

Deteksi Emosi Pada Citra Wajah Dengan Deep Learning Sebagai Alat Pendukung Terapi Bagi Pengidap Alexithymia

Alfin Yogi Setyawan^{1*}
Jumadi²
Eva Nurlatifah³

^{1,2,3}Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sunan Gunung Djati Bandung, Jalan Cileunyi, Kec. Cileunyi, Kabupaten Bandung, Jawa Barat 40622, Indonesia
¹alfinyogisetiawan@gmail.com, ²jumadi@uinsgd.ac.id, ³evanurlatifah@uinsgd.ac.id

***Penulis Korespondensi:**
Alfin Yogi Setyawan
alfinyogisetiawan@gmail.com

Abstrak

Alexithymia adalah kondisi yang ditandai dengan kesulitan mengidentifikasi dan mengungkapkan emosi secara verbal, yang dapat menghambat kemampuan seseorang dalam memahami dan mengelola emosi mereka. Penelitian ini bertujuan untuk menerapkan dan mengembangkan sebuah model yang dapat mendeteksi emosi dengan menggunakan arsitektur MobileNetV2 dalam terapi untuk individu yang mengalami alexithymia. Metode ini menggunakan data FER-2013, yang terdiri dari 35.887 gambar wajah grayscale dalam 7 kategori emosi yaitu marah, jijik, takut, senang, netral, sedih, dan terkejut. Menggunakan pendekatan deep learning berdasarkan CRISP-DM, penelitian ini memulai dengan normalisasi dan peningkatan data untuk meningkatkan ketahanan model terhadap variasi citra. Model yang dikembangkan mencapai akurasi pelatihan sebesar 67,7% dan akurasi validasi sebesar 65,3%, menunjukkan kemampuan yang signifikan dalam mengenali dan mengklasifikasikan emosi dari citra wajah. Evaluasi Menggunakan confusion matrix, model menghasilkan nilai precision sebesar 64,9%, recall sebesar 65,4%, dan F1-score sebesar 63,7% untuk setiap kelas emosi. Penelitian ini mengimplikasikan potensi pengembangan sistem yang dapat mendukung terapi psikologis, khususnya untuk membantu individu dengan alexithymia dalam memahami dan mengelola emosi mereka melalui analisis ekspresi wajah, menghadirkan teknologi yang sensitif terhadap ekspresi emosional.

Kata Kunci: Alexithymia; Deep Learning; Deteksi Emosi; FER-2013; MobileNetV2

Abstract

Alexithymia is a condition characterized by difficulty in identifying and verbally expressing emotions, which can hinder an individual's ability to understand and manage their emotions. This study aims to implement and develop a model that can detect emotions using the MobileNetV2 architecture for therapy purposes for individuals experiencing alexithymia. The method uses the FER-2013 dataset, which consists of 35,887 grayscale facial images in 7 emotion categories: anger, disgust, fear, happiness, neutral, sadness, and surprise. Using a deep learning approach based on CRISP-DM, the research begins with normalization and data augmentation to improve the model's resilience to image variations. The developed model achieved a training accuracy of 67.7% and a validation accuracy of 65.3%, demonstrating significant capability in recognizing and classifying emotions from facial images. Evaluation using a confusion matrix showed that the model produced a precision of 64.9%, a recall of 65.4%, and an F1-score of 63.7% for each emotion class. This research implies the potential for developing systems that can support psychological therapy, especially to help individuals with alexithymia understand and manage their emotions through facial expression analysis, providing technology sensitive to emotional expressions.

Keywords: Alexithymia; Deep Learning; Emotion Detection; FER-2013; MobileNetV2

1. Pendahuluan

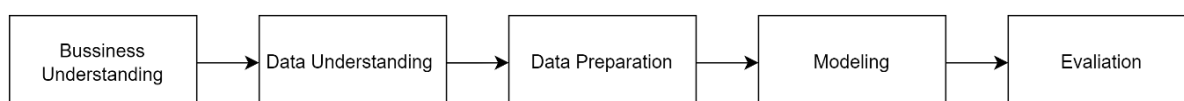
Beberapa penelitian di Indonesia menunjukkan adanya data dari masyarakat Indonesia yang menunjukkan prevalensi alexithymia. Penelitian yang dilakukan di Malang pada tahun 2016 menunjukkan bahwa 70 dari 150 remaja berusia 18-23 tahun memiliki tingkat alexithymia yang tinggi [1]. Selanjutnya, sebuah penelitian pada tahun 2018 menemukan bahwa 32,2% dari 215 remaja berusia 18-22 tahun mengalami alexithymia [2]. Pada tahun 2020, sebanyak 207 remaja di Jakarta Selatan yang berusia 13-19 tahun juga mengalami alexithymia [3]. Penelitian lain di Panti Asuhan Kabupaten Tegal pada tahun 2021 menunjukkan bahwa responden memiliki

alexithymia sedang dengan nilai rata-rata 58,39. [4]. *Alexithymia* merupakan kondisi psikologis yang ditandai dengan ketidakmampuan individu untuk mengenali, memahami, dan mengungkapkan emosi mereka secara verbal. Ciri-ciri utamanya mencakup kesulitan dalam interaksi sosial, kurangnya empati, dan tantangan dalam memahami diri sendiri serta orang lain. Faktor *psikogenik*, seperti gaya *kelekatan* dan pengalaman trauma, faktor *biogenik*, seperti paparan zat-zat dan struktur otak, serta faktor *alexithymia* primer dan sekunder, seperti stres, kecemasan, dan kelelahan, dapat menjadi penyebab *alexithymia* [5]. Pada akhirnya, *alexithymia* dapat memengaruhi kualitas hubungan intrapersonal dan interpersonal. Mereka yang mengalami *alexithymia* mungkin menunjukkan perilaku atau sikap tak acuh yang menghalangi mereka untuk berperilaku prososial. Meskipun penanganan *alexithymia* biasanya melibatkan tahapan terapi, belum ada terapi baru yang didukung oleh bukti yang berhasil dikembangkan untuk meningkatkan kualitas hidup bagi individu yang mengalaminya. Penanganan *alexithymia* dapat dianggap sebagai pelengkap terapi klinis yang sudah ada, sehingga individu dengan tingkat *alexithymia* yang tinggi dapat mendapatkan manfaat lebih dari psikoterapi [6]. Orang biasanya membagi emosi mereka menjadi dua kategori utama: emosi positif dan negatif. Emosi positif, juga dikenal sebagai efek positif, mendorong perasaan menyenangkan dan menenangkan. Contoh dari emosi positif termasuk ketenangan, rasa santai, kebahagiaan, hiburan, kehangatan emosional, dan kegembiraan. Pengalaman emosi positif ini sering dikaitkan dengan kondisi psikologis yang positif. Di sisi lain, emosi negatif, atau efek negatif, menciptakan perasaan tidak nyaman dan mengganggu. Saat mengalami emosi negatif, dampaknya biasanya tidak menyenangkan dan mengganggu. Beberapa contoh emosi negatif termasuk kesedihan, kekecewaan, putus asa, depresi, rasa tidak berdaya, frustrasi, kemarahan, dan dendam, serta lainnya [7]. Dalam bidang kesehatan mental dan fisik, emosi negatif yang tidak dikelola dengan baik dapat menyebabkan peningkatan tingkat stres, yang pada gilirannya dapat mengurangi sistem kekebalan tubuh dan meningkatkan risiko penyakit. Di sisi lain, emosi positif dipahami sebagai tanggapan terhadap situasi atau stimulus yang menggembirakan. Memahami bagaimana emosi diekspresikan menjadi penting karena dapat memberikan petunjuk tentang kondisi psikologis seseorang dan potensi risiko kesehatan yang mungkin timbul [8]. Wajah bukan hanya sebagai alat identifikasi seseorang, tetapi juga sebagai media untuk menyampaikan informasi tentang keadaan emosional individu. Ekspresi wajah menunjukkan interpretasi dari apa yang diucapkan serta mengontrol alur percakapan. Kemampuan manusia untuk menginterpretasi emosi sangat penting untuk komunikasi yang efektif [9]. Ekspresi emosional adalah reaksi spontan individu terhadap situasi, yang dapat berasal dari internal maupun eksternal. Hal ini tercermin melalui perubahan biologis, fisiologis, serta sikap dan perilaku [10]. Ekspresi emosional pada wajah manusia memberikan gambaran yang penting untuk membuat penilaian sosial. Dalam pengenalan ekspresi, wajah berperan lebih dominan dibandingkan dengan suara atau gestur tubuh lainnya [11]. Meskipun ekspresi emosional singkat, pengenalan dan kategorisasinya kompleks, membutuhkan sensitivitas terhadap detail-detail subtan dalam interval waktu tertentu. Ekspresi emosional adalah aspek penting yang memengaruhi kesejahteraan dan kesehatan individu. *Macro* dan *micro* adalah dua jenis ekspresi wajah yang dikenal dalam ilmu psikologi. *Macro expression* adalah ekspresi emosional yang tampak jelas dan mudah diamati oleh orang di sekitarnya. Ini melibatkan perubahan ekspresi wajah, bahasa tubuh, atau suara yang mencolok dan mudah dikenali. Sedangkan, *micro expression* adalah ekspresi emosional yang lebih kecil, tidak disadari, dan terjadi dengan cepat. Karena itu, ekspresi mikro seringkali sulit diamati [12]. Gerakan otot di bawah kulit wajah menghasilkan ekspresi yang dikendalikan oleh sistem saraf otak. Sebagai contoh, senyum mencerminkan kebahagiaan dengan gerakan otot pipi dan bibir ke atas. Sebaliknya, wajah sedih, terindikasi karena mata tidak fokus, bibir tertarik ke bawah, dan kelopak atas terkulai [13]. Selama bertahun-tahun, terjadi kemajuan besar dalam bidang kecerdasan buatan, terutama dalam bagian utamanya, *machine learning*. *Deep learning*, metode yang lebih kompleks tetapi lebih canggih, telah meningkatkan *machine learning*. Keunggulan utamanya adalah tingkat akurasi yang lebih tinggi, khususnya dalam pengenalan objek atau aktivitas. Saat ini, pengembangan model deteksi emosi melalui analisis citra menggunakan *deep learning* sedang dilakukan. Di masa depan,

model ini diharapkan dapat diintegrasikan ke dalam aplikasi mobile untuk membantu individu dengan *alexithymia* dalam terapi dan memenuhi kebutuhan mereka untuk berkomunikasi [14]. Terdapat penelitian yang mengusulkan penggunaan deep learning untuk Facial Emotion Recognition (FER) dalam konteks perawatan kesehatan pengemudi dengan pendekatan transfer learning dan penerapan SqueezeNet 1.1 untuk pengenalan emosi wajah terkait keselamatan pengemudi. Dengan teknik preprocessing data seperti resizing gambar dan augmentasi data, penelitian ini berhasil mencapai akurasi maksimum sebesar 95,83% pada basis data KMU-FED, serta menunjukkan kinerja yang sebanding dengan basis data benchmark lainnya seperti CK+, KDEF, dan FER2013 [15]. Penelitian lain mengusulkan penggunaan deep learning untuk Facial Emotion Recognition (FER) pada anak-anak dengan autisme, yang sering mengalami kesulitan dalam mengekspresikan emosi melalui ekspresi wajah. Model MobileNetV2 diterapkan untuk mengenali emosi, dengan hasil yang menunjukkan performa terbaik pada kelas "Marah" (F1-Score 90%), diikuti oleh "Bahagia" dan "Terkejut". Namun, terdapat ketidakseimbangan pada kelas "Sedih" dan "Takut", serta performa yang lebih rendah pada kelas "Netral" [16]. Dalam penelitian ini, digunakan metode deep learning dengan model arsitektur MobileNetV2 yang telah dilatih sebelumnya untuk mendeteksi emosi. Dalam penelitian ini, digunakan metode deep learning dengan model arsitektur MobileNetV2 yang telah dilatih sebelumnya untuk mendeteksi emosi. MobileNetV2 merupakan jenis algoritma Convolutional Neural Network (CNN) yang dikembangkan dari Multi-Layer Perceptron (MLP) dan dioptimalkan untuk pengolahan data dalam bentuk dua dimensi. Algoritma ini memiliki struktur jaringan yang mendalam dan sering digunakan dalam analisis gambar [17]. Tujuan dari penelitian ini adalah mengembangkan model yang diharapkan dapat mencapai tingkat akurasi dan efisiensi yang lebih tinggi, sehingga dapat menjadi dasar untuk pengembangan sistem terapi bagi pengidap alexithymia melalui analisis citra wajah. Meskipun fokus saat ini adalah pada pengembangan model, diharapkan model ini dapat ditingkatkan di masa depan dan diadaptasi untuk perangkat mobile sebagai alat terapi. MobileNetV2 dipilih karena desainnya yang ringan dan efisien, menjadikannya cocok untuk perangkat seluler dengan sumber daya terbatas tanpa mengorbankan kinerja dalam analisis emosi [16].

2. Metode Penelitian

Sebuah standar industri untuk pengembangan data mining, *CRISP-DM* (*Cross Industry Standard Process for Data Mining*), digunakan sebagai metodologi penelitian [18].



Gambar 1. Metode CRISP-DM








Business Understanding

Penelitian ini bertujuan untuk mengembangkan sebuah model pengklasifikasi emosi sebagai alat terapi bagi pengidap *alexithymia*. *Alexithymia* adalah kondisi yang menyulitkan mereka dalam mengenali dan mengungkapkan emosi mereka secara verbal. Untuk mencapai tujuan ini, diperlukan data dari tujuh kelas emosi: marah, jijik, takut, senang, netral, sedih, dan terkejut. Data ini akan digunakan untuk membangun model pengklasifikasi emosi menggunakan rancangan *MobileNetV2*. Pola yang telah dibuat akan tersimpan dengan format *.keras*.

Data Understanding

Dataset *FER-2013* (*Facial Expression Recognition 2013*), diumumkan pada *ICML* (*International Conference on Machine Learning*) pada tahun 2013, digunakan dalam penelitian ini [19]. Dataset ini terdiri dari 35.887 citra wajah *grayscale* berukuran 48 x 48 piksel, yang dibagi menjadi tujuh emosi: marah, jijik, takut, senang, netral, sedih, dan terkejut. Dataset ini terstruktur dengan baik, dengan setiap gambar sudah dilabeli menggunakan indeks kelas dari 0 hingga 6.

Tabel 1. Kelas emosi FER-2013

| Label | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|--------|---|---|---|---|---|---|---|
| Gambar |  |  |  |  |  |  |  |
| Emosi | marah | jijik | takut | senang | netrall | sedih | terkejut |
| Jumlah | 4953 | 547 | 5121 | 8989 | 6198 | 6077 | 4002 |

Data Preparation

Dataset *FER-2013* terbagi menjadi dua jenis, yaitu data *train* dan data *test*, dengan proporsi lebih banyak data dialokasikan untuk pelatihan. Sebelum melatih model, setiap gambar harus diberi label untuk mengidentifikasi jenis emosi yang terdapat di dalamnya seperti pada Tabel 1, yang terdiri dari tujuh kategori emosi: marah, jijik, takut, senang, netral, sedih, dan terkejut. Setiap gambar ditempatkan dalam folder sesuai dengan emosi yang tercermin di dalamnya. Setelah persiapan dataset, ukuran gambar yang dimasukkan diubah menjadi 224 x 224 piksel, ukuran umum yang digunakan pada arsitektur *MobileNetV2* yang telah dilatih sebelumnya. Nilai piksel gambar ini berada dalam rentang [0, 255]. Untuk mempersiapkan data untuk model *deep learning*, normalisasi dilakukan dengan membagi setiap nilai piksel dengan 255.0, sehingga nilai piksel berada dalam rentang [0, 1]. Selanjutnya, kumpulan gambar yang terdiri dari 35.887 gambar dibagi menjadi dua bagian sesuai dengan metode pembagian.

Tabel 2. Pembagian data

| Data Kategori | Jumlah | Persentase |
|---------------|--------|------------|
| <i>Train</i> | 28709 | 79.99% |
| <i>Test</i> | 7178 | 20.01% |

Augmentasi data adalah metode yang digunakan untuk meningkatkan jumlah citra dalam dataset dengan melakukan transformasi pada citra asli. Penggunaan *augmentasi data* dapat meningkatkan akurasi model, terutama untuk kelas-kelas yang sering salah diklasifikasikan.[20]. Transformasi umum yang biasanya digunakan meliputi *flipping*, *rotating*, *scaling*, dan *cropping* [21]. Dalam penelitian ini, beberapa transformasi yang diterapkan adalah sebagai berikut:

1. *Rotation* untuk merotasi citra secara acak.
2. *Shear* untuk mencukur citra secara acak.
3. *Width shift* untuk menggeser citra secara acak.
4. *Height shift* untuk menggeser citra secara acak.
5. *Zoom* untuk memperbesar secara acak.
6. *Horizontal flip* untuk membalik citra.

Modeling

Setelah tahapan pengolahan awal data selesai, langkah selanjutnya adalah mengembangkan model menggunakan *MobileNetV2* yang telah di-*train* sebelumnya. *MobileNet*, dikembangkan oleh Google, dirancang untuk menjadi ringan dan efisien, terutama untuk perangkat dengan sumber daya terbatas. Salah satu fitur utama arsitektur ini adalah penggunaan *depthwise separable convolution*. *Depthwise separable convolution* adalah teknik yang melakukan pembagian metode *convolution* menjadi 2 fase yaitu konvolusi mendalam dan konvolusi titik. Konvolusi mendalam melakukan filter per saluran, sedangkan konvolusi titik melakukan konvolusi 1x1 untuk menggabungkan keluaran dari konvolusi mendalam. Pendekatan ini secara signifikan mengurangi jumlah parameter dan komputasi yang terlibat jika dibandingkan dengan konvolusi biasa, menghasilkan model yang lebih efisien tanpa mengorbankan akurasi. Untuk penelitian ini, arsitektur *MobileNetV2* diadaptasi dan dijelaskan lebih rinci dalam Tabel 3, yang menguraikan konfigurasi lapisan dan parameter yang digunakan untuk membangun model pengklasifikasi emosi.

Tabel 3. Arsitektur MobileNetV2

| Input Size | Operator | t | c | n | s |
|------------|--------------------------|---|------|---|---|
| 224×224×1 | Conv2D | - | 3 | 1 | 1 |
| 224×224×3 | BatchNormalization | - | - | 1 | - |
| 224×224×3 | MobileNetV2 (Conv2D) | - | 32 | 1 | 2 |
| 112×112×32 | MobileNetV2 (Bottleneck) | 1 | 16 | 1 | 1 |
| 112×112×16 | MobileNetV2 (Bottleneck) | 6 | 24 | 2 | 2 |
| 56×56×24 | MobileNetV2 (Bottleneck) | 6 | 32 | 3 | 2 |
| 28×28×32 | MobileNetV2 (Bottleneck) | 6 | 64 | 4 | 2 |
| 14×14×64 | MobileNetV2 (Bottleneck) | 6 | 96 | 3 | 1 |
| 14×14×96 | MobileNetV2 (Bottleneck) | 6 | 160 | 3 | 2 |
| 7×7×160 | MobileNetV2 (Bottleneck) | 6 | 320 | 1 | 1 |
| 7×7×320 | MobileNetV2 (Conv2D 1x1) | - | 1280 | 1 | 1 |
| 7×7×1280 | GlobalAveragePooling2D | - | - | - | - |
| 1280 | Dropout | - | - | - | - |
| 1280 | Dense (Output Layer) | - | 7 | 1 | - |

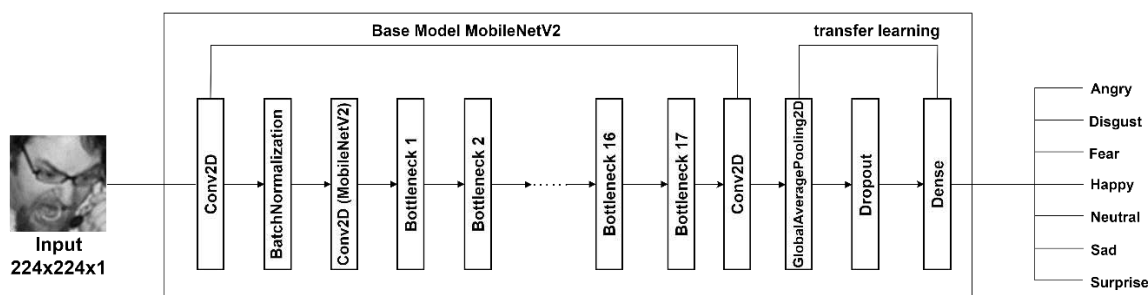
Keterangan:

n adalah jumlah iterasi atau pengulangan yang dilakukan sebanyak **n** kali.

c adalah ukuran saluran output (output channel).

s adalah ukuran langkah (stride) yang digunakan dalam blok.

t adalah faktor ekspansi, yang berarti menambah jumlah saluran (channels).



Gambar 2. Penerapan arsitektur MobileNetV2

Dalam model ini, input awalnya memiliki ukuran 224x224 piksel dengan 1 channel warna grayscale, yang kemudian diubah menjadi 3 channel untuk digunakan dalam *MobileNetV2*. Langkah konvolusi awal menggunakan 3 filter untuk mendapatkan fitur mendasar seperti tekstur dan tepi dari gambar. Hasil konvolusi dinormalkan menggunakan lapisan *BatchNormalization*. Model dilengkapi dengan serangkaian lapisan *bottleneck* yang menggunakan teknik konvolusi terpisah mendalam (*depthwise separable convolution*) dengan jumlah filter yang bervariasi dari 16 hingga 320. Teknik ini membagi konvolusi menjadi *konvolusi terpisah*, yang melakukan konvolusi pada setiap saluran input secara individu, dan konvolusi titik yang menyatukan outputnya. Setelah melewati lapisan *bottleneck*, dilakukan konvolusi 1x1 dengan 1280 filter untuk penyesuaian fitur sebelum *average pooling* dengan kernel 7x7, yang menurunkan dimensi spasial dari fitur map. Tahap akhir melibatkan konvolusi 1x1 untuk penggabungan fitur yang disederhanakan oleh *average pooling*, diikuti oleh lapisan *softmax* dan *fully connected* untuk klasifikasi peluang. Peneliti menggunakan *MobileNetV2* sebagai *base model* dengan *transfer learning* untuk mengubah *classifier*-nya, menambahkan *Global Max Pooling 2D*, lapisan *Dense* dengan 256 neuron serta *ReLU* sebagai fungsi aktivasi, serta *dropout layer* untuk mengurangi overfitting. Output model ini adalah probabilitas kelas emosi, yang menentukan prediksi berdasarkan nilai probabilitas tertinggi dari *softmax*.

Evaluasi

Dilakukan pengujian dan evaluasi dengan data uji yang telah disiapkan untuk mengukur kinerja model yang dibuat. Kinerja model klasifikasi menunjukkan seberapa baik model dapat mengkategorikan data ke dalam kelas tertentu. *Confusion matrix* adalah salah satu cara untuk mengukur kinerja model tersebut. Metode yang umum untuk mengevaluasi seberapa baik sistem klasifikasi bekerja adalah akurasi. Tujuan memperhitungkan akurasi adalah mengetahui berapa tingkat keberhasilan algoritma yang menunjukkan probabilitas nilai yang sebenarnya dan label kelas keseluruhan. Singkatnya, akurasi mengukur kinerja keseluruhan dari algoritma [22].

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + False\ Negative + True\ Negative} \quad (1)$$

Keterangan:

True Positive (TP): Kasus di mana data yang benar-benar positif terklasifikasi dengan tepat sebagai positif.

False Positive (FP): Kasus di mana data yang seharusnya negatif salah diklasifikasikan sebagai positif.

True Negative (TN): Kasus di mana data yang benar-benar negatif terklasifikasi dengan tepat sebagai negatif.

False Negative (FN): Kasus di mana data yang seharusnya positif salah diklasifikasikan sebagai negatif.

Precision adalah persentase dari prediksi yang relevan dari hasil klasifikasi. *Recall*, di sisi lain, adalah persentase dari total hasil yang relevan yang berhasil diklasifikasikan dengan benar.

$$Precision = \frac{True\ Positive}{True\ Positif + False\ Positif} \quad (2)$$

Recall adalah nilai yang mencerminkan seberapa banyak sampel positif yang berhasil diklasifikasikan dengan benar dibandingkan dengan jumlah keseluruhan sampel positif, sesuai dengan rumus yang diberikan.

$$Recall = \frac{True\ Positive}{True\ Positif + False\ Negatif} \quad (3)$$

F1-Score adalah rata-rata harmonik antara *precision* dan *recall*.

$$F1 - Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \quad (4)$$

3. Hasil

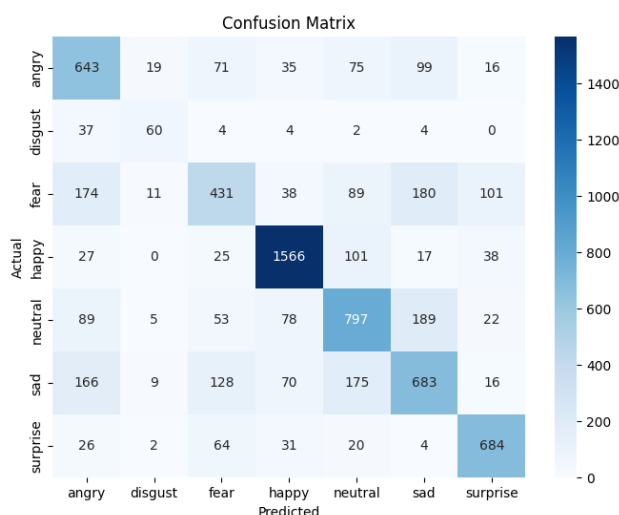
Transformasi citra yang digunakan dalam pelatihan model deteksi emosi menggunakan arsitektur *MobileNetV2* melibatkan berbagai teknik augmentasi untuk meningkatkan ketahanan dan generalisasi model. Transformasi tersebut termasuk rotasi acak hingga 20 derajat, yang membantu model belajar mengenali emosi dari citra yang tidak terorientasi secara standar. *Shear* citra dengan faktor 0.2 digunakan untuk meniru efek perspektif dan distorsi. Penggeseran citra secara horizontal sebesar 20% dari lebar citra dan vertikal sebesar 20% dari tinggi citra membantu model untuk belajar mengenali emosi bahkan ketika subjek tidak berada di pusat frame. Zoom acak hingga 20% membantu model memfokuskan dan memahami fitur penting dalam berbagai ukuran citra.



Gambar 3. Hasil augmentasi data

Gambar ini memperlihatkan citra asli dan beberapa citra hasil *augmentasi*, yang masing-masing menunjukkan efek dari transformasi yang diterapkan. Transformasi ini bertujuan untuk membuat model lebih tahan terhadap variasi dan generalisasi lebih baik saat mendeteksi emosi dari citra wajah yang berbeda.

Confusion Matrix digunakan untuk mengevaluasi kinerja sebuah model dengan memperhitungkan akurasi, *precision*, *recall*, dan *F1-score*. Matrix ini mengandung empat kemungkinan kombinasi hasil prediksi dan nilai aktual dalam masalah klasifikasi [23]. Seberapa baik model mengklasifikasikan setiap kelas ditunjukkan oleh Confusion Matrix pada Gambar 5.



Gambar 4. Confusion Matrix

Berdasarkan Gambar 4, terdapat jumlah data yang sesuai untuk setiap kategori emosi ialah 643 marah, 60 jijik, 431 takut, 1566 senang, 797 netral, 683 sedih, dan 684 terkejut. Meskipun demikian, terdapat kesalahan prediksi emosi yang signifikan karena jumlah data yang diuji cukup besar.

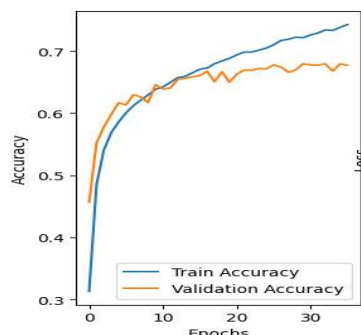
Tabel 4 menunjukkan hasil percobaan terbaik dengan epoch ke-36, batch size bernilai 32, learning rate 0,001, dan optimizer *Adam*. Akurasi pelatihan mencapai 67.7% dari keseluruhan data uji, sedangkan akurasi validasisebesar 65.3% dari keseluruhan data latih.

Tabel 4. Disajikan clasification report

| Expression | Precision | Recall | F1-Score |
|---------------------|-----------|--------|----------|
| Marah | 0.553 | 0.671 | 0.607 |
| Jijik | 0.566 | 0.541 | 0.553 |
| Takut | 0.555 | 0.421 | 0.479 |
| Senang | 0.860 | 0.883 | 0.871 |
| Netral | 0.633 | 0.646 | 0.640 |
| Sedih | 0.581 | 0.548 | 0.564 |
| Terkejut | 0.780 | 0.823 | 0.801 |
| Macro Average | 0.649 | 0.654 | 0.637 |
| Train Accuracy | 0.677 | | |
| Validation Accuracy | 0.653 | | |

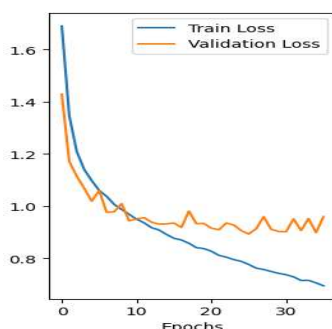
4. Pembahasan

Setelah dilakukan *transformasi data* dan pelatihan model deteksi emosi menggunakan arsitektur *MobileNetV2*, hasil yang dihasilkan biasanya mencakup empat metrik utama yaitu *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss*.



Gambar 5. Akurasi train dan validasi

Grafik akurasi pada Gambar 4 menunjukkan bahwa akurasi pelatihan meningkat dari sekitar 0,3 pada epoch pertama hingga lebih dari 0,75 pada epoch ke-30, menandakan bahwa model semakin efektif dalam mempelajari pola dari data pelatihan. Akurasi validasi juga meningkat signifikan di awal, menunjukkan kemampuan model untuk menggeneralisasi dengan baik. Namun, setelah epoch ke-10, akurasi validasi cenderung stabil di sekitar 0,65, mengindikasikan bahwa model mendekati batas kinerja optimal. Pelatihan berakhir pada epoch ke-36, dengan akurasi pelatihan 0,742 dan akurasi validasi 0,677.



Gambar 6. Loss train dan validasi

Pada Gambar 5, grafik loss menunjukkan penurunan tajam pada *train loss* dari sekitar 1,6 pada epoch pertama hingga mendekati 0,6 pada epoch ke-30, menandakan peningkatan kemampuan model dalam memprediksi data pelatihan. *Validation loss* juga menurun signifikan di awal, namun stabil di sekitar 1,0 setelah epoch ke-10, menunjukkan bahwa model mendekati batas kemampuannya dalam generalisasi. Pelatihan berakhir pada epoch ke-36, dengan *train loss* 0,694 dan *validation loss* 0,959.

5. Penutup

Hasil model menunjukkan bahwa model yang dilatih menggunakan *MobileNetV2* menunjukkan performa yang cukup baik. Model mendapatkan *train accuracy* sebesar 67,7% dan *validation accuracy* sebesar 65,3%. Evaluasi tambahan menggunakan *confusion matrix* menunjukkan nilai *precision* 64,9%, *recall* 65,4%, dan *F1-score* 63,7%, yang perlu diperinci untuk setiap kelas emosi. Model ini menggambarkan potensi yang baik dalam deteksi emosi untuk aplikasi mobile, terutama sebagai alat pendukung terapi bagi pengidap *alexithymia*. Namun, untuk penerapan lebih lanjut, disarankan untuk melakukan penelitian selanjutnya guna meningkatkan akurasi dan

validitas model agar dapat memaksimalkan hasilnya dalam aplikasi mobile sebagai alat terapi yang efektif.

6. Referensi

- [1] L. Lestari, "Pengaruh Kecenderungan Alexithymia Terhadap Kecemburuan Dalam Hubungan Berpacaran," 2016. [Online]. Available: <https://api.semanticscholar.org/CorpusID:150522190>
- [2] T. W. Harjanah, "Hubungan Antara Level Alexithymia Dengan Perilaku Prososial Dewasa Muda."
- [3] Y. M. Lestari, S. Y. Dewi, and A. Chairani, "SCRIPTA SCORE Scientific Medical Journal Penelitian Hubungan Alexithymia dengan Kecanduan Media Sosial pada Remaja di Jakarta Selatan."
- [4] K. Emosional dan Kecenderungan Alexithymia pada Remaja yang tinggal di Panti Asuhan, S. Ayu Novita, T. Suprihatin, A. Fitriani, F. Psikologi, and U. Islam Sultan Agung, "Dipresentasikan dalam Seminar Nasional dan Call for Paper "Penguatan Kesehatan Mental di Masa Pandemi," 2021.
- [5] I. Gusti, A. Agung, M. Pradnyadewi, and P. N. Wideasavitri, "Faktor-faktor Penyebab Alexithymia pada Remaja: Literature Review."
- [6] D. Samur, M. Tops, C. Schlinkert, M. Quirin, P. Cuijpers, and S. L. Koole, "Four decades of research on alexithymia: Moving toward clinical applications," *Front Psychol*, vol. 4, no. NOV, 2013, doi: 10.3389/fpsyg.2013.00861.
- [7] G. Istiqomah, D. Wahyuni, P. Islam Anak Usia Dini, F. Ilmu Tarbiyah, and U. Raden Fatah Palembang, "Pengenalan Emosi Positif dan Emosi Negatif Pada Anak Usia Dini," *Jurnal Ilmiah Multidisiplin*, vol. 3, no. 1, 2023.
- [8] A. Larasati, Z. Fahma Auliya, F. Ekonomi, and D. Bisnis, "The influence of the store environment characteristics toward impulse buying, positive emotional response of consumers as a mediation," 2023.
- [9] M. F. Ali and M. Khatun, "Facial Emotion Detection Using Neural Network." [Online]. Available: <http://www.ijser.org>
- [10] A. B. Jala, T. W. Purboyo, and R. A. Nugrahaeni, "Implementation of Convolutional Neural Network (CNN) Algorithm for Classification of Human Facial Expression in Indonesia," in *2020 International Conference on Information Technology Systems and Innovation, ICITSI 2020 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020, pp. 256–262. doi: 10.1109/ICITSI50517.2020.9264940.
- [11] G. B. Vasani, P. R. S. Senjaliya, P. V Kathiriyai, A. J. Thesiya, and H. H. Joshi, "Human Emotional State Recognition Using Facial Expression Detection," 2013.
- [12] "1712. Gesture - Mengungkap Makna Dibalik Bahasa Tubuh Orang Lain - Zaka Putra Ramdani - The Blackdof (1)[1]"
- [13] L. Zahara, P. Musa, E. Prasetyo Wibowo, I. Karim, and S. Bahri Musa, "The Facial Emotion Recognition (FER-2013) Dataset for Prediction System of Micro-Expressions Face Using the Convolutional Neural Network (CNN) Algorithm based Raspberry Pi," in *2020 5th International Conference on Informatics and Computing, ICIC 2020*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020. doi: 10.1109/ICIC50835.2020.9288560.
- [14] C. R. W. Morin and W. Rahardjo, "Kecemasan Sosial, Kecenderungan Alexithymia Dan Adiksi Internet Pada Mahasiswa," *Jurnal Psikologi*, vol. 14, no. 1, pp. 11–24, 2021, doi: 10.35760/psi.2021.v14i1.3439.
- [15] T. I. Hermanto and Y. Muhyidin, "Analisis Data Sebaran Bandwidth Menggunakan Algoritma DbSCAN Untuk Menentukan Tingkat Kebutuhan Bandwidth Di Kabupaten Purwakarta," *Rabit : Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 5, no. 2, pp. 130–137, Jul. 2020, doi: 10.36341/rabit.v5i2.1388.
- [16] F. B. Hoirot, "Implementasi Mobilenetv2 Untuk Mendeteksi Emosi Pada Anak Autis Melalui Ekspresi Wajah Berbasis Mobile," Politeknik Negeri Jember, 2024.

- [17] A. A. Handoko, M. A. Rosid, and U. Indahyanti, "Implementasi Convolutional Neural Network (CNN) Untuk Pengenalan Tulisan Tangan Aksara Bima," *SMATIKA JURNAL*, vol. 14, no. 01, pp. 96–110, Jul. 2024, doi: 10.32664/smatika.v14i01.1196.
- [18] T. I. Hermanto and Y. Muhyidin, "Analisis Data Sebaran Bandwidth Menggunakan Algoritma Dbscan Untuk Menentukan Tingkat Kebutuhan Bandwidth Di Kabupaten Purwakarta," *Rabit : Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 5, no. 2, pp. 130–137, Jul. 2020, doi: 10.36341/rabit.v5i2.1388.
- [19] I. J. Goodfellow *et al.*, "Challenges in representation learning: A report on three machine learning contests," *Neural Networks*, vol. 64, pp. 59–63, Apr. 2015, doi: 10.1016/j.neunet.2014.09.005.
- [20] O. A. Shawky, A. Hagag, E. S. A. El-Dahshan, and M. A. Ismail, "Remote sensing image scene classification using CNN-MLP with data augmentation," *Optik (Stuttg)*, vol. 221, Nov. 2020, doi: 10.1016/j.ijleo.2020.165356.
- [21] A. M. Saleh and T. Hamoud, "Analysis and best parameters selection for person recognition based on gait model using CNN algorithm and image augmentation," *J Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-020-00387-6.
- [22] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond Accuracy, F-score, and ROC: A Family of Discriminant Measures for Performance Evaluation," 2006. [Online]. Available: www.aaii.org
- [23] D. Safitri, S. S. Hilabi, and F. Nurapriani, "Analisis Penggunaan Algoritma Klasifikasi Dalam Prediksi Kelulusan Menggunakan Orange Data Mining," *Rabit : Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 8, no. 1, pp. 75–81, Jan. 2023, doi: 10.36341/rabit.v8i1.3009.