=en&lr=&id=b06qDwAAQBAJ&oi=fnd&pg=PP9&dq=what+is+deep+learning&ots=_pASRJpWVO&sig=p9oL59JbaWSnwP_QR4c0Rfjmgno&redir_esc=y#v=onepage&q=what%20is%20deep%20learning&f=false)

- Lecun. (1999). *MNIST dataset (Modified National Institute of Standards and Technology)*. <http://yann.lecun.com/exdb/mnist/index.html>
- Li, T., Wang, J., & Zhang, T. (2022). L-DETR: A light-weight detector for end-to-end object detection with transformers. *IEEE Access*, 10(October), 105685–105692. <https://doi.org/10.1109/ACCESS.2022.3208889>
- Liliana, D. Y., Basaruddin, T., & Oriza, I. I. D. (2018). The Indonesian Mixed Emotion Dataset (IMED): A facial expression dataset for mixed emotion recognition. *ACM International Conference Proceeding Series*, 56–60. <https://doi.org/10.1145/3293663.3293671>
- Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8693 LNCS(PART 5), 740–755. [https://doi.org/10.1007/978-3-319-10602-1\\_48](https://doi.org/10.1007/978-3-319-10602-1_48)
- Liu, Y., Wang, F., Deng, J., Zhou, Z., Sun, B., & Li, H. (2022). MogFace: Towards a deeper appreciation on face detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2022-June*, 4083–4092. <https://doi.org/10.1109/CVPR52688.2022.00406>
- Lv, W., Zhao, Y., Xu, S., Wei, J., Wang, G., Cui, C., Du, Y., Dang, Q., & Liu, Y. (2023). *DETRs beat YOLOs on real-time object detection*. <http://arxiv.org/abs/2304.08069>
- Meena, G., Mohbey, K. K., & Kumar, S. (2023). Sentiment analysis on images using convolutional neural networks based Inception-V3 transfer learning approach. *International Journal of Information Management Data Insights*, 3(1), 100174. <https://doi.org/10.1016/j.jjime.2023.100174>
- Mesuga, R., & Bayanay, B. J. (2021). *A deep transfer learning approach on identifying glitch wave-form in gravitational wave data*. May. <https://doi.org/10.48550/arXiv.2107.01863>
- Meyers, R. A. (2001). *Encyclopedia of physical science and technology*. <https://www.sciencedirect.com/referencework/9780122274107/encyclopedia-of-physical-science-and-technology#book-info>

- Naqa, I. El, & Murphy, M. J. (2015). Machine learning in radiation oncology. *Machine Learning in Radiation Oncology*, 3–11. <https://doi.org/10.1007/978-3-319-18305-3>
- Ouyang, H. (2023). *DEYOv3: DETR with YOLO for real-time object detection*. <http://arxiv.org/abs/2309.11851>
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1994). Alternative scales for measuring service quality: A comparative assessment based on psychometric and diagnostic criteria. *Journal of Retailing*, 70(3), 201–230. [https://doi.org/10.1016/0022-4359\(94\)90033-7](https://doi.org/10.1016/0022-4359(94)90033-7)
- Plastiras, G., Terzi, M., Kyrkou, C., & Theodoridis, T. (2018). Edge intelligence: Challenges and opportunities of near-sensor machine learning applications. *Proceedings of the International Conference on Application-Specific Systems, Architectures and Processors, 2018-July*. <https://doi.org/10.1109/ASAP.2018.8445118>
- Pratomo, A. H., Kaswidjanti, W., & Mu'arifah, S. (2020). Implementasi algoritma region of interest (ROI) untuk meningkatkan performa algoritma deteksi dan klasifikasi kendaraan. *J. Teknol. Inf. Dan Ilmu Komput*, 7(1), 155–162.
- Qi, D., Tan, W., Yao, Q., & Liu, J. (2023). YOLO5Face: Why reinventing a face detector. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13805 LNCS, 228–244. [https://doi.org/10.1007/978-3-031-25072-9\\_15](https://doi.org/10.1007/978-3-031-25072-9_15)
- Ramla, M., Sangeetha, S., & Nickolas, S. (2022). Chapter 6 - towards building an efficient deep neural network based on YOLO detector for fetal head localization from ultrasound images. In R. Sridhar, G. R. Gangadharan, M. Sheng, & R. B. T.-E.-T. in P. H. S. S. Shankaran (Eds.), *Cognitive Data Science in Sustainable Computing* (pp. 137–156). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-323-90585-5.00005-9>
- Razak, A. A., & Shamsudin, M. F. (2019). The influence of atmospheric experience on theme park tourist's satisfaction and loyalty in Malaysia. *International Journal of Innovation, Creativity and Change*, 6(9), 10–20.
- Revina, I. M., & Emmanuel, W. R. S. (2021). A survey on human face expression

- recognition techniques. *Journal of King Saud University-Computer and Information Sciences*, 33(6), 619–628.
- Ringler, C. (2021). Truth and lies: The impact of modality on customer feedback. *Journal of Business Research*, 133(May), 376–387.  
<https://doi.org/10.1016/j.jbusres.2021.05.014>
- Robinson, D. L. (2008). Brain function, emotional experience and personality. *Netherlands Journal of Psychology*, 64(4), 152–168.  
<https://doi.org/10.1007/bf03076418>
- Sajjad, M., Ullah, F. U. M., Ullah, M., Christodoulou, G., Alaya Cheikh, F., Hijji, M., Muhammad, K., & Rodrigues, J. J. P. C. (2023). A comprehensive survey on deep facial expression recognition: challenges, applications, and future guidelines. *Alexandria Engineering Journal*, 68, 817–840.  
<https://doi.org/10.1016/j.aej.2023.01.017>
- Slim, M., Kachouri, R., & Atitallah, A. Ben. (2018). Customer satisfaction measuring based on the most significant facial emotion. *2018 15th International Multi-Conference on Systems, Signals and Devices, SSD 2018*, 502–507. <https://doi.org/10.1109/SSD.2018.8570588>
- Sugianto, N., Tjondronegoro, D., & Tydd, B. (2018). Deep Residual Learning for Analyzing Customer Satisfaction using Video Surveillance. *Proceedings of AVSS 2018 - 2018 15th IEEE International Conference on Advanced Video and Signal-Based Surveillance*, 1–6.  
<https://doi.org/10.1109/AVSS.2018.8639478>
- Szeliski, R. (2010). *Computer vision: Algorithms and applications*. Springer London. <https://books.google.co.id/books?id=bXzAlkODwa8C>
- TensorFlow. (2019). *TensorFlow Lite*. Tensorflow.Org.
- Terven, J., Cordova-Esparza, D. M., Ramirez-Pedraza, A., Chavez-Urbiola, E. A., & Romero-Gonzalez, J. A. (2023). *Loss functions and metrics in deep learning*. <http://arxiv.org/abs/2307.02694>
- Thompson, N., Greenewald, K., Lee, K., & Manso, G. F. (2023). *The computational limits of deep learning*.  
<https://doi.org/10.21428/bf6fb269.1f033948>

- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *CoRR*, abs/1706.0. <http://arxiv.org/abs/1706.03762>
- Vemou, K., Horvath, A., & Zerdick, T. (2021). Facial emotion recognition. *Lecture Notes in Mechanical Engineering*, 1, 751–761. [https://doi.org/10.1007/978-981-15-9956-9\\_73](https://doi.org/10.1007/978-981-15-9956-9_73)
- Venkatesan, R., & Li, B. (2018). *Convolutional neural networks in visual computing: A concise guide*. CRC Press. <https://books.google.co.id/books?id=aOtPAQAACAAJ>
- Vinh, T. Q., & Tran Dac Thinh, P. (2019). Advertisement system based on facial expression recognition and convolutional neural network. *Proceedings - 2019 19th International Symposium on Communications and Information Technologies, ISCIT 2019*, 476–480. <https://doi.org/10.1109/ISCIT.2019.8905134>
- Wen, Z., Lin, W., Wang, T., & Xu, G. (2023). Distract your attention: Multi-head cross attention network for facial expression recognition. *Biomimetics*, 8(2). <https://doi.org/10.3390/biomimetics8020199>
- Wu, W., Peng, H., & Yu, S. (2023). YuNet: A tiny millisecond-level face detector. *Machine Intelligence Research*, 20. <https://doi.org/10.1007/s11633-023-1423-y>
- Yang, M. H., Kriegman, D. J., & Ahuja, N. (2002). Detecting faces in images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1), 34–58. <https://doi.org/10.1109/34.982883>
- Yang, S., Luo, P., Loy, C.-C., & Tang, X. (2016). WIDER face: A face detection benchmark. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5525–5533.
- Zhu, Y., Cai, H., Zhang, S., Wang, C., & Xiong, Y. (2020). TinaFace: Strong but simple baseline for face detection. <http://arxiv.org/abs/2011.13183>