Jurnal Sistem Komputer dan Informatika (JSON)
Volume 5, Nomor 1, September 2023
e-ISSN 2685-998X
DOI 10.30865/json.v5i1.6768

Facial Expression Recognition inCovid-19 Pandemic EraUsing 2 Stage Convolutional Neural Network

Kevin Kurniawan*, Lili Ayu Wulandhari, Antonius Rildo Pramudya Gondosiswojo

Computer Science Department, Master of Computer Science, BINUS Graduate Program, Bina Nusantara University, Jakarta, Indonesia
Email: 1,*kevin.kurniawan004@binus.ac.id, 2lili.wulandhari@binus.ac.id, 3antonius.gondosiswojo@binus.ac.id
Corresponding Author Email: kevin.kurniawan004@binus.ac.id
Submitted: **07/09/2023**; Accepted: **28/09/2023**; Published: **29/09/2023**

Abstract−In the Covid-19 pandemic era, the use of face mask has become mandatory for all citizens to prevent the spread of the virus. This regulation has becomes a big problem for the Facial Expression Recognitions (FER) applications because face mask cover more than halfof human faces. Starting from this problem, this experiment attemps to produce a robust network which can perform well in both conditions: recognizing expressions with and without a face mask. Dataset used in this experiment is FER2013, with a preprocessing step to produce a FER2013 masked. This research uses 2-stage network, where the first network is used to recognize whether the subject is wearing a mask or not, and then the second network is used to recognize the expression based on the result from the first stage. The network in this experiment is based on Convolutional Neural Network (CNN), with Imagenet as our pretrainedmodel and EfficientNet as our architecture model.Ourproposed model has shown quite good performance in recognizing expressions, even when the data consists of subjects who use and do not use masks with an accuracy of 57.55%.

Keyword: Pandemic Era; Covid-19; Facial Expression Recognition; Face Mask; Occlusion; Convolutional Neural Network

1. INTRODUCTION

Machine learning is one of the most popular areas of digitalization at the moment. Machine learning allows systems to perform tasks that were previously only possible for humans. One of the interesting topics from the field of machine learning is Facial Expression Recognition (FER). FER has become increasingly popular in recent years due to its many applications in everyday life, such as market research, in-depth interviews, and more[1][2].

Due to its many applications in everyday life, many researchers have conducted various studies on this topic. Research related to facial expression recognition has been conducted by Minaee et al., whoused the FER2013 dataset and compared Aff-Wild2(VGG Backbone), Attentional Convolutional Network (ACN), GoogleNet, etcTheir experiment concluded that the Aff-Wild2(VGG Backbone) method is the best solution for facial expression recognition without mask, with an accuracy of 75%. The second-best method was Attentional Convolutional Network (ACN), with an accuracy of 70.02%[3][4][5]. Another study was conducted by Luan et al. (2022), who used the FER2013 dataset and compared MaskingBlock& ResNet34 (their proposed solution), CBAM_ResNet50, and ResNet152. Their experiment concluded that MaskingBlock with ResNet34 performed the best, with an accuracy of 74.14%. CBAM_ResNet50 achieved an accuracy of 73.39%, and ResNet152 achieved an accuracy of 73.22%[6]. From those experiments, we knew that state-of-the-art (SOTA) of FER field is already quite good, with accuracy up to 75% with the FER2013 scenario[7].

But in everyday life, there are various challenges to be able to build an efficient FER model. One of them is if there is occlusion at the face area. Face mask is one of the common examples of occlusion and become a big challenge for FER because more than half of human faces are covered by masks. This is further strengthened by the existence of government policies that requires the use of masks in Covid19 pandemic era. Figure 1 shows a map of the distribution of Covid19 until April 2022[8].

Figure 1. Growth of Covid19 in Indonesia [8]

The problem of facial expression recognition with face masks has been realized by researchers, as well as the possibility of future pandemics like COVID-19. Based on the results of research conducted by Minaee et al. (2020) titled "Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network," it can be concluded that a lot of information is needed to accurately recognize facial expressions, including the region of

Copyright © 2023 **Kevin Kurniawan**, Page 102
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Jurnal Sistem Komputer dan Informatika (JSON) Hal: 102−111
Volume 5, Nomor 1, September 2023
e-ISSN 2685-998X
DOI 10.30865/json.v5i1.6768

interest around the eyes, mouth, and nose. Figure 2 shows the region of interest obtained from the results of this research[3][9].

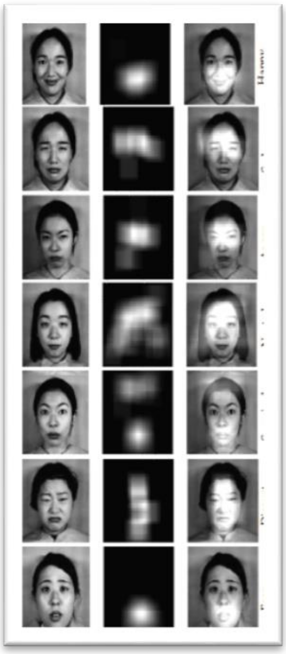

Figure 2. Region Interest of Facial Expression [3]

Based on the region of interest above, if the subject has face mask at the face area, then that will become a big problem to recognize expression. Especially for expressions that depend on the area of the nose down, for example expressions like fear, happy, surprised, and disgust.There are several studies that have begun to focus on this problem. Among them is the research conducted by Yang et al, where the experiment used the LFW dataset. Where the purpose of this experiment is to be able to recognize someone's facial expressions when the subject is wearing a mask. Emotions that are recognized only consist of 3 basic human emotions, namely positive, negative, and neutral. This experiment compares mobile net with vgg19. The results obtained show that mobile net has greater accuracy than vgg19 with an accuracy of 79.09%[10]. But keep in mind that this experiment only recognizes 3 basic human expressions. The other experiment established by Castellano et al., where in this experiment using FER2013 which was preprocessed to crop the nose down. This is done to be able to simulate as if the subject is wearing a mask, where the mask area is not required to analyze the expression. The architecture used in this experiment is CNN. The results of this experiment show that the difficulty level of being able to perform expression recognition when using a mask is still very difficult, as evidenced by the accuracy results obtained of only 43%.Where the model is completely unable to classify disgust expressions[11]. From those experiments, we can conclude that use of mask is really become a big challenge for FER field.

Starting from this problem, this experiment attemps to produce a robust network which can perform well in both conditions: recognizing expressions with and without a face mask.This experiment used the FER2013 dataset, which consists of 7 labeled emotions with up to 28,709 training images and 7,178 test images. We used the FER2013 dataset to generate a FER2013 masked version dataset. From the preprocessing step, we achieved 11,708 training images and 4,962 test images. After that, our proposed method tried to construct a robust model to achieve our goal, which approach consists of 3 deep learning models: Masked/Unmasked Model, FER Masked Model, and FER Unmasked Model. We built 3 models to be able to narrow the region of interest in the model analysis so that it can provide good accuracy results in accordance with the conditions of each model specialization.

2. RESEARCH METHODOLOGY

This section would try to provide detail methodology and step by step of this experiment, from detail dataset used, preprocessing step, until our proposed architecture model and its comparation scenario.

2.1 Dataset Preprocessing

The FER2013 dataset consists of subjects, which each data already have an expression label . But the problem is there is no data for masked images. So to produce FER2013 mask version dataset, we use modeling results from Aqeel Anwar named MaskTheFace[12].First of all, Aqeel Anwar's team prepared 3 pictures of masks from different angles. Then the subject of the image will be detected related to its face landmarks, where the dataset used is VGGFace2. After that the system will estimate the area where the mask will be affixed using the facenet

Copyright © 2023 Kevin Kurniawan, Page 103
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 5, Nomor 1, September 2023
e-ISSN 2685-998X
DOI 10.30865/json.v5i1.6768

architecture. Finally, the mask image that has been prepared will be pasted in the area that has been estimated as the nose down area. In our experiment, we immediately implemented the model developed by the Aqeel Anwar team with the FER2013 dataset.Figure 4 explains how to produce FER2013 mask version[13], [14].

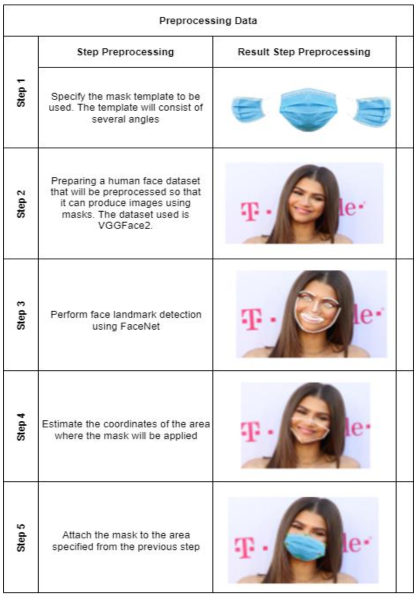

Figure 3. Step Preprocess FER2013 Dataset

Figure 3 shows an example result of the preprocessed FER2013 Masked version dataset. Data generated by this step is used to construct both for Preprocessed FER2013 Masked version Dataset and FER2013 Masked &UnMasked Dataset.

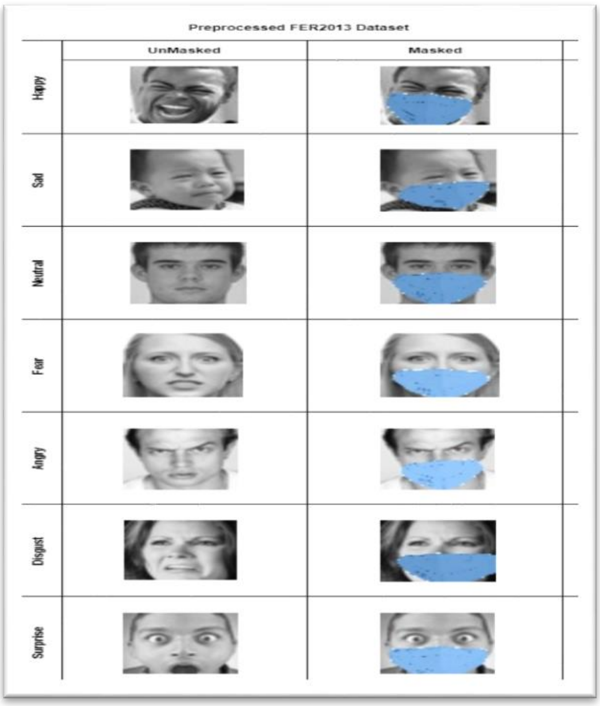

Figure 4. Result Preprocess FER2013 Dataset

2.2 Dataset

The dataset that will be used in this experiment is taken from a public dataset, namely the FER2013 dataset. However, because this dataset only consists of subject images without using a mask, then the data will be preprocessed first which already explained in section 2.1. Dataset used consist of 3 set of data:

1) **FER2013 Masked &UnMasked Dataset**
   This dataset is used for the 1st stage model to identify object use mask or not. The dataset used here is a combination of the FER2013 dataset and the preprocessed FER2013 dataset which using a mask. Then the

Copyright © 2023 **Kevin Kurniawan**, Page 104
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 5, Nomor 1, September 2023
e-ISSN 2685-998X
DOI 10.30865/json.v5i1.6768

results are reduced so that dataset become balanced between classes. Table 1 shows the quantity of each class of FER2013 Masked &UnMasked Dataset.

Table 1. FER2013 Masked &UnMasked Dataset Proportion

Class	%Train	%Test
Masked	50%	50.18%
UnMasked	50%	49.82%
Total	100%	100%

2) FER2013 UnMasked Dataset

This dataset is used for the 2nd stage modelto recognize expressions for unmasked data. This dataset is the original version of FER2013 dataset. Table 2 shows the proportion of each class.

Table 2. FER2013 UnMasked Dataset Proportion

Class	%Train	%Test
Angry	13.91%	13.34%
Disgust	1.51%	1.54%
Fear	14.27%	14.26%
Happy	25.13%	24.71%
Neutral	17.29%	17.17%
Sad	16.82%	17.37%
Surprise	11.04%	11.57%
Total	100%	100%

3) FER2013 Masked Dataset

This dataset is used for the 2nd stage modelto recognize expressions for masked data. This data is generated by a preprocessed step to add a mask to the FER2013 dataset. Table 3 shows the quantity of each class.

Table 3. FER2013 Masked Dataset Proportion

Class	%Train	%Test
Angry	9.26%	12.93%
Disgust	2.91%	1.73%
Fear	9%	13.01%
Happy	25.87%	27.83%
Neutral	22.14%	18.76%
Sad	12.35%	13.40%
Surprise	18.44%	12.31%
Total	100%	100%

2.3 Proposed Model

Model proposed in this experiment using 2 stage network,which is Mask/UnMasked Classifier & FER Classifier. Both stage networks are CNN based. The first model used to detect expressions without using a mask and the second model used to detect expressions when the subject is wearing a mask[15], [16]. The following is an explanation of each stage used :

1) First Stage:Mask/UnMasked Classifier

This model used to classify subjects whether using masks or not. If the subject is usinga mask, then the system will pass this data to the FER2013_Mask classifier, else it will pass to FER2013_UnMask classifier. Model used here using imagenet architecture and EfficientNet-B0 as a pre-trained model. Figure 5 shows how this experiment achieves this first stage model.

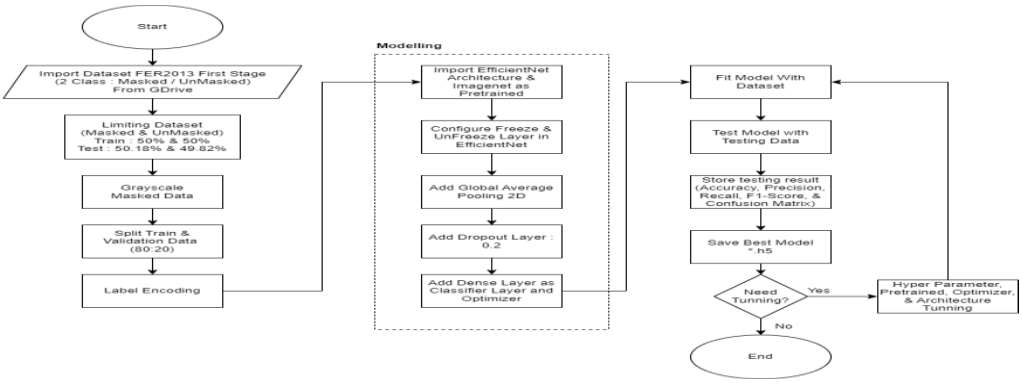

Figure 5. Mask/UnMask Step by Step Model

Copyright © 2023 **Kevin Kurniawan**, Page 105
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Jurnal Sistem Komputer dan Informatika (JSON) Hal: 102−111
Volume 5, Nomor 1, September 2023
e-ISSN 2685-998X
DOI 10.30865/json.v5i1.6768

As seen in figure 5, we first prepared FER2013 Masked &UnMasked as our dataset (section 2.2 point 1). After that we applied a greyscale process to ensure the color factor did not affect the mask detection process. As seen in figure 4, the mask used is blue, whereas the FER2013 dataset only consists of black and white. After that we do label encoding and start the model training process. In the experiments carried out, we used imagenet as the pretrained model and efficientnet as the model architecture. After going through the feature extraction and dropout layers, the data will be classified using a neural network layer. The parameters that we tune include hyper parameter tuning, freeze unfreeze layer, optimizer layer, architecture model.

2) **Second Stage:FER Classifier**
This stage consists of 2 networks, one to classify FER with mask, another for classify FER without mask. Those networks were neededbecause the region of interest between data mask and without mask is different. Model used here using EfficientNetV2-B3 as pretrained model as architecture, and imagenet as pretrained model. Figure 6 and 7 shows how this experiment achieves this second stage model: FER UnMask& FER Mask Model[17], [18].

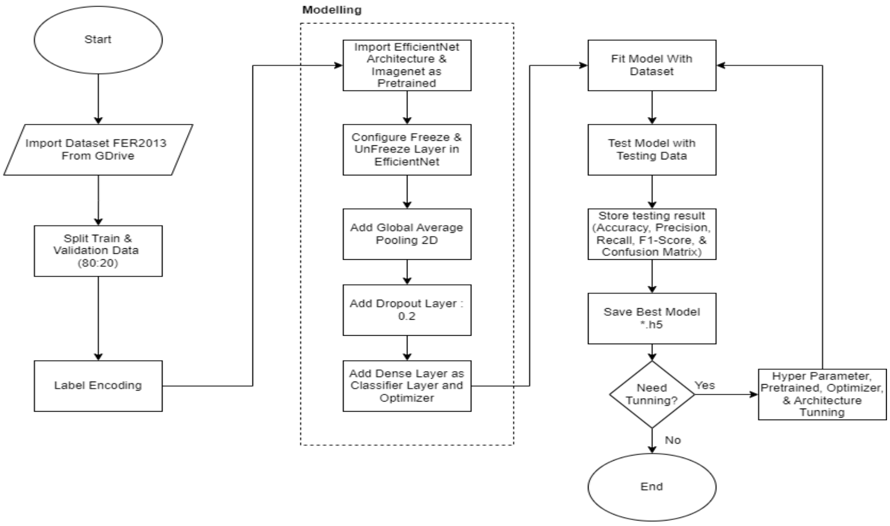

Figure 6. FER UnMask Step by Step Model

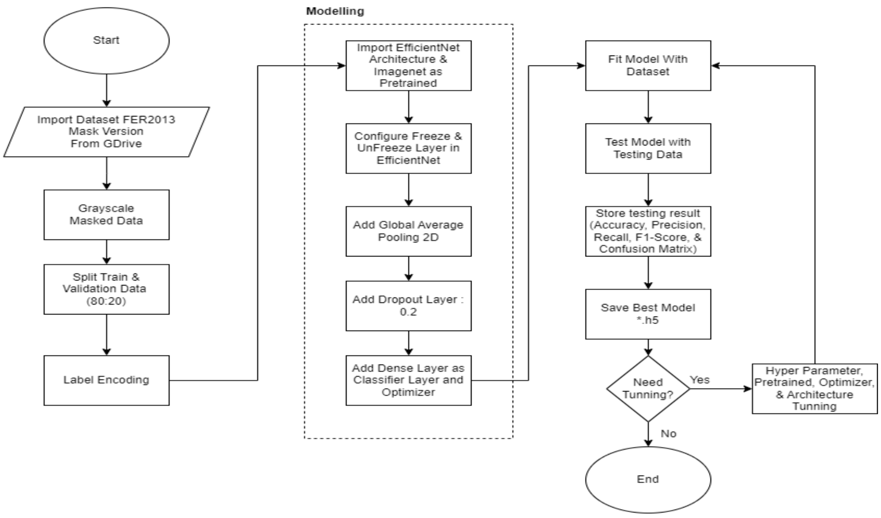

Figure 7. FER Mask Step by Step Model

As shown in figure 6 and figure 7, first we prepare a dataset to be used in the model creation process according to the purpose of each model. Figure 6 describes the process of forming the FER model for detecting expressions when not using a mask, therefore the dataset used is the original version of FER2013 (section 2.2 point 2). While for the process of forming the FER model using masks as reflected in Figure 7, the dataset used is used is the FER2013 version using a mask (section 2.2 point 3). After that, the difference in the modeling process lies in the implementation of greyscale in the FER Masked version to ensure that the color factor does not affect the facial expression detection process. After that we did label encoding and started the model training process. In the experiments, we used imagenet as the pretrained model and efficientnet as the model architecture. After going through the feature extraction and dropout layers, the data will be classified using a

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Jurnal Sistem Komputer dan Informatika (JSON) Hal: 102−111
Volume 5, Nomor 1, September 2023
e-ISSN 2685-998X
DOI 10.30865/json.v5i1.6768

dense layer. The parameters that we tune include hyper parameter tuning, freeze unfreeze layer, optimizer layer, architecture model.

3) **Ensemble Model (Proposed Model)**
After get best model for each scenario, then we try to assembled model to produce our proposedarchitecture. Figure 8 shows detail our proposed architecture model.

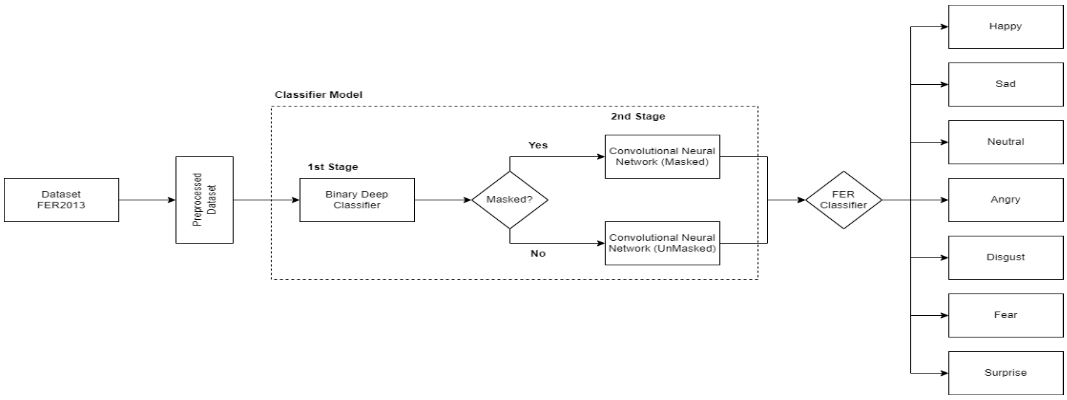

Figure 8. Proposed Architecture

As can be seen in Figure 8, the ensemble process is carried out by applying the first stage to detect whether the subject is wearing a mask or not. This model that will decide the FER process will be carried out by the FER Mask or FER UnMask model in the second stage process. This process is carried out so that the region of interest observed by the model can be right in the area that really needs to be observed. This idea aims to ensure that the entire model maintains its accuracy and effectiveness.

2.4 Comparation Model: Highest Probability Model

As previously explained, the final goal of this experiment is to find a combination of models that have good model performance with a good level of effectiveness. The effectiveness in question is how much the level of computational complexity will certainly affect the required computational time. Therefore, we try to create a scenario where the model we propose will be compared with a comparison scheme which we call the highest probability model[19], [20]. Which is the same as the proposed model, this scheme will use the best model from each FER scenario, namely FER Mask & FER UnMask (section 2.3 point 2). We do not use the first stage of the proposed model. The two FER models will run for each data, and then from the probability matrix produced by the two models, we will take the highest probability value for determining the decision on the expression being recognized. Figure 9 shows the scheme of the highest priority model used as a comparison model[21], [22].

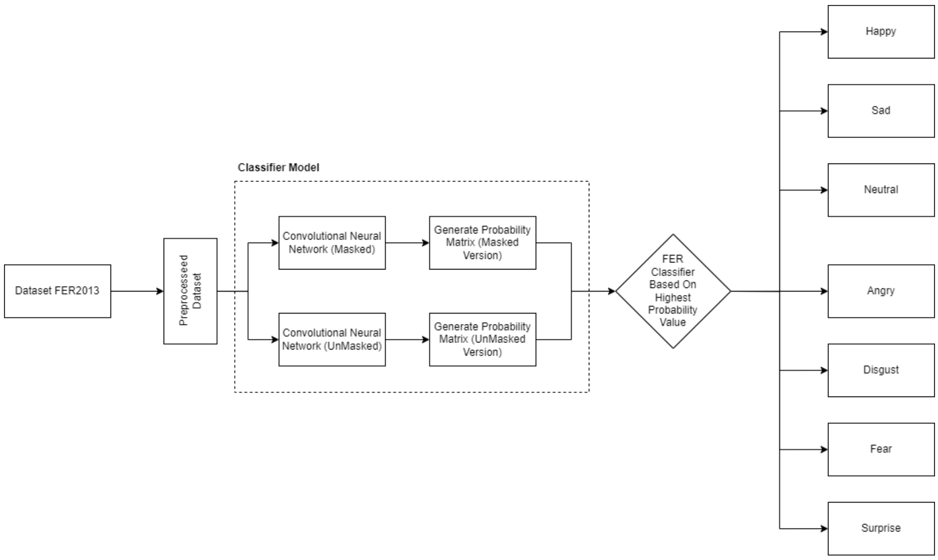

Figure 9. Highest Probability Architecture

3. RESULTS AND DISCUSSION

In this experiment, we use 2 phases of experiments. Phase 1 tries to get the best model for each scenario. The first scenario is used to detect the use of the mask and the second scenario to do FER with mask and FER without mask.

Copyright © 2023 **Kevin Kurniawan**, Page 107
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 5, Nomor 1, September 2023
e-ISSN 2685-998X
DOI 10.30865/json.v5i1.6768

Then Phase 2 tries to combine best model from phase 1 to build our proposed model and our comparison model. Detail explanation about those phases can be found at section 4.1 and 4.2.

3.1 Phase 1 – Best Model of Each Scenario

As described before, this experiment will produce 3 best models for each scenario to construct 2 stage network FER model. Those models are mask/unmasked model, FER UnMasked Model, and FER Masked Model. The following is a complete description of the best model results obtained from each scenario:

1) **CNN Model to Classify Mask or Unmask**
   This model is used to identify object data using a mask or not. As described before, we use imagenet as our architecture and EfficientNetB0 as our pretrained model. As a result, we get 99.85% accuracy. The result from this model will decide whether objects going to classify use FER with mask model or FER without mask model in our proposed model. Table 4 shows the result summary from the model and Table 5 shows the confusion matrix from this model.

Table 4. Detail Result of Masked/UnMasked Model

Expression	Precision	Recall	F1-Score
Masked	100%	99.87%	99.93%
UnMasked	99.87%	100%	99.93%

Table 5. Confusion Matrix Masked/UnMasked Model

Actual\predicted	Masked	Unmask
Masked	1603	2
Unmask	2	1591

2) **CNN Model to Classify Expression Mask**
   This model is used to classify expressions when objects are using Mask. As described before, we use imagenet as our architecture and EfficienetV2-B3 as our pretrained model. We also compare with some architecture and optimizer models shown at Table 6.

Table 6. FER Model Mask Comparison

Architecture	Optimizer	Accuracy	F1-Score
EfficientNetV2-B3	Adam	57.82%	52.30%
EfficientNet-B0	Adam	54.78%	52.64%
EfficientNet-B0	SGD	48.35%	40.57%
EfficientNetV2-B2	Adam	53.79%	50.06%

As we see in table 6, the best configuration for FER mask classification is with our proposed method using EfficientNetV2-B3 for our architecture and Adam Optimizer with 57.82% accuracy. Table 7 shows detail precision, recall, and F1-Score for each expression (class).

Table 7. Detail Result of FER Mask Model

Expression	Precision	Recall	F1-Score
angry	51.06%	41.12%	45.56%
disgust	48.42%	53.49%	50.83%
fear	54.26%	24.61%	33.87%
happy	72.05%	76.18%	74.06%
neutral	47.70%	71.32%	57.17%
sad	47.49%	27.07%	34.48%
surprise	61.01%	82.49%	70.15%

3) **CNN Model to Classify Expression Unmask**
   This model is used to classify expressions when an object is not using a mask. As described before we propose imagenet as our architecture and EfficienetV2-B3 as our pretrained model. We also compare with some architectures and optimizers shown at table 8.

Table 8. FER Model Without Mask Comparison

Architecture	Optimizer	Accuracy	F1-Score
EfficientNetV2-B3	Adam	68.08%	65.63%
EfficientNet-B0	Adam	65.31%	63.34%
EfficientNet-B0	SGD	63.15%	57.75%

Copyright © 2023 **Kevin Kurniawan**, Page 108
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 5, Nomor 1, September 2023
e-ISSN 2685-998X
DOI 10.30865/json.v5i1.6768

As we see in table 6, the best configuration for FER without mask classification is with our proposed method using EfficientNetV2-B3 for our architecture and Adam Optimizer with 68.08% accuracy. Table 9 shows detail precision, recall, and F1-Score for each expression.

Table 9. Detail Result of FER Without Mask Model

Expression	Precision	Recall	F1-Score
angry	52.27%	70.88%	60.17%
disgust	65.26%	55.86%	60.19%
fear	62.83%	45.90%	53.05%
happy	84.89%	88.39%	86.61%
neutral	62.04%	64.56%	63.28%
sad	58.90%	50.44%	54.34%
surprise	81.50%	82.19%	81.85%

3.2 Phase 2 – Assembled Model (Proposed Model vs Highest Probability Model)

In phase 2 we try to compare each scenario when data is assembled between masked & unmasked. Phase 2 experiment try to compare our proposed model with the comparison model namely highest probability model.

1) **CNN Model to Classify Expression from Highest Probability Recognition**
In this scenario, we do a test for the model scheme where the 2 best models from phase 1, Masked Model &UnMasked Model are asked to perform expression recognition simultaneously. Then the results of the probability matrix generated from those models will be compared based on their probability values to be able to determine which expression is considered most appropriate to the object/data used. From this scenario, quite good results are obtained with a model accuracy of 57.58%. Figure 7 is an illustrative explanation of the scenario that was carried out. Table 10 is the result of the experiments conducted in this scenario.

Table 10. Detail Result Highest Probability FER Model

Expression	Precision	Recall	F1-Score
angry	56.22%	40.13%	46.83%
disgust	52.27%	46.70%	49.33%
fear	56.32%	22.70%	32.35%
happy	71.00%	81.31%	75.81%
neutral	51.87%	59.10%	55.25%
sad	42.74%	57.43%	49.01%
surprise	68.36%	75.66%	71.82%

Table 11. Confustion Matrix Highest Probability FER Model

actual \ predicted	angry	disgust	fear	happy	neutral	sad	surprise
angry	642	35	84	160	244	356	79
disgust	30	92	8	16	16	26	9
fear	173	22	379	91	236	515	254
happy	47	4	11	2032	220	114	71
neutral	79	7	48	302	1279	394	55
sad	139	14	74	166	384	1098	37
surprise	32	2	69	95	87	66	1091

2) **CNN Model to Classify Expression Mask with First Stage Model (Proposed Model)**
This is our proposed model, we try to add the first stage for mask/unmask classifier to our model to identify whether data is classified with FER Mask or FER UnMask. This proposed model managed to get a good accuracy with an accuracy of 57.55%. Table 12 and table 13 show detailed results for each expression and confusion matrix of this model.

Table 12. Detail Result Proposed Model

Expression	Precision	Recall	F1-Score
angry	57.04%	40.50%	47.37%
disgust	50.00%	49.75%	49.87%
fear	53.88%	24.97%	34.12%
happy	72.35%	82.39%	77.04%
neutral	49.64%	63.08%	55.56%
sad	43.32%	51.57%	47.09%
surprise	68.86%	71.91%	70.35%

Copyright © 2023 **Kevin Kurniawan**, Page 109
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Table 13. Confustion Matrix Proposed Model

actual \ predicted	angry	disgust	fear	happy	neutral	sad	surprise
angry	648	39	96	161	279	299	78
disgust	31	98	7	14	19	23	5
fear	164	28	417	110	267	463	221
happy	31	6	17	2059	235	73	78
neutral	62	8	55	252	1365	368	54
sad	163	15	86	151	478	986	33
surprise	37	2	96	99	107	64	1037

From the results of the comparison above between the highest probability value model and our proposed model, it was found that the highest probability model gave slightly better results in terms of accuracy compared to the proposed model. Table 14 shows us head-to-head comparison between those models.

Table 14. Highest Probability Model vs Proposed Model

Model	Accuracy	Processing Time
Highest Probability Model	57.58%	29.95s
Proposed Model	57.55%	25.58s

As can be seen from the results above, the highest probability model has results that outperform the proposed model. There is a performance increase of 0.03%, but it should be noted that this performance increase is also followed by a much greater increase in cost computing compared to our proposedmodel in this experiment. Table 15 is a comparison of the required cost computing between scenario 1 and scenario 2 in terms of the number of parameters required based on the pretrained model used.

Table 15. Cost Computing Comparison

Model	Description	Pretrained Used	QtyParam	Total Param
Highest Probability Model	FER Masked Model	EfficientNetV2-B3	12M	24M
	FER UnMasked Model	EfficientNetV2-B3	12M	
Proposed Model	Mask/UnMask Model	EfficientNet-B0	5.3M	17.3M
	FER Mask/UnMask Model	EfficientNetV2-B3	12M	

It can be seen in table 15, where there is a cost computing efficiency of 28% with param differences of up to 6.7 million parameters. This is also evidenced from table 14, which shows the processing time of the proposed model is up to 15% faster which difference of almost 5 seconds compared to the highest probability model. It can be concluded that our proposed model has a better level of performance and efficiency compared to the model performed in highest probability model.

4. CONCLUSION

From the overall results of the experiments, it shows how the facemask become a big problem on the results of the accuracy of facial expression recognition. Where in the phase 1 experiment, it shows that if a FER model tries to recognize a face without a mask, the results obtained are quite good with an accuracy of 68.08%. However, for facial recognition cases using masks, the maximum accuracy obtained is only 57.82%. Keep in mind that these results were obtained without combining data between FER2013 data without masks and FER2013 with masks. After that, in the second phase of the experiment, we try to compare our proposed model with the comparison model called highest probability model. Where in this phase, dataset used is assembled data which include FER2013 Masked andUnMasked FER2013. From the results it can be concluded that the highest probability model provides slightly better accuracy compared to the proposed model. Highest probability model obtains an accuracy of 57.58% while the proposed model obtains an accuracy of 57.55%. However, this difference in accuracy pays off with computational efficiency that is much lighter in the proposed model, where there is a cost computing efficiency of up to 28% based on the pretrained parameters used and the processing time efficiency of 15% with difference of up to 5 seconds using the proposed model. So, it can be concluded that the proposed model has better performance and efficiency than the comparison scenario. In carrying out this experiment, several obstacles were also found which might become suggestions for future research, including the imbalance quantity of data from each expression, the preprocessing carried out could not cover all of the existing FER2013 data which resulting in the number of FER2013 using masks not the same in quantity as FER2013 without a mask, and also the classifier used in the future might be able to start applying attention mechanisms or use other algorithms so that the overall accuracy results can also get better.

Copyright © 2023 **Kevin Kurniawan**, Page 110
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Volume 5, Nomor 1, September 2023
e-ISSN 2685-998X
DOI 10.30865/json.v5i1.6768

REFERENCES

[1] P. Saxena, “Real-time emotion recognition: Potential use cases and challenges,” indiaai.gov.in, 2021. .
[2] Tinaliah and T. Elizabeth, “Penerapan Convolutional Neural Network Untuk Klasifikasi Citra Ekspresi Wajah Manusia Pada MMA Facial Expression Dataset,” J. Tek. Inform. dan Sist. Inf., vol. 8, no. 4, pp. 2051–2059, 2021, doi: 10.35957/jatisi.v8i4.1437.
[3] S. Minaee, M. Minaei, and A. Abdolrashidi, “Deep-emotion: Facial expression recognition using attentional convolutional network,” Sensors, vol. 21, no. 9, pp. 1–16, 2021, doi: 10.3390/s21093046.
[4] S. T. Zulkarnain and N. Suciati, “Modified Local Ternary Pattern With Convolutional Neural Network for Face Expression Recognition,” JUTI J. Ilm. Teknol. Inf., vol. 19, no. 1, p. 10, 2021, doi: 10.12962/j24068535.v19i1.a1031.
[5] R. Septian, D. I. Saputra, and S. Sambasri, “Klasifikasi Emosi Menggunakan Convolutional Neural Networks,” 2019, doi: 10.13733/j.jcam.issn.2095-5553.2022.11.026.
[6] L. Pham, T. H. Vu, and T. A. Tran, “Facial expression recognition using residual masking network,” Proc. - Int. Conf. Pattern Recognit., pp. 4513–4519, 2020, doi: 10.1109/ICPR48806.2021.9411919.
[7] Syahromi, “Facial Expression Recognition Menggunakan Algoritma Mtcnn (Multi-Task Cascaded Convolutional Neural Network) Dan Vgg16,” Apl. dan Anal. Lit. Fasilkom UI, pp. 4–25, 2022.
[8] Ministry of Health, “Peta Sebaran,” Https://Covid19.Go.Id/Peta-Sebaran, p. https://covid19.go.id/peta-sebaran, 2022.
[9] Haeruddin, Herman, and Patrick Pratama Hendri, “PENGEMBANGAN APLIKASI EMOTION RECOGNITION DAN FACIAL RECOGNITION MENGGUNAKAN ALGORITMA LOCAL BINARY PATTERN HISTOGRAM (LBPH) DAN CONVOLUTIONAL NEURAL NETWORK (CNN),” J. Teknol. Terpadu, vol. 9, no. 1, pp. 49–55, 2023.
[10] B. Yang, J. Wu, and G. Hattori, “Facial Expression Recognition with the advent of face masks,” ACM Int. Conf. Proceeding Ser., pp. 335–337, 2020, doi: 10.1145/3428361.3432075.
[11] G. Castellano, B. De Carolis, and N. Macchiarulo, “Automatic emotion recognition from facial expressions when wearing a mask,” ACM Int. Conf. Proceeding Ser., 2021, doi: 10.1145/3464385.3464730.
[12] A. Anwar and A. Raychowdhury, “Masked Face Recognition for Secure Authentication,” pp. 1–8, 2020.
[13] A. Lioga Seandrio, A. Hendrianto Pratomo, and M. Y. Florestiyanto, “Implementation of Convolutional Neural Network (CNN) in Facial Expression Recognition,” Telemat. J. Inform. dan Teknol. Inf., vol. 18, no. 2, pp. 211–221, 2021, doi: 10.31515/telematika.v18i2.4823.
[14] M. Fakhrudin and A. Fatoni, “Pengaruh Risiko Spesifik Bank dan Good Corporate Governance Terhadap Stabilitas Bank Pembiayaan Rakyat Syariah,” vol. 4, no. 1, pp. 116–127, 2023, doi: 10.47065/jtear.v4i1.734.
[15] D. Alamsyah and D. Pratama, “Implementasi Convolutional Neural Networks (CNN) untuk Klasifikasi Ekspresi Citra Wajah pada FER-2013 Dataset,” J. Teknol. Inf., vol. 4, no. 2, pp. 350–355, 2020, doi: 10.36294/jurti.v4i2.1714.
[16] Lia Farokhah, “Implementasi Convolutional Neural Network untuk Klasifikasi Variasi Intensitas Emosi pada Dynamic Image Sequence,” J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 4, no. 6, pp. 1070–1076, 2020, doi: 10.29207/resti.v4i6.2644.
[17] S. Sriyati, A. Setyanto, and E. E. Luthfi, “Literature Review: Pengenalan Wajah Menggunakan Algoritma Convolutional Neural Network,” J. Teknol. Inf. dan Komun., vol. 8, no. 2, 2020, doi: 10.30646/tikomsin.v8i2.463.
[18] P. Musa, W. K. Anam, S. B. Musa, W. Aryunani, R. Senjaya, and P. Sularsih, “Pembelajaran Mendalam Pengklasifikasi Ekspresi Wajah Manusia dengan Model Arsitektur Xception pada Metode Convolutional Neural Network,” Rekayasa, vol. 16, no. 1, pp. 65–73, 2023, doi: 10.21107/rekayasa.v16i1.16974.
[19] Y. Fajriyanti, A. H. Rahmah, S. Ulfa, and E. Hadiyanti, “Analisis Motivasi Kerja Generasi Z Yang Dipengaruhi Oleh Lingkungan Kerja dan Komitmen Kerja,” vol. 4, no. 1, pp. 107–115, 2023, doi: 10.47065/jtear.v4i1.808.
[20] S. P. Ristiawanto, B. Irawan, and C. Setianingsih, “PENGENALAN EKSPRESI WAJAH BERBASIS CONVOLUTIONAL NEURAL NETWORK MENGGUNAKAN ARSITEKTUR RESIDUAL NETWORK-50,” e-Proceeding Eng., vol. 8, no. 5, pp. 6442–6454, 2021.
[21] P. P. Kusdiananggalih, E. Rachmawat, and Risnandar, “Pengenalan Ekspresi Wajah Menggunakan Deep Convolutional Neural Network,” J. Tugas Akhir Fak. Inform. e-Proceeding, vol. 8, no. 2, p. 75, 2021, doi: 10.35200/explore.v11i2.478.
[22] J. Omar, N. H. Shabrina, A. N. Bhakti, and A. Patria, “Emotion Recognition Using Convolutional Neural Network (CNN),” Ultim. Comput. J. Sist. Komput., vol. 13, no. 1, 2021, doi: 10.1088/1742-6596/1962/1/012040.

This Journal is licensed under a Creative Commons Attribution 4.0 International License