

## Perancangan Deteksi Emosi Manusia berdasarkan Ekspresi Wajah Menggunakan Algoritma VGG16

Dimas Setiawan<sup>1\*</sup>, Suprih Widodo<sup>2</sup>, Taufik Ridwan<sup>3</sup>, Rifqi Ambari<sup>4</sup>

<sup>1,2</sup>Program Studi Pendidikan Sistem dan Teknologi Informasi, Universitas Pendidikan Indonesia

Jl. Veteran 08, Purwakarta

Email: <sup>1\*</sup>dimassetiawan@upi.edu, <sup>2\*</sup>supri@upi.edu,

<sup>3\*</sup>taufikridwan@upi.edu, <sup>4</sup>rifqi4mbari@upi.edu

**Abstrak.** *Machine Learning* merupakan bagian dari kecerdasan buatan yang banyak digunakan untuk memecahkan berbagai masalah khususnya di domain *Computer Vision*. Salah satu kajian di bidang ini adalah deteksi emosi yang menjadi salah satu kajian yang paling kompleks dan menantang. Besarnya variasi yang disebabkan oleh perubahan dalam penampilan wajah, pencahayaan dan ekspresi wajah itu sendiri. Dalam penelitian ini, riset yang dilakukan adalah mempelajari lebih lanjut perihal deteksi emosi manusia berdasarkan ekspresi wajah menggunakan Algoritma *Convolutional Neural Network* dimana data yang digunakan untuk proses training data berupa sampel citra dari dataset FER 2013. Metode pengembangan perangkat lunak menggunakan framework AI Project Life Cycle. Arsitektur dari *Convolutional Neural Network* yang akan digunakan yaitu VGG 16 dengan pendeteksian wajah menggunakan Metode MTCNN.

**Kata kunci:** Deteksi Emosi, Ekspresi Wajah, VGG 16, MTCNN

### 1 Pendahuluan

*Machine Learning* dapat didefinisikan sebagai aplikasi komputer dan algoritma matematika yang diadopsi dengan cara pembelajaran yang berasal dari data dan menghasilkan prediksi di masa yang akan datang. Selain itu, proses pembelajaran yang dimaksud adalah suatu usaha dalam memperoleh kecerdasan yang melalui dua tahap yakni tahap training dan testing[1].

Kemajuan di bidang *Machine Learning* termasuk pengolahan citra digital membentuk banyak penemuan baru dalam suatu sistem, secara tak eksklusif teknologi pengolahan citra mulai menggeser sistem-sistem yg masih memakai teknologi usang. tidak hanya teknologi pengolahan gambaran saja yg berkembang tetapi metode-metode pengenalan juga ikut berkembang seperti misalnya adalah sistem yang menggunakan sosialisasi wajah menjadi sistem keamanan serta pemantauan, termasuk dalam hal yang berkaitan dengan penguatan ketepatan

sistem pada mengenali paras atau ekspresi pada wajah manusia, sebagai akibatnya wajah masih menjadi objek yang terus menerus dikembangkan peneliti.

Wajah adalah sebuah objek yang mana dalam mengidentifikasi seorang. dari paras atau wajah manusia yang dalam hal ini dapat menunjukkan aktualisasi diri atau emosi mereka pada berinteraksi sosial, dalam interaksi tersebut, syarat emosional yang baik sebagai hal penting pada menjalin komunikasi yang baik. ekspresi atau mimik paras merupakan bentuk komunikasi nonverbal yang mana dalam ekspresi tersebut dapat memberikan keadaan emosi dari seseorang kepada orang yang mengamatinya[2]

Ekspresi emosi merupakan usaha yang dilakukan oleh seseorang untuk mengkomunikasikan status perasaan (emosi) sebagai respon terhadap situasi tertentu baik internal maupun eksternal yang terlihat dari perubahan biologis, fisiologis dan serangkaian kecenderungan tindakan (sikap dan tingkah laku) berorientasi pada tujuan[3]. Pengenalan emosi seseorang dapat membantu dalam menentukan kegiatan yang membangkitkan suasana salah satunya ialah dalam kegiatan belajar mengajar[4]. Kestabilan emosi penting bagi keberhasilan proses belajar mengajar, karena emosi mempengaruhi siswa dalam memusatkan perhatian dan menggunakan pikirannya dengan baik. Perlu diketahui bahwa Domain *Computer Vision* khususnya dalam pendeteksian emosi berdasarkan ekspresi wajah data yang didapatkan berupa data yang sudah ada, karena dalam *Machine Learning* cara kerjanya mesin membaca dari data yang diberikan untuk belajar dan seperti manusia dari data yang diberikan memberikan informasi yang sebelumnya telah diberikan[5].

Azhari dan Fitriyani (2020)) dalam penelitiannya mengatakan Penerapan *Deep Learning* dengan metode *Convolutional Neural Network* untuk pendeteksian emosi berdasarkan ekspresi wajah secara realtime ini menemui hasil yang memuaskan namun disisi lain memakan resource hardware yang cukup besar untuk prediksi nilai emosi. Dataset yang digunakan dalam penelitian ini bersumber dari kaggle FER 2013. Hasil dari pengkombinasian antara metode CNN, *Viola-Jones* dan *Haar Cascade* dengan tingkat keakuratan yang terbilang cukup bagus antara rentang 67-83% [6].

Omar, dkk. (2021) menyimpulkan bahwa pada penelitian ini, model CNN di training menggunakan library tensorflow-gpu *version 2.3.0*. Dataset dari training yang digunakan adalah dataset FER-2013 yang didapatkan dari kaggle. Hasil yang didapatkan, bahwa model berhasil untuk melakukan klasifikasi ekspresi wajah manusia. Hal ini didukung dengan akurasi sebesar 73,56%. Dengan menggunakan algoritma Haar Cascade. Model CNN yang di training mencapai akurasi sebesar 73.56%. Selain itu, sistem juga berhasil untuk melakukan pengenalan berbagai ekspresi wajah partisipan terhadap input gambar yang berupa gambar screenshot dari virtual meeting zoom. Grafik kurva dari training dan validation juga menunjukkan bahwa model CNN yang terdapat pada penelitian ini tidak terjadi overfitting [7].

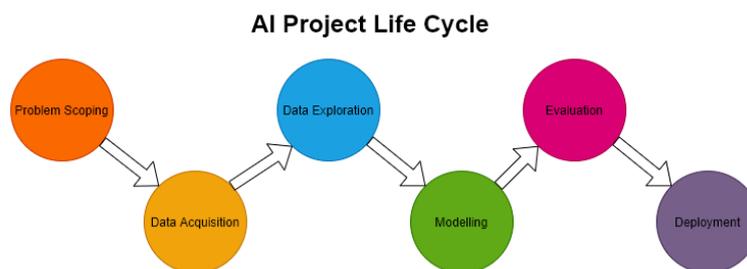
Seandrio, dkk. (2021) dalam penelitiannya menyimpulkan penelitian yang telah dilakukan, banyaknya lapisan pada lapisan konvolusi tidak menjadi jaminan baik dan buruknya hasil akurasi yang diperoleh. Arsitektur *Convolutional Neural Network* menggunakan model *Visual Group Geometri* (VGG) dengan 16 lapisan, terbukti dapat menangani data emosi wajah yang diklasifikasi. Hal itu dapat dibuktikan dengan pengujian *confusion matrix* yang menghasilkan nilai rata-rata akurasi sebesar 70,47%, presisi rata-rata sebesar 70,86%, dan recall rata-rata sebesar 69,57%. Pengujian dilakukan dengan menggunakan data uji sebanyak 3589 dataset, yang terdiri dari 491 data marah, 55 data jijik, 528 data takut, 879 data senang, 594 data sedih, 416 data kaget, dan 626 data biasa. Hasil klasifikasi yang kurang tepat pada penelitian ini dipengaruhi oleh faktor tertentu seperti buram pada kamera dan kemiripan antar emosi[5].

Pratama, dkk (2019) mengemukakan bahwa dalam hasil pengujian secara otomatis menggunakan berbagai pola variasi dalam kumpulan data gambar Aberdeen, pengujian hingga 60 gambar dari 30 orang berbeda yang digunakan dalam Sistem Pencarian Pengenalan Wajah menggunakan *Cascade Multitasking Convolutional Neural Network* (MTCNN) menghasilkan kesimpulan, beberapa parameter citra yang mempengaruhi seperti variasi kecerahan (gelap, terang, sedang), variasi posisi objek (antara lain dapat diambil dekat dengan wajah, dari jarak dengan kemiringan yang berbeda), kemudian posisi grip (sedikit miring, ekspresi wajah dalam gambar objek yang sangat berbeda (marah, sedih, kesal, tersenyum, datar, cemberut), sistem pencarian pengenalan wajah menggunakan MTCNN (*Multi-Task Cascaded Convolutional Neural Network*) berhasil melakukan pengenalan wajah 100%. [8].

Berdasarkan penelitian terdahulu, penulis melakukan perancangan dengan menggabungkan keduanya Arsitektur VGG 16 dengan MTCNN, metode algoritma *Convolutional Neural Network* dengan model arsitektur VGG 16 sebagai model arsitektur dalam Perancangan Deteksi Emosi Manusia berdasarkan Ekspresi Wajah. Selain itu akan menggunakan pengenalan wajah dengan metode *Multi-Task Convolutional Neural Network* (MTCNN). Algoritma ini dipilih karena setelah melihat hasil penelitian-penelitian terdahulu ternyata algoritma ini lebih efektif untuk digunakan dalam kasus seperti ini karena Arsitektur *Convolutional Neural Network* menggunakan model *Visual Group Geometri* (VGG) dengan 16 lapisan, kemudian Pendeteksian wajah dengan metode MTCNN berhasil melakukan pengenalan wajah 100% berdasarkan penelitian terdahulu, maka dari itu dengan menggunakan CNN Arsitektur VGG 16 dan pengenalan wajah menggunakan MTCNN terbukti dapat menangani data emosi wajah yang diklasifikasi 100% dalam pengenalan wajah, sehingga dalam pendeteksian menggunakan Algoritma VGG 16 menjadi maksimal, oleh sebab itu akan dilakukan perancangan tersebut.

## 2 Metode Penelitian

Metode penelitian ini menggunakan *Framework AI Project Life Cycle* yang terdiri dari *Problem Scoping*, *Data Acquisition*, *Data Exploration*, *Modelling*, *Evaluation*, dan *Deployment*[9]. Adapun proses dari *framework* yang akan dilaksanakan penulis dimulai dari *Problem scoping* yaitu proses awal dari *AI project life cycle* dalam hal ini terkait lingkup masalah sehingga dengan hal ini masalah cepat diketahui akar permasalahannya dengan menggunakan 4Ws mulai dari *who*, *where*, *what*, dan *why*, *Data Acquisition* yaitu pada tahap ini penulis melaksanakan proses pengukuran, pengumpulan data dan validitas data yang dibuat dalam menyelesaikan masalah tersebut dengan cara bagaimana cara mengetahui data yang dibutuhkan dan bagaimana cara mendapatkannya dalam hal ini mengambil data dari Kaggle yaitu Dataset FER2013, *Data Exploration* yaitu penulis pada tahap ini menindaklanjuti data yang didapat dalam *Data Acquisition*, pada tahap ini memahami karakteristik dari data yang didapatkan dengan preprocessing terkait data yang ada tersebut dengan cara Augmentasi dataset FER 2013, *Modelling* yaitu pada tahap *modelling* ini tahap selanjutnya dari *Data Exploration* yaitu pengembangan model yang meliputi proses pemilihan atau menentukan algoritma, membuat infrastruktur model, pembagian dataset menjadi train, validation dan testing serta menentukan *Hyperparameter*, *Evaluation* yaitu pada tahap *evaluation* ini tahap lanjutan dari model setelah model dapat akan dilakukan evaluasi dengan memperhatikan akurasi dari data tersebut setelah training sudah baik atau belum jika belum akan dilakukan proses training data dan preprocessing ulang[10]. Jika sudah selesai lanjut ke tahap *deployment* dari model yang sudah baik dan *Deployment* Pada metode yang digunakan menggunakan Algoritma CNN dengan arsitektur VGG 16 dengan metode pengenalan wajah menggunakan *Multi-Task Convolutional Neural Network* (MTCNN). Adapun tahapan dalam melakukan penelitian ini menggunakan *Framework AI Project Life Cycle* sebagai berikut



**Gambar 1.** Framework AI Project Life Cycle  
(Sumber: (Dhingra et al., 2020))

Adapun dalam *AI Project Life Cycle* lebih jelasnya penulis sampaikan dalam poin-poin, penulis akan melakukan yaitu

- 1) *Problem Scoping*

Problem scoping adalah proses awal dari AI project life cycle dalam hal ini terkait lingkup masalah sehingga dengan hal ini masalah cepat diketahui akar permasalahannya dengan menggunakan 4Ws mulai dari who, where, what, dan why.

2) *Data Acquisition*

Data Acquisition adalah proses setelah permasalahan dapat dan solusinya, beralih apa yang dibutuhkan terkait penyelesaian yaitu dalam hal ini terkait data, jadi dalam Data Acquisition ini proses pengukuran, pengumpulan data dan validitas data yang dibuat dalam menyelesaikan masalah tersebut dengan cara bagaimana cara mengetahui data yang dibutuhkan dan bagaimana cara mendapatkannya dalam hal ini mengambil data dari Kaggle yaitu dataset FER2013

3) *Data Exploration*

Dalam Data Exploration ini menindaklanjuti data yang didapat dalam Data Acquisition, pada tahap ini memahami karakteristik dari data yang didapatkan dengan *preprocessing* terkait data yang ada tersebut.

4) *Modelling*

Pada tahap modelling ini tahap selanjutnya dari Data Exploration yaitu pengembangan model yang meliputi proses pemilihan algoritma dan training data

5) *Evaluation*

Pada tahap *evaluation* ini tahap lanjutan dari model setelah model dapat akan dilakukan evaluasi dengan memperhatikan akurasi dari data tersebut setelah training sudah baik atau belum jika belum akan dilakukan proses training data dan preprocessing ulang. Jika sudah selesai lanjut ke tahap deployment dari model yang sudah baik tersebut

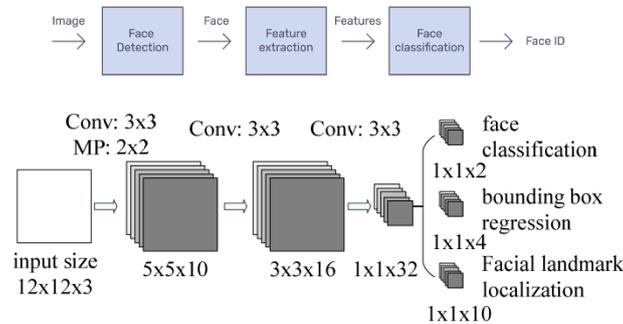
6) *Deployment*

Pada tahap deployment ini tahap lanjutan dari Evaluation setelah didapatkan model baik atau yang siap digunakan tahap deployment ini implementasi dari model yang dibuat dalam hal ini akan deteksi emosi menggunakan ekspresi wajah.

Pada metode ini yang akan dijalankan sampai ke tahap Modelling yang mana data yang diperoleh akan dilakukan preprocessing data yang digunakan yaitu berasal dari kaggle yakni data FER 2013.

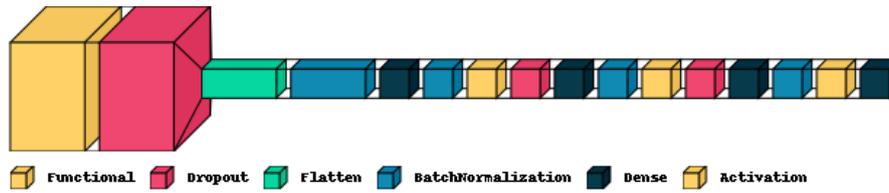
Adapun arsitektur yang digunakan yaitu VGG 16 yang mana VGG 16 ini termasuk dalam arsitektur dari CNN, Secara umum tipe layer CNN terdiri dari *feature learning/extraction* dan *classification layer*. Dimana feature extraction terdiri dari *convolution layer*, *activation* dan *pooling layer*. Sementara *classification layer* terdiri dari *flatten*, *fully connected layer* dan *output layer*[11]. CNN memiliki kinerja yang sangat baik dalam mengatasi masalah pembelajaran mesin (*Machine Learning*) dan deep learning. Khususnya aplikasi yang berhubungan dengan data gambar, seperti kumpulan data klasifikasi gambar

terbesar (*ImageNet*), *Computer Vision Dan Natural Language Processing* kemudahan pendeteksian wajah menggunakan MTCNN yaitu sebuah *library* dalam pengenalan wajah. Adapun alurnya sebagai berikut

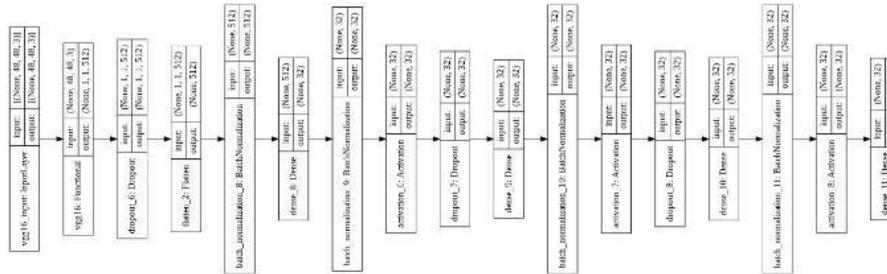


**Gambar 2.** Alur deteksi wajah dengan MTCNN  
(Sumber(Zhang et al., 2016))

Alur diatas merupakan alur bagaimana cara mendeteksi sebuah wajah, pada *library* MTCNN *Feature Extraction* menggunakan *FaceNet* dan penggunaan *activation softmax* untuk klasifikasi wajah[2]. *Multi-task Cascaded Convolutional Neural Network* (MTCNN), MTCNN adalah pengembangan dari CNN yang mana di embed dalam bentuk *Library* MTCNN untuk pendeteksian wajah agar lebih mudah digunakan. Perbedaan CNN dengan MTCNN tidak hanya klasifikasi wajah seperti CNN, tetapi arsitekturnya terdiri dari pendeteksian objek, seperti Algoritma pada YOLO, *Bounding box regression* maksudnya mencari titik koordinat pada wajah tersebut, dan *Facial Landmark Localization* yang berarti 5 titik poin face key point pada wajah yaitu dua mata, hidung dan mulut kanan dan kiri [12] . Cara kerja dari MTCNN dengan mendeteksi garis wajah dari gambar yang disediakan kemudian membuat model jika wajah pengguna dihalangi. MTCNN bekerja secara paralel sehingga dapat memberikan hasil yang lebih cepat. Setelah menggunakan MTCNN, akan dilakukan image post-processing. Image post processing digunakan untuk memberikan bounding box serta memperkecil ukuran gambar yang sudah ada. Selanjutnya *feature extraction* menggunakan *FaceNet* Selanjutnya pada arsitektur yang ada dalam VGG 16. VGG16 adalah arsitektur *Convolutional Neural Network* yang terdiri dari 16 layer. VGG16 merupakan model CNN yang memanfaatkan convolutional layer dengan spesifikasi convolutional filter yang kecil (3x3). Dengan ukuran convolutional filter tersebut, kedalaman neural network dapat ditambah dengan lebih banyak lagi convolutional layer. Hasilnya, model CNN menjadi lebih akurat daripada model-model CNN sebelumnya yang dilakukan sebagai berikut



Gambar 3. Arsitektur VGG 16



Gambar 4 . Arsitektur VGG 16 dengan penambahan layer

Dalam arsitektur VGG 16 terdiri atas *functional*, *Dropout*, *Flatten*, *Batch Normalization*, *Dense*, dan *Activation*. VGG-16 adalah versi arsitektur VGG yang memiliki kedalaman layer sebanyak 16. Pada penelitian dari Simonyan & Zisserman (2015) menunjukkan nilai *error* pada VGG-16 merupakan nilai error yang terkecil dibandingkan versi VGG yang lainnya[13]. Pada penelitian tersebut juga diajukan metode *multiscale training* untuk dibandingkan dengan *single-scale training*. *Single-scale training* adalah mengubah ukuran gambar menjadi suatu ukuran kemudian gambar dipotong untuk disesuaikan dengan layer input. Sedangkan *multi-scale training* adalah mengubah ukuran gambar menjadi suatu ukuran dengan jarak tertentu kemudian gambar tersebut dipotong untuk disesuaikan dengan layer input. Pada *single-scale training* besar kemungkinan untuk kehilangan objek pada suatu gambar dibandingkan dengan *multi-scale training*. VGG-16 menghasilkan nilai eror sebesar 8,8% pada *single-scale training* dan 8,1% pada *multi scale training*, sedangkan VGG-19 menghasilkan nilai eror sebesar 9% pada *single-scale training* dan 8% pada *multi-scale training*. Berikut arsitektur VGG

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64	conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128	conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

**Gambar 5.** Arsitektur VGG  
(Sumber: (Simonyan & Zisserman, 2015))

### 3 Hasil dan Pembahasan

Pada penelitian ini melakukan proses dari problem scoping sampai dengan *modelling*, pada penelitian ini menggunakan data yang berasal dari kaggle yaitu FER2013. Dataset FER2013 berbentuk gambar. Pada gambar yang telah ada ini dilakukan Preprocessing yang artinya dilakukan eksplorasi data dengan memperhatikan data yang sudah ada dilakukan pengecekan. Pada data FER2013 ini dengan data mulai dari ekspresi *angry, disgusted, fearful, happy, neutral, sad*, dan *surprised* yang berarti 7 Jenis ekspresi wajah. Dataset akan dilakukan pembagian yaitu data training 80% dan 20% data test.

Model CNN dibuat dengan arsitektur VGG 16 dibuat dengan menggunakan library keras dan Tensorflow. Hasil dari arsitektur VGG 16 sebagai berikut

```

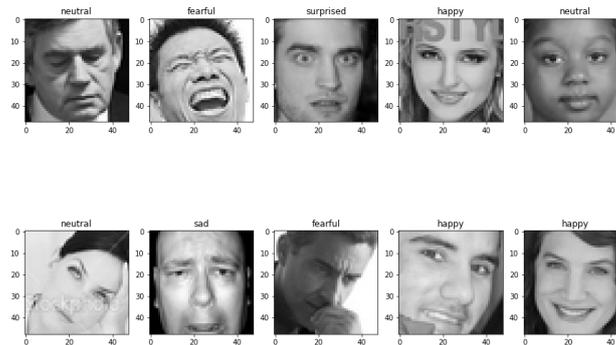
Model: "sequential_2"
Layer (type)                Output Shape                Param #
-----
vgg16 (Functional)          (None, 1, 1, 512)          14714688
dropout_6 (Dropout)         (None, 1, 1, 512)          0
flatten_2 (Flatten)         (None, 512)                 0
batch_normalization_8 (Batch Normalization) (None, 512)                 2048
dense_8 (Dense)              (None, 32)                  16416
batch_normalization_9 (Batch Normalization) (None, 32)                  128
activation_6 (Activation)    (None, 32)                  0
dropout_7 (Dropout)         (None, 32)                  0
dense_9 (Dense)              (None, 32)                  1056
batch_normalization_10 (Batch Normalization) (None, 32)                  128
activation_7 (Activation)    (None, 32)                  0
dropout_8 (Dropout)         (None, 32)                  0
dense_10 (Dense)             (None, 32)                  1056
batch_normalization_11 (Batch Normalization) (None, 32)                  128
activation_8 (Activation)    (None, 32)                  0
dense_11 (Dense)             (None, 7)                   231
-----
Total params: 14,735,879
Trainable params: 7,099,399
Non-trainable params: 7,636,480

```

**Gambar 6.** Hasil dari model summary

Model klasifikasi ekspresi wajah yang penulis gunakan dalam perancangan ini menggunakan arsitektur model VGG-16 yang digunakan sebagai lapisan ekstraksi, output dari lapisan VGG-16 adalah matriks 3 dimensi (1, 1, 512) kemudian menggunakan lapisan *batch normalization* untuk melakukan normalisasi matrix, pada setiap lapisan batch normalisasi diterapkan lapisan dropout sehingga arsitektur model tidak overfitting, matriks yang telah di ekstraksi akan melewati lapisan *full connected layer* dengan menggunakan kernel *he uniform* dan aktivasi fungsi berupa sigmoid. Total lapisan *batch normalisation* yang digunakan adalah 4 lapisan, lapisan *dropout* sebanyak 3 lapisan, lapisan *full connected* yang menggunakan kernel *he uniform* sebanyak 3 lapisan, dan terakhir lapisan fully connected menggunakan aktivasi sigmoid sebagai output yang terdiri dari 7 channel sebanyak 1 lapisan.

Adapun data training yang digunakan dan setelah dilakukan training sebagai berikut



**Gambar 7.** Sebagian Data Trainig

## 4 Kesimpulan

Dapat disimpulkan bahwa perancangan ini menghasilkan model dan arsitektur VGG 16 yang telah ditambahkan layer dan data training yang digunakan dalam data FER 2013. Model klasifikasi ekspresi wajah yang penulis gunakan dalam perancangan ini menggunakan arsitektur model VGG-16 yang digunakan sebagai lapisan ekstraksi, output dari lapisan VGG-16 adalah matriks 3 dimensi (1, 1, 512) kemudian menggunakan lapisan batch normalization untuk melakukan normalisasi matrix, pada setiap lapisan batch normalisasi diterapkan lapisan dropout sehingga arsitektur model tidak *overfitting*, matriks yang telah di ekstraksi akan melewati lapisan full connected layer dengan menggunakan kernel *he uniform* dan aktivasi fungsi berupa sigmoid. Total lapisan *batch normalisation* yang digunakan adalah 4 lapisan, lapisan *dropout* sebanyak 3 lapisan, lapisan *full connected* yang menggunakan kernel *he uniform* sebanyak 3 lapisan, dan terakhir lapisan *fully connected* menggunakan aktivasi sigmoid sebagai output yang terdiri dari 7 channel sebanyak 1 lapisan.

Pada Arsitektur VGG 16 ini dipadukan dengan pengenalan wajah menggunakan MTCNN ini akan memberikan hasil yang maksimal berdasarkan penelitian sebelumnya, kemudian ketika dilakukan training data arsitektur VGG 16 ini merespon dengan baik dan dapat menghasilkan arsitektur yang baik pula.

Untuk pengembangan selanjutnya dari model yang sudah didapatkan dengan model data training dan data test yang didapat dapat dilakukan evaluasi dengan menerapkan atau menggunakan akurasi, lebih tepatnya mengusi model tersebut apakah baik atau tidak dan terakhir dilakukan deployment sehingga mampu mendeteksi emosi berdasarkan ekspresi wajah menggunakan Algoritma VGG 16 ini

## 5 Referensi

- [1] A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan Machine Learning dalam

- Berbagai Bidang: Review paper,” *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 5, no. 1, pp. 75–82, 2020, doi: 10.31294/ijcit.v5i1.7951.
- [2] X. Yin and X. Liu, “Multi-Task Convolutional Neural Network for Pose-Invariant Face Recognition,” *IEEE Trans. Image Process.*, vol. 27, no. 2, pp. 964–975, 2018, doi: 10.1109/TIP.2017.2765830.
- [3] P. P. Kusdiananggalih and E. Rachmawati, “Pengenalan Ekspresi Wajah Dari Cross Dataset Menggunakan Convolutional Neural Network ( CNN ),” vol. 8, no. 2, pp. 3429–3445, 2021.
- [4] K. K. V and V. Bahel, “Transfer Learning Approach for Analyzing Attentiveness of Students in an Online Classroom Environment with Emotion Detection,” no. May, pp. 1–9, 2021, doi: 10.20944/preprints202105.0303.v1.
- [5] A. L. Seandrio, A. H. Pratomo, and M. Yanu, “Implementation of Convolutional Neural Network ( CNN ) in Facial Expression Recognition Implementasi Convolutional Neural Network ( CNN ) Pada Pengenalan Ekspresi Wajah,” vol. 18, no. 2, pp. 211–221, 2021, doi: 10.31515/telematika.v18i2.4823.
- [6] I. Azhari, A. R. Sanjaya, A. R. Sanjaya, D. Wajah, D. Learning, and C. N. Network, “Implementasi Algoritma Convolutional Neural Network Dalam Deteksi Emosi Manusia,” vol. 1, no. 1, pp. 112–118, 2020.
- [7] N. A. S. Badrulhisham and N. N. A. Mangshor, “Emotion Recognition Using Convolutional Neural Network (CNN),” *J. Phys. Conf. Ser.*, vol. 1962, no. 1, 2021, doi: 10.1088/1742-6596/1962/1/012040.
- [8] Y. Pratama, M. Istoningtyas, and E. Rasywir, “Pengujian Algoritma MTCNN (Multi-task Cascaded Convolutional Neural Network) untuk Sistem Pengenalan Wajah,” *J. Media Inform. Budidarma*, vol. 3, no. 3, p. 240, 2019, doi: 10.30865/mib.v3i3.1324.
- [9] Siddhartha, “AI Project Cycle, teh 5 stages.” 2021, [Online]. Available: [https://7-hiddenlayers.com/ai\\_project\\_cycle/](https://7-hiddenlayers.com/ai_project_cycle/).
- [10] H. Dhingra, R. Bansal, S. Banerjee, and S. Verma, *A Text Book of Artificial Intelligence*, Pertama. New Delhi: Goyal Brothers Prakashan, 2020.
- [11] M. A. Pangestu and H. Bunyamin, “Analisis Performa dan Pengembangan Sistem Deteksi Ras Anjing pada Gambar dengan Menggunakan Pre-Trained CNN Model,” *J. Tek. Inform. dan Sist. Inf.*, vol. 4, pp. 337–344, 2018.
- [12] K. Zhang, Z. Zhang, Z. Li, S. Member, Y. Qiao, and S. Member, “Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks,” *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1499–1503, 2016.
- [13] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–14, 2015.