

Prediksi Keputusan Juri Pencarian Bakat Berdasarkan Ekspresi Wajah Menggunakan Analisis VGG16 dan Fuzzy

Fakhrudin Fikriansyah, Anggunmeka Luhur Prasasti, Marisa W. Paryasto

Teknik Komputer, Fakultas Teknik Elektro,

Telkom University

Jl. Telekomunikasi No. 1, Terusan Buahbatu - Bojongsoang, Sukapura, Kec. Dayeuhkolot, Kabupaten Bandung, Jawa Barat 40257, (022) 7566456

fikriansyah@student.telkomuniversity.ac.id, anggunmeka@telkomuniversity.ac.id,

marisaparyasto@telkomuniversity.ac.id

Diterima: 19 Agustus 2022. Disetujui: 25 September 2022. Dipublikasikan: 5 November 2022.

Abstract - Expression is a human communication tool to express feelings. Currently, facial expression recognition technology continues to be developed by research to improve the quality of the technology. With this, existing models will be developed in order to predict human decisions based on facial expressions. This study uses the modified VGG16 architecture as a facial expression classification model. The datasets used for classification are FER-2013 Modification into five expressions, namely angry, disgust, happy, neutral, and surprised, with a total data of 23910. The classification model is used to read the jury's facial expressions via video in the prediction model. The result of the expression that is read will be calculated using Fuzzy Logic in determining the jury's decision with the result of 'yes' or 'no'. The results of testing the Modified VGG16 Architecture using the best combination of parameters were obtained with epoch 100, batch size 32, learning rate 0.0001, and data split 10% for validation to obtain training accuracy of 93% and validation accuracy of 86%. The model is evaluated by testing data of 10% outside of the training data, obtaining a test accuracy of 85%. The classification report from the evaluation obtained a precision of 84%, a recall of 82%, and an f1-score of 83%. The resulting model has a good performance in classifying facial expressions compared to VGG16. The results of the judge's decision prediction using fuzzy obtained a correct prediction of 20:20 from the number of test samples.

Keywords: accuracy, expression, fuzzy, prediction, VGG16

Abstrak-- Ekspresi merupakan suatu alat komunikasi manusia untuk mengungkapkan rasa emosional. Saat ini teknologi pengenalan ekspresi wajah terus dikembangkan oleh penelitian untuk meningkatkan kualitas dari teknologi tersebut. Dengan hal tersebut akan dikembangkan model yang sudah ada agar dapat memprediksi keputusan manusia berdasarkan ekspresi wajahnya. Penelitian ini menggunakan arsitektur VGG16 Modifikasi sebagai model klasifikasi ekspresi wajah. Datasets yang digunakan untuk klasifikasi adalah FER-2013 Modifikasi menjadi lima ekspresi yaitu angry, disgust, happy, neutral, dan surprise dengan jumlah data 23910. Model klasifikasi digunakan untuk pembacaan ekspresi wajah juri melalui video pada model prediksi. Hasil ekspresi yang terbaca akan dihitung menggunakan Logika Fuzzy dalam menentukan keputusan juri dengan hasil 'yes' atau 'no'. Hasil pengujian Arsitektur VGG16 Modifikasi menggunakan kombinasi parameter terbaik diperoleh dengan epoch 100, batch size 32, learning rate 0,0001, dan split data 10% untuk validasi memperoleh akurasi latih 93% dan akurasi validasi 86%. Model dievaluasi dengan data testing sebesar 10% diluar dari data training memperoleh akurasi uji sebesar 85%. Classification Report dari evaluasi memperoleh precision 84%, recall 82%, dan f1-score 83%. Model yang dihasilkan memiliki performansi yang bagus dalam mengklasifikasikan ekspresi wajah dibandingkan dengan VGG16. Hasil prediksi keputusan juri menggunakan fuzzy memperoleh prediksi yang benar sebanyak 20:20 dari jumlah sampel pengujian.

Kata kunci: akurasi, ekspresi, fuzzy, prediksi, VGG16

I. PENDAHULUAN

Seperti yang kita ketahui dalam komunikasi antar manusia, wajah merupakan media untuk menyampaikan ekspresi. Dengan adanya ekspresi tersebut, orang dapat mengetahui yang sedang terjadi pada suatu individu. Ekspresi wajah dapat disebut juga sebagai karakteristik perilaku. Ekspresi wajah sebagai media penyampaian tentang apa yang

ada di dalam pikiran seseorang[1]. Manusia bisa mengekspresikan sesuatu tanpa sengaja maupun tidak disengaja. Hal tersebut terjadi karena perasaan atau ungkapan dari emosi orang tersebut. Terkadang menyembunyikan emosi tertentu dari ekspresi amat sulit. Ekspresi pada wajah dapat diketahui dengan mudah artinya, misalnya ekspresi sedang marah dan senang[2].

Berkembangnya teknologi saat ini, komputer menjadi peranan penting dalam interaksi antara komputer dengan manusia untuk mengenali ekspresi pada manusia. Pendeteksi wajah contohnya mata dan mulut adalah komponen penting dalam hal memproses citra wajah untuk mengidentifikasi wajah dan pendeteksi ekspresi. Pemrosesan citra wajah ini dapat digunakan sebagai langkah awal untuk mendeteksi ekspresi. Pengenalan ekspresi manusia saat ini banyak diimplementasikan untuk kelembagaan, bidang militer, kamera pengawas, dan masih banyak lagi[3].

Penelitian terkait yang dilakukan oleh Zhenhue et.al menggunakan algoritma CNN untuk mengklasifikasikan ekspresi wajah dengan CK+ datasets. Optimizer yang digunakan yaitu SGD[4]. Bin Li et.al melakukan penelitian pengenalan ekspresi wajah menggunakan arsitektur ResNet-50. Datasets yang digunakan berjumlah 700. Penambahan Batch Normalization dapat meningkatkan akurasi pada model. Akurasi yang diperoleh 95% dengan total gambar 700[5]. Penelitian pengenalan ekspresi wajah oleh Kulandai et.al menggunakan fitur ekstraksi HOG (Histogram of oriented Gradient memperoleh akurasi 87% dan menggunakan fitur ekstraksi LBP (Local Binary Pattern) memperoleh akurasi 64%[6]. Krishnadas et.al melakukan penelitian pengenalan ekspresi wajah menggunakan Arsitektur VGG16 dikombinasikan dengan LSTM (Long Short-Term Memory). Pengenalan ekspresi wajah dilakukan secara real-time. Datasets yang digunakan adalah CK+ total gambar 1562 dengan enam label ekspresi yaitu *angry, disgust, fear, happy, sad, dan surprise*. Ukuran data yang digunakan 128x128 pixel. Akurasi yang diperoleh sebesar 99,45%[7].

Terdapat penelitian yang sejenis dengan merancang sistem pengenalan ekspresi wajah oleh Bodavarapu et.al menggunakan berbagai arsitektur, satu diantaranya adalah arsitektur VGG16[8]. Pengujian dilakukan menggunakan datasets FER-2013 dengan jumlah data 35887 terdiri dari 7 label ekspresi, yaitu *angry, disgust, fear, happy, sad, neutral, surprise*. Penelitian tersebut memperoleh akurasi latih 63% dengan akurasi uji 60%. Dengan perolehan nilai akurasi tersebut, perlu adanya pengembangan agar akurasi pada model VGG16 dalam sistem pengenalan ekspresi wajah dapat ditingkatkan.

Tujuan penelitian ini yaitu mengembangkan model pengenalan ekspresi wajah dari arsitektur VGG16 yang dapat digunakan untuk memprediksi

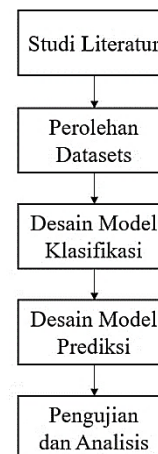
keputusan juri pencarian bakat berdasarkan ekspresi wajah. Arsitektur yang diusulkan adalah Arsitektur VGG16 Modifikasi berbasis Convolution Neural Network (CNN)[9], [10]. Model VGG16 dikembangkan dengan cara menambahkan layer normalisasi dan dropout untuk meningkatkan akurasi[11]. Datasets yang digunakan untuk melatih model klasifikasi adalah FER2013 dengan 7 label ekspresi yang diunduh dari Kaggle. FER-2013 dimodifikasi menjadi 5 label ekspresi dengan total gambar 29310. Hasil model klasifikasi digunakan untuk membaca ekspresi wajah melalui video juri pada model prediksi. Prediksi keputusan juri dilakukan berdasarkan perhitungan fuzzy dengan fungsi input berupa ekspresi wajah dan fungsi output berupa keputusan juri yaitu 'yes' dan 'no'. Penelitian ini akan menghasilkan model yang dapat memprediksi keputusan manusia sebagai contoh keputusan juri pencarian bakat berdasarkan ekspresi wajah.

II. METODE PENELITIAN

Penelitian yang dilakukan berfokus pada performansi dari Model Klasifikasi agar dapat mengenali ekspresi wajah secara optimal dan hasil model akan digunakan untuk pembacaan ekspresi pada model prediksi. Berikut adalah pemaparan dari Metode Penelitian yang dilakukan.

A. Tahapan Penelitian

Pada bagian ini dijelaskan tahapan-tahapan yang dilakukan untuk penelitian. Tahapan dilakukan secara berurutan untuk memperoleh hasil akhir seperti ditampilkan pada gambar 1.



Gambar 1. Tahapan Penelitian

Berikut adalah penjelasan untuk tahapan dari penelitian yang dilakukan :

- 1) Studi Literatur

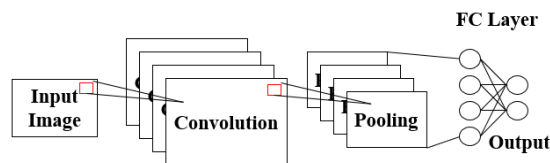
Langkah pertama yaitu melakukan studi literatur dengan mencari referensi yang terkait dengan penelitian. Untuk mencari referensi yang terkait bisa diperoleh dari buku, jurnal ilmiah, prosiding, dan dari berbagai sumber lainnya. Referensi dijadikan sebagai ilmu dasar selama proses penelitian.

2) Perolehan *Datasets*

Datasets yang digunakan pada penelitian ini dibagi menjadi tiga bagian, yaitu *Datasets* FER-2013, *Datasets* Video Juri, dan *Datasets* Nilai Keanggotaan. *Datasets* FER-2013 diunduh dari *Kaggle* dengan 7 label ekspresi, yaitu *angry*, *disgust*, *fear*, *happy*, *sadness*, *neutral*, dan *surprise*. *Datasets* Video Juri diperoleh dari *platform* *YouTube* dengan mengunduh video pencarian bakat untuk mengambil sampel ekspresi juri. *Datasets* Nilai Keanggotaan diperoleh dari jumlah dari masing-masing ekspresi yang terbaca dari video ekspresi juri.

3) Desain Model Klasifikasi

Model Klasifikasi dirancang menggunakan algoritma CNN (*Convolutional Neural Network*). *Convolutional Neural Network* (CNN) merupakan algoritma yang biasanya digunakan dalam mengenali suatu citra atau objek pada gambar dengan menerima *input* berupa gambar untuk melatih mesin mempelajari mengenali suatu objek untuk dibedakan dengan objek lainnya[3]. CNN memiliki struktur jaringan hirarki yang banyak, contohnya transmisi maju dan transmisi mundur. Transmisi maju merupakan *input* fitur melewati lapisan-lapisan operasi sampai menghasilkan *output*. Sedangkan transmisi mundur adalah ketika terjadi perhitungan yang salah antara hasil transmisi maju dengan *tag* sampel yang diberikan sebelumnya. Proses yang salah tersebut akan dikembalikan disetiap *layer*[12], [13].



Gambar 2. Arsitektur CNN Umum

Gambar 2 menampilkan arsitektur CNN. Gambar yang di-*input* akan diproses pada *convolutional layer* yang akan memecah gambar dan menghasilkan *feature map*. *Feature map* tersebut akan disimpan untuk dilakukan *downsampling*. *Downsampling* dilakukan untuk mengambil *pixel* terbesar. Hasilnya akan diproses pada *Fully-*

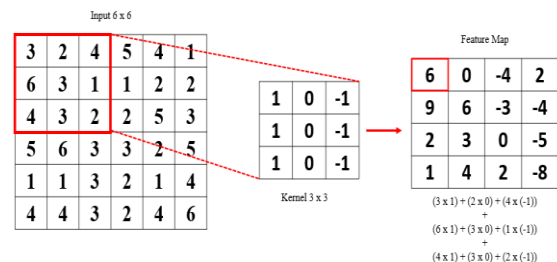
Connected Layer untuk dilakukan proses klasifikasi[14].

Pada CNN tiga *layer* utama, yaitu *Convolutional Layer*, *Pooling Layer*, dan *Fully-Connected Layer*.

a) *Convolutional Layer*

Lapisan konvolusi ini merupakan lapisan pertama pada CNN. Lapisan konvolusi berbentuk sebuah *filter* dengan *pixel* yang memiliki panjang dan tinggi. Filter tersebut akan digeser keseluruhan bagian dari gambar. *Output* yang dihasilkan berupa *feature map*[4].

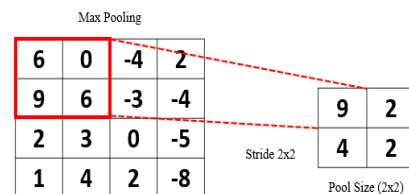
Dapat dilihat pada Gambar 3 proses konvolusi dengan *Input* 6x6 dengan *Kernel* 3x3 dan *Stride* 1x1. Penggunaan *Stride* berukuran 1x1 dapat diartikan bahwa *kernel* akan melakukan pergeseran sebanyak 1 *pixel*[14]. *Kernel* atau *filter* dengan ukuran 3x3 tersebut akan bergeser keseluruh bagian dari *input*. Proses yang terjadi adalah ukuran dari setiap *input* yang dibaca dikalikan dengan ukuran *filter* sehingga menghasilkan *output* berupa *feature map* dengan ukuran yang lebih kecil dari *input*.



Gambar 3. Proses Konvolusi dari Input

b) *Pooling Layer*

Tujuan dari *layer* ini adalah untuk mengurangi dimensi dari *feature map*, sehingga proses komputasi dapat dilakukan lebih cepat karena parameter yang semakin sedikit[3].



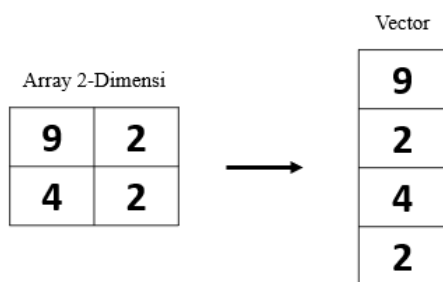
Gambar 4. Proses pada *Pooling Layer*

Proses yang terjadi pada *maxpooling* ini adalah hasil *feature map* dari *convolutional layer* akan diambil nilai *pixel* tertinggi untuk dijadikan sebagai nilai *invariant* baru dari *input*[15]. *Pool Size* merupakan ukuran untuk nilai baru setelah diambil

nilai tertinggi dari *feature map*. Dapat dilihat pada Gambar 4 ukuran *pool* nya 2x2 dengan *stride* 2x2 berarti pembacaannya akan bergeser sebanyak 2 *pixel* ke seluruh *feature map* dengan menghasilkan *invariant* baru dengan ukuran 2x2.

c) *Fully-Connected Layer*

Semua data yang diolah sudah melewati *convolution layer*, *pooling layer*, lalu pada lapisan ini data dari *layer* sebelumnya akan diterima lalu diolah untuk menentukan hasil klasifikasi dari data input pada lapisan konvolusi. Pada *Fully-Connected Layer* ini terdapat *Dense* untuk menghubungkan setiap *neuron*[16].

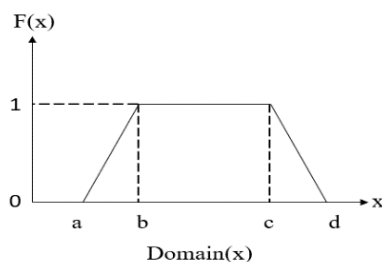


Gambar 5. Perubahan Array 2-Dimensi ke Vector

Output dari *maxpooling* berupa *array* 2-Dimensi. *Array* 2-Dimensi ini akan dijadikan *Vector* untuk dihubungkan ke semua *neuron* pada *FC-Layer* seperti ditampilkan pada Gambar 5.

4) Desain Model Prediksi

Fuzzy merupakan algoritma dengan memproses data *input* dan *output* untuk menentukan hasil akhir. *Fuzzy* pada dasarnya memiliki dua nilai yaitu 0 yang berarti salah dan 1 yang berarti benar[17]. Algoritma *Fuzzy* pada sistem menggunakan *library* dari *scikit-fuzzy* dengan bahasa *python* 3. Model Prediksi dapat membaca ekspresi wajah melalui *datasets* video juri menggunakan algoritma *haarcascade*. Sistem akan membaca ekspresi wajah untuk dijumlahkan masing-masing ekspresi yang terbaca. Hasil penjumlahan ekspresi wajah akan dijadikan sebagai nilai keanggotaan pada sistem *fuzzy*.



Gambar 6. Representasi Trapesium

Gambar 6 menampilkan representasi yang digunakan pada sistem *fuzzy* yaitu representasi trapesium. Representasi trapesium memiliki beberapa nilai derajat keanggotaan 1[18].

Fungsi Keanggotaan Representasi Trapesium menggunakan persamaan (1):

$$\mu_A = \begin{cases} 0 & ; x \leq a \text{ atau } x \geq d \\ \frac{x-a}{b-a} & ; a \leq x \leq b \\ 1 & ; b \leq x \leq c \\ \frac{d-x}{d-c} & ; c \leq x \leq d \end{cases} \quad (1)$$

Keterangan :

$\mu_A[X]$ = Fungsi Keanggotaan Himpunan *Fuzzy*
 a, b, c, dan d = nilai real domain

5) Pengujian dan Analisis

Pengujian dibagi menjadi dua, yaitu pengujian parameter pada Model Klasifikasi dan pengujian prediksi pada Model Prediksi. Model Klasifikasi diuji dengan berbagai parameter yaitu *batch size* dan *learning rate* bertujuan untuk mencari kombinasi parameter dengan performansi terbaik. Model yang memiliki performansi terbaik akan dievaluasi dengan *data testing* untuk mengetahui performansi klasifikasi diluar dari *data training*. Hasil evaluasi akan ditampilkan dalam bentuk *Confusion Matrix*. *Confusion Matrix* adalah matriks untuk melaporkan hasil klasifikasi[6]. Pada matriks terdapat label kelas, *precision*, *recall*, *F1-Score*, dan Akurasi, seperti ditampilkan pada gambar 7.

| | POSITIVE | NEGATIVE |
|----------|-----------|-----------|
| POSITIVE | TP | FN |
| NEGATIVE | FP | TN |

Gambar 7. *Confusion Matrix*

Terdapat beberapa istilah pada *Confusion Matrix*, yaitu :

- a) TP (*True Positive*)
Data positif yang benar diprediksi positif
- b) TN (*True Negative*)
Data negatif yang benar diprediksi negatif
- c) FP (*False Positive*)
Prediksi data positif tetapi salah
- d) FN (*False Negative*)
Prediksi data negatif tetapi salah

Berikut adalah persamaan nilai *Classification Report* sebagai performansi model yang dihasilkan dari *Confusion Matrix* :

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

$$F1 - Score = \frac{2*Precision*Recall}{Precision+Recall} \tag{4}$$

$$Akurasi = \frac{TP+TN}{TP+FP+TN+FN} \tag{5}$$

III. HASIL DAN PEMBAHASAN

A. Perolehan Datasets

Datasets dibagi menjadi tiga, yaitu :

1) FER-2013

Datasets untuk klasifikasi menggunakan data FER2013 yang diunduh dari *Kaggle*. Data yang diunduh dengan format gambar tersebut memiliki 7 label ekspresi, yaitu *angry, disgust, fear, happy, sad, neutral, dan surprise*. Data ini akan dimodifikasi menjadi 5 label ekspresi yaitu *angry, disgust, happy, neutral, dan surprise*.

TABEL I. JUMLAH FER-2013 MODIFIKASI

| Label Ekspresi | Jumlah Gambar |
|-----------------|---------------|
| <i>Angry</i> | 4546 |
| <i>Disgust</i> | 547 |
| <i>Happy</i> | 8829 |
| <i>Neutral</i> | 6078 |
| <i>Surprise</i> | 3910 |
| Total | 23910 |

Modifikasi dilakukan dengan menghapus dua *label* ekspresi secara manual yaitu *fear* dan *sad* karena pada umumnya juri pencarian bakat lebih sering mengekspresikan ke-lima ekspresi pada Tabel 1. Data dibagi untuk *training* 19368 gambar, untuk *validation* 2150 gambar, dan data *testing* 2392 gambar. Data *train* dan data *validation* digunakan proses *training* pada model. Hasil dari model akan dievaluasi menggunakan data *testing* untuk mengukur performansi model dalam membaca data diluar proses *training*.

2) Video Juri

Datasets video juri didapatkan dari video pencarian bakat di *YouTube* dengan mengambil potongan ekspresi juri selama pertunjukkan berlangsung. Durasi video dipotong untuk pengambilan setiap ekspresi dari juri. Juri disorot

oleh kamera ketika melakukan ekspresi yang menunjukkan bahwa juri suka atau tidak pada kontestan. Pada video tersebut terdapat tiga sampai lima juri.

Pengambilan sampel video ekspresi ini dengan cara meng-*capture* satu juri untuk dijadikan satu sampel. Saat pengambilan sampel data, video ditonton terlebih dahulu dengan melihat keputusan juri untuk mengetahui hasilnya, seperti ditampilkan pada gambar 8.



Gambar 8. Contoh Video Juri

3) Nilai Keanggotaan

Datasets Nilai Keanggotaan digunakan dalam memprediksi keputusan juri melalui ekspresi wajah. Nilai yang terbaca merupakan hasil perhitungan dari ekspresi yang terbaca dengan kualitas video 25 FPS (*frame per second*).

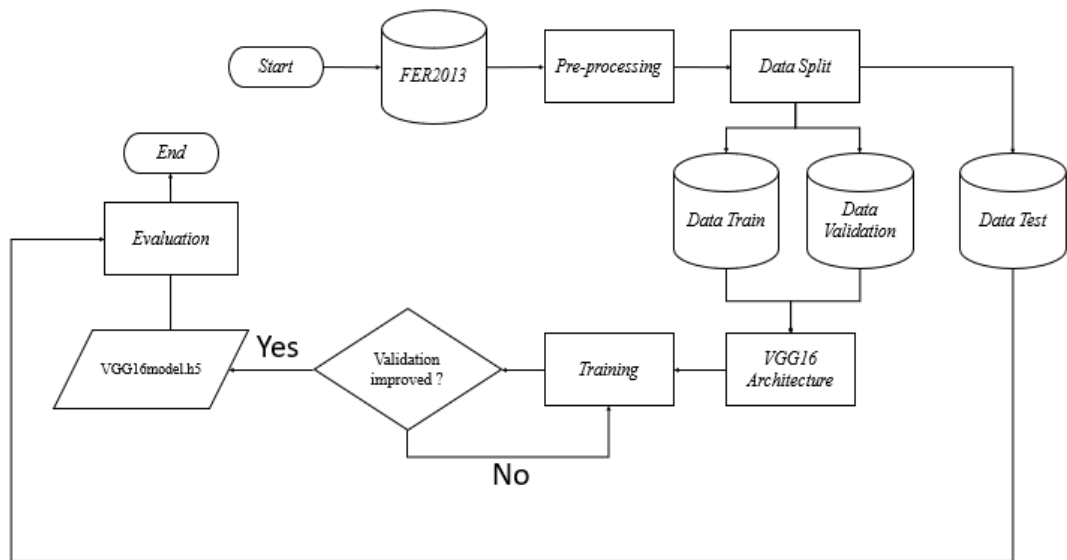
TABEL II. NILAI UNIVERSAL EKSPRESI

| Fungsi Keanggotaan | Nilai Universal |
|--------------------|-----------------|
| <i>Angry</i> | [0, 20] |
| <i>Disgust</i> | [0, 20] |
| <i>Happy</i> | [0, 120] |
| <i>Neutral</i> | [0, 100] |
| <i>Surprise</i> | [0, 100] |

Tabel 2 menampilkan nilai universal ekspresi. Nilai universal dari setiap label ekspresi diperoleh dari nilai tertinggi jumlah ekspresi wajah yang terbaca selama pemutaran video. Untuk memperoleh nilai minimum dan maksimum dilakukan dengan mengambil jumlah ekspresi tertinggi dan terendah dari setiap label ekspresi menggunakan sampel yang berbeda agar nilai lebih bervariasi.

B. Desain Model Klasifikasi

Gambar 9 menampilkan model klasifikasi yang akan digunakan. Arsitektur yang digunakan untuk Klasifikasi ekspresi wajah adalah VGG16 Modifikasi. Model akan diuji menggunakan berbagai kombinasi nilai parameter dengan *batch*



Gambar 9. Flowchart Model Klasifikasi

size 16, 32, 64, dan learning rate 0,01; 0,001; 0,0001; 0,00001 untuk memperoleh nilai akurasi terbaik. Akurasi tersebut akan disimpan ke dalam file VGG16model.h5.

1) *Pre-processing*

Pre-processing merupakan tahap untuk memproses *datasets* dari FER2013. Tahap ini berguna untuk menyamakan ukuran gambar dari *datasets* dengan ukuran 48 x 48 pixel agar proses *training* dapat berjalan. Data dibaca dalam format warna *grayscale* karena data asli pada FER2013 berwarna *grayscale*. Terdapat tahap augmentasi data yang berguna untuk memanipulasi data agar pembacaan *datasets* lebih optimal. Berikut beberapa fitur augmentasi data yaitu *rescale*, *rotation_range*, *width_shift_range*, *height_shift_range*, *zoom_range*, *shear_range*, dan *horizontal_flip*.

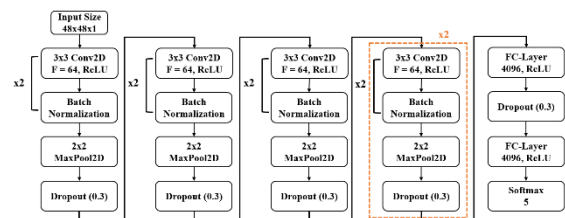
2) *Data Split*

Datasets dibagi menjadi tiga yaitu untuk *training*, *validation*, dan *testing*. Pada saat proses *training*, *datasets* yang digunakan hanya Data *training* dan *validation*. Data *testing* digunakan untuk evaluasi model pada tahap terakhir. Evaluasi bertujuan mengukur performansi model dalam mengklasifikasikan data diluar data *training*.

3) *VGG16 Architecture*

Arsitektur VGG16 pada umumnya memiliki 13 *convolutional layer* dan 3 *fully-connected layer*, seperti ditampilkan pada gambar 10. Pada pengujian yang dilakukan, ada dua jenis *layer* yang ditambahkan ke dalam arsitektur, yaitu *BatchNormalization* dan *Dropout*. Penambahan

BatchNormalization agar model dapat mempelajari partisi data saat proses *training* sehingga proses *training* lebih cepat[5]. Penambahan *Dropout* bertujuan untuk menonaktifkan *neuron* yang tidak diperlukan agar mencegah terjadinya *overfitting*[19]. Penggunaan *Dropout* sebesar 0,3 artinya 30% jumlah *neuron* akan diatur bobotnya menjadi nol bertujuan sebagai lintasan maju atau mundur dalam satu iterasi.



Gambar 10. Arsitektur VGG16 Modifikasi

- a) Pada *Block* ke-1 terdapat *Convolutional Layer* pertama dan kedua memiliki *kernel* dengan ukuran 3x3 dengan *filter* sebesar 64 dan *input shape* sebesar 48x48x1. *Activation* yang digunakan adalah ReLU (*Rectified Linear Unit*). Pada *Convolutional Layer* setelah *activation* ada *BatchNormalization* lalu disusul oleh *Pooling Layer* dengan *stride* sebesar 2 dan penutup *block* 1 menggunakan *Dropout* dengan nilai 0,3.
- b) Pada *Block* ke-2 terdapat *Convolutional Layer* sebanyak dua *layer* dengan ukuran *kernel* 3x3, *filter* sebesar 128, dengan *Activation* ReLU. Setelah *Activation* akan ada *BatchNormalization*.

- Layer ditutup dengan *Pooling Layer* dengan *stride* sebesar 2 dan *Dropout* dengan nilai 0,3.
- c) Pada *Block* ke-3 terdapat *Convolutional Layer* sebanyak tiga *layer* dengan ukuran *kernel* 3x3, *filter* sebesar 512, dengan *Activation* ReLU. Setelah *Activation* akan ada *BatchNormalization*. *Layer* ditutup dengan *Pooling Layer* dengan *stride* sebesar 2 dan *Dropout* dengan nilai 0,3.
 - d) Pada *Block* ke-4 terdapat *Convolutional Layer* sebanyak tiga *layer* dengan ukuran *kernel* 3x3, *filter* sebesar 512, dengan *Activation* ReLU. Setelah *Activation* akan ada *BatchNormalization*. *Layer* ditutup dengan *Pooling Layer* dengan *stride* sebesar 2 dan *Dropout* dengan nilai 0,3.
 - e) Pada *Fully-connected Layer* dengan 4096 unit pada satu *Dense layer* dengan *Activation* ReLU ditambah *Dropout* dengan nilai 0.3 dan 4096 unit pada satu *Dense layer* dengan *Activation* ReLU yang diikuti oleh *Softmax*.

4) *Training*

Model VGG16 Modifikasi ini dibangun dengan menyusun *layer* dari Keras API dan menggunakan *library Tensorflow*. Proses *training* menggunakan data dari FER2013 yang sudah dimodifikasi menjadi lima ekspresi. Data yang digunakan memiliki ukuran 48x48 *pixel* dengan mode warna *grayscale* (1 *channel*). Partisi data dibagi menjadi dua, ada yang untuk *training* dan untuk *validation*. *Training* dilakukan berdasarkan jumlah *epoch*, *batch size*, dan *learning rate*.

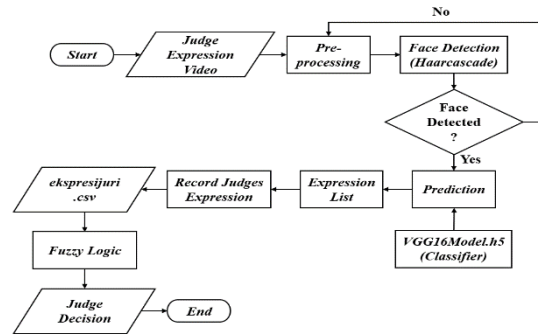
5) *Evaluatoin*

Evaluasi model yang dilakukan bertujuan untuk menguji performansi model yang sudah dilatih. Model dievaluasi menggunakan data *testing* sebesar 10% yaitu 2392 gambar di luar data *training* untuk menguji performansi model dalam mengklasifikasikan data diluar data *training*. Hasil evaluasi ditampilkan dalam *Confusion Matrix* yang berisi nilai *precision*, *recall*, *f1-score*, dan akurasi data uji.

C. *Desain Model Prediksi*

Model prediksi ini menggunakan video *input* berupa video ekspresi juri pencarian bakat khususnya menyanyi. Video diambil dari *YouTube* untuk dijadikan sampel dari ekspresi juri. Pada satu sampel video hanya terdapat satu juri dalam *frame*. Ekspresi juri didapatkan saat *talent* sedang tampil.

Pada perancangan mesin prediksi ini terdapat beberapa proses utama, yaitu *input video*, *pre-processing*, pembacaan ekspresi, *fuzzy logic*, seperti ditampilkan pada gambar 11.



Gambar 11. Flowchart Model Prediksi

1) *Pre-processing*

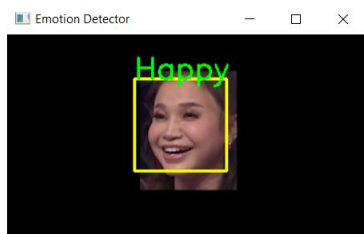
Pre-Processing pada mesin prediksi merupakan tahap untuk memproses ukuran citra wajah pada video *input*. Video *input* akan diproses dengan mendeteksi fitur wajah dan memberikan *boundingbox* pada wajah yang dideteksi. Proses deteksi wajah menggunakan logika *haarcascade*. Pada saat wajah dideteksi, video dengan format RGB (*Red Green Blue*) akan diolah menjadi format *grayscale*. Hal ini dilakukan agar pembacaan wajah sesuai dengan model VGG16 yang sudah di-*training* pada model klasifikasi. Proses *resize* dilakukan untuk mengubah ukuran dari citra menjadi 48x48 *pixel* agar sesuai dengan *input shape* pada model VGG16 Modifikasi.

2) *Pembacaan Ekspresi*

Video yang sudah diproses pada tahap sebelumnya, akan dilakukan untuk pembacaan ekspresi. Pembacaan ekspresi ini menggunakan *classifier* dari model VGG16.h5. Mesin akan membaca ekspresi dari video yang di-*input* dan hasil pembacaan ekspresi tersebut akan ditampilkan pada *frame* selama pemutaran video berlangsung.

Sistem akan langsung membaca ekspresi tersebut dan menampilkannya pada *frame*. Setiap ekspresi yang muncul pada video dijumlahkan dengan masing-masing label ekspresi. Gambar 12 menampilkan hasil pembacaan ekspresi wajah selama pemutaran video durasi 4 detik.

Ekspresi yang dibaca selama video berlangsung akan disimpan pada file *ekspresijuri.csv* yang dijadikan nilai fungsi keanggotaan untuk proses prediksi keputusan juri berdasarkan ekspresi wajah.



Gambar 12. Pembacaan Ekspresi Wajah Melalui Video

3) *Fuzzy Logic*

Fuzzy Logic pada sistem menggunakan modul dari *scikit-fuzzy*. Gambar 13 menampilkan *perhitungan Fuzzy* yang dilakukan dengan membuat nilai universal dengan lima label ekspresi berdasarkan nilai ekspresi dari *ekspresijuri.csv*. Lima label ekspresi akan dijadikan sebagai fungsi *input* dan dua keputusan juri yaitu 'yes' dan 'no' akan dijadikan sebagai fungsi *output*. Dari *input* dan *output* yang sudah dibuat, akan dibuat *rules* untuk menentukan keputusan. *Rules* dibuat sebanyak 32 berdasarkan jumlah fungsi *output* dipangkatkan dengan jumlah fungsi *input*.

Jumlah dari Setiap Ekspresi yang ditampilkan oleh juri :
 Angry : 0
 Disgust : 0
 Happy : 99
 Neutral : 18
 Surprise : 0

Gambar 13. Hasil penjumlahan ekspresi yang muncul

D. *Pengujian dan Analisis*

Pengujian yang dilakukan dibagi menjadi dua, yaitu pengujian parameter pada Model Klasifikasi dan Pengujian Model Prediksi.

1) *Pengujian Parameter pada Model Klasifikasi*

Pengujian yang dilakukan menggunakan rasio data 90:10. Pada pengujian ini, data akan diuji nilai parameter yang berbeda dari parameter *batch size* (BS) dan nilai *learning rate* (Lr) untuk dibandingkan akurasi (Acc) dari model VGG16 Modifikasi dan VGG16, seperti ditampilkan pada Tabel 3.

TABEL III. HASIL PENGUJIAN PARAMETER VGG16 MODIFIKASI

| Arsitektur | Rasio Data | BS | Lr | Acc |
|------------------|------------|----|---------|------|
| VGG16 Modifikasi | 90:10 | 16 | 0,01 | 37 % |
| | | | 0,001 | 84 % |
| | | | 0,0001 | 85 % |
| | | 32 | 0,00001 | 82 % |
| | | | 0,01 | 37 % |
| | | | 0,001 | 85 % |
| | | | 0,0001 | 86 % |

| Arsitektur | Rasio Data | BS | Lr | Acc |
|------------|------------|----|---------|------|
| | | 64 | 0,00001 | 80 % |
| | | | 0,01 | 37 % |
| | | | 0,001 | 83 % |
| | | | 0,0001 | 84 % |
| | | | 0,00001 | 75 % |

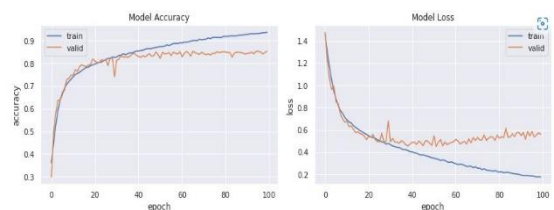
Untuk mengetahui peningkatan performa dari VGG16 Modifikasi, dilakukan pengujian parameter pada Arsitektur VGG16 Asli seperti ditampilkan pada Tabel 4.

TABEL IV. HASIL PENGUJIAN PARAMETER VGG16

| Arsitektur | Rasio Data | BS | Lr | Acc |
|------------|------------|------|---------|------|
| VGG16 | 90:10 | 16 | 0,01 | 37 % |
| | | | 0,001 | 37 % |
| | | | 0,0001 | 37 % |
| | | | 0,00001 | 83 % |
| | | | 0,01 | 37 % |
| | | 32 | 0,001 | 37 % |
| | | | 0,0001 | 83 % |
| | | | 0,00001 | 81 % |
| | | | 0,01 | 37 % |
| | | | 0,001 | 37 % |
| 64 | 0,0001 | 83 % | | |
| | 0,00001 | 79 % | | |

Dapat disimpulkan bahwa VGG16 Modifikasi memiliki peningkatan Akurasi Validasi sebesar 3% yaitu 86% dibandingkan dengan VGG16. Model dengan performa terbaik akan dilakukan Analisis untuk mengetahui performa model dalam proses klasifikasi ekspresi wajah.

VGG16 Modifikasi kombinasi parameter terbaik dengan pembagian rasio data 90:10, *batch size* 32, dan *learning rate* 0,0001. Akurasi latihan yang diperoleh 0,9320 dengan *loss* 0,1874 dan akurasi validasi 0,8535 dengan validasi *loss* 0,5857, seperti ditampilkan pada gambar 14. Model dengan kombinasi parameter terbaik ini akan dievaluasi dalam bentuk *confusion matrix* menggunakan data *testing* sebanyak 2392 gambar diluar dari data *training* untuk menguji performansi model dalam mengklasifikasikan gambar.



Gambar 14. Grafik akurasi dan *loss* VGG16 Modifikasi

TABEL V. HASIL EVALUASI VGG16 MODIFIKASI

| Ekspresi | Precision | Recall | F1-Score |
|----------|-----------|--------|----------|
| Angry | 77% | 72% | 75% |
| Disgust | 85% | 73% | 78% |
| Happy | 93% | 89% | 91% |
| Neutral | 78% | 85% | 81% |
| Surprise | 87% | 91% | 89% |

Berdasarkan hasil evaluasi pada Tabel 5, diperoleh hasil evaluasi model VGG16 Modifikasi kombinasi terbaik memperoleh nilai *precision* 84%, *recall* 82%, *f1-score* 83%, dan akurasi dari data *testing* 85%. Dapat disimpulkan bahwa VGG16 Modifikasi memiliki performansi lebih baik dalam mengklasifikasikan ekspresi wajah dan daya ingat model bagus dengan data yang dibaca secara benar dibandingkan dengan model VGG16 Asli. Nilai *F1-Score* merupakan nilai dari rata-rata antara *precision* dan *recall* dan nilai akurasi merupakan total keseluruhan gambar yang diklasifikasikan secara benar

2) Pengujian Model Prediksi

Pengujian Model Prediksi dilakukan dengan mengambil nilai keanggotaan dari *datasets* pembacaan ekspresi juri menggunakan 10 sampel yang berbeda. Lalu pengujian dilakukan dengan video diluar dari *datasets* dengan 20 sampel pengujian seperti ditampilkan pada Tabel 6.

TABEL VI. HASIL PENGUJIAN PREDIKSI MENGGUNAKAN FUZZY

| D | 1 | 2 | 3 | 4 | 5 | K | P |
|----|----|---|-----|-----|----|-----|-----|
| 1 | 0 | 0 | 0 | 86 | 0 | No | No |
| 2 | 40 | 0 | 0 | 43 | 0 | No | No |
| 3 | 48 | 0 | 0 | 32 | 0 | No | No |
| 4 | 0 | 0 | 0 | 111 | 0 | No | No |
| 5 | 14 | 0 | 0 | 67 | 32 | No | No |
| 6 | 0 | 0 | 111 | 1 | 0 | Yes | Yes |
| 7 | 0 | 0 | 62 | 29 | 1 | Yes | Yes |
| 8 | 0 | 0 | 40 | 24 | 34 | Yes | Yes |
| 9 | 0 | 0 | 83 | 3 | 0 | Yes | Yes |
| 10 | 0 | 0 | 81 | 65 | 0 | Yes | Yes |
| 11 | 0 | 0 | 111 | 31 | 1 | Yes | Yes |
| 12 | 0 | 0 | 51 | 96 | 0 | Yes | Yes |
| 13 | 0 | 0 | 82 | 53 | 0 | Yes | Yes |
| 14 | 4 | 0 | 131 | 0 | 0 | Yes | Yes |
| 15 | 35 | 0 | 89 | 7 | 4 | Yes | Yes |
| 16 | 0 | 0 | 134 | 0 | 0 | Yes | Yes |
| 17 | 0 | 0 | 95 | 38 | 0 | Yes | Yes |
| 18 | 0 | 0 | 73 | 59 | 0 | Yes | Yes |
| 19 | 0 | 0 | 104 | 13 | 0 | Yes | Yes |
| 20 | 0 | 0 | 110 | 27 | 0 | Yes | Yes |

Keterangan :

- D = Data
- 1 = *Angry*
- 2 = *Disgust*
- 3 = *Happy*
- 4 = *Neutral*

5 = *Surprise*

K = Keputusan (Keputusan Asli Juri)

P = Prediksi (Hasil prediksi *Fuzzy*)

Hasil pengujian prediksi keputusan juri menggunakan sistem *fuzzy* menguji 20 sampel video yang berbeda dengan 5 sampel keputusan ‘no’ dan 15 sampel keputusan ‘yes’. Keputusan asli dari juri pencarian bakat dapat diprediksi menggunakan *fuzzy* dengan jumlah data diprediksi secara benar 20:20 dari jumlah data pengujian

IV. KESIMPULAN DAN SARAN

A. Kesimpulan

Penelitian yang dilakukan merupakan pengembangan dari model VGG16. Model dikembangkan dengan menambahkan *batchnormalization* dan *dropout* dengan tujuan meningkatkan performansi, mengurangi waktu pelatihan dan mencegah terjadinya *overfitting*. Model VGG16 Modifikasi diuji menggunakan berbagai nilai parameter untuk mencari kombinasi parameter dengan performansi yang terbaik. Diperoleh kombinasi parameter terbaik menggunakan *Batch Size* 32, *Epoch* 100, *Rasio Data* 90% data *train* dengan jumlah data 19368, 10% data *validation* dengan jumlah data 2150, dan *Learning Rate* 0,0001. Model VGG16 Modifikasi memperoleh akurasi data *training* 93%, akurasi data *validation* 86%. Model tersebut dievaluasi menggunakan data *testing* dengan jumlah data 2392 yang bertujuan mengukur performansi model dengan mengklasifikasikan ekspresi diluar dari *data* pelatihan. Hasil evaluasi memperoleh akurasi data *testing* 85%. *Classification Report* yang diperoleh dengan nilai *precision* 84%, *recall* 82%, dan *F1-Score* 83%. Dengan nilai tersebut, Model VGG16 Modifikasi dapat mengklasifikasikan ekspresi wajah secara optimal dan performansinya lebih baik dibandingkan dengan Model VGG16. Pengujian pada memprediksi keputusan juri pencarian bakat menggunakan 20 sampel video dengan 5 sampel keputusan ‘no’ dan 15 sampel keputusan ‘yes’. Penggunaan *fuzzy* dinilai cukup efektif dalam memprediksi keputusan juri berdasarkan ekspresi wajah dengan jumlah data benar 20:20 dari jumlah data pengujian.

B. Saran

Dari penelitian yang dilakukan, model masih dapat dikembangkan dengan menambahkan metode fitur ekstraksi agar performa model dapat meningkat. Dari *datasets* yang digunakan, perlu adanya penyeimbangan jumlah *datasets* dari masing-masing kelas.

REFERENSI

[1] S. Zhang, X. Pan, Y. Cui, X. Zhao, and L. Liu, "Learning Affective Video Features for Facial Expression Recognition via Hybrid Deep Learning," *IEEE Access*, vol. 7, pp. 32297–32304, 2019, doi: 10.1109/ACCESS.2019.2901521.

[2] C. Qi *et al.*, "Facial Expressions Recognition Based on Cognition and Mapped Binary Patterns," *IEEE Access*, vol. 6, pp. 18795–18803, 2018, doi: 10.1109/ACCESS.2018.2816044.

[3] G. Yue and L. Lu, "Face Recognition Based on Histogram Equalization and Convolution Neural Network," *Proceedings - 2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2018*, vol. 1, pp. 336–339, 2018, doi: 10.1109/IHMSC.2018.00084.

[4] Z. Zhang and M. Li, "Research on facial expression recognition based on neural network," *Proceedings - 2020 International Conference on Computer Network, Electronic and Automation, ICCNEA 2020*, pp. 78–81, 2020, doi: 10.1109/ICCNEA50255.2020.00025.

[5] B. Li and D. Lima, "Facial expression recognition via ResNet-50," *International Journal of Cognitive Computing in Engineering*, vol. 2, pp. 57–64, 2021, doi: 10.1016/j.ijcce.2021.02.002.

[6] J. K. Josephine Julina and T. S. Sharmila, "Facial Emotion Recognition in Videos using HOG and LBP," *2019 4th IEEE International Conference on Recent Trends on Electronics, Information, Communication and Technology, RTEICT 2019 - Proceedings*, pp. 56–60, 2019, doi: 10.1109/RTEICT46194.2019.9016766.

[7] N. Krishnadas and N. T. Bhuvan, "Facial Expression Recognition Using VGG16 and LSTM," *Turkish Online Journal of Qualitative Inquiry (TOJQI)*, vol. 12, no. 7, pp. 9657–9674, 2021.

[8] P. N. R. Bodavarapu and P. V. V. S. Srinivas, "Facial expression recognition for low resolution images using convolutional neural networks and denoising techniques," *Indian Journal of Science and Technology*, vol. 14, no. 12, pp. 971–983, 2021, doi: 10.17485/ijst/v14i12.14.

[9] N. Nour, M. Elhebir, and S. Viriri, "FACE EXPRESSION RECOGNITION USING CONVOLUTION NEURAL NETWORK (CNN) MODELS," vol. 11, no. 1, pp. 1–11, 2020, doi: 10.5121/ijca.2020.11401.

[10] F. Altekin and H. Demir, "Emotion Detection from Facial Expression Using Different Feature Descriptor Methods with Convolutional Neural Networks," vol. 4, no. July, pp. 14–17, 2021.

[11] B. Ko, H. G. Kim, and H. J. Choi, "Controlled dropout: A different dropout for improving training speed on deep neural network," *2017 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2017*, vol. 2017-Janua, pp. 972–977, 2017, doi: 10.1109/SMC.2017.8122736.

[12] L. Hui and S. Yu-Jie, "Research on face recognition algorithm based on improved convolution neural network," *Proceedings of the 13th IEEE Conference on Industrial Electronics and Applications, ICIEA 2018*, pp. 2802–2805, 2018, doi: 10.1109/ICIEA.2018.8398186.

[13] F. A. Isman, A. L. Prasasti, and R. A. Nugrahaeni, "Expression Classification For User Experience Testing Using Convolutional Neural Network," in *2021 International Conference on Artificial Intelligence and Mechatronics Systems (AIMS)*, Apr. 2021, pp. 1–6. doi: 10.1109/AIMS52415.2021.9466088.

[14] S. Tammina, "Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images," *International Journal of Scientific and Research Publications (IJSRP)*, vol. 9, no. 10, p. p9420, 2019, doi: 10.29322/ijsrp.9.10.2019.p9420.

[15] A. Ghosh, A. Sufian, F. Sultana, A. Chakrabarti, and D. De, *Fundamental concepts of convolutional neural network*, vol. 172, no. January. 2019. doi: 10.1007/978-3-030-32644-9_36.

[16] 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," *2020 International Conference on Information Technology Systems and Innovation, ICITSI 2020 - Proceedings*, pp. 256–262, 2020, doi: 10.1109/ICITSI50517.2020.9264940.

[17] E. D. S. Mulyani and J. P. Susanto, "Classification of maturity level of fuji apple fruit with fuzzy logic method," *2017 5th International Conference on Cyber and IT Service Management, CITSM 2017*, 2017, doi: 10.1109/CITSM.2017.8089294.

[18] F. Y. Mulato, "Klasifikasi Kematangan Buah Jambu Biji Merah (Psidium Guajava) dengan Menggunakan Model Fuzzy," pp. 1–155, 2015.

[19] B. Taha and D. Hatzinakos, "Emotion Recognition from 2D Facial Expressions," *2019 IEEE Canadian Conference of Electrical and Computer Engineering, CCECE 2019*, pp. 1–4, 2019, doi: 10.1109/CCECE.2019.8861751.