SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12785
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Optimizing Gender Classification Accuracy in Facial Images Using Data Augmentation and Inception V-3

Juliansyah Putra Tanjung [1)*], Mhd. Rio Faldi[2)], Haggai Sitompul [3)], Muhammad Ridho[4)], Jojor Putri Ambarita [5)]
[1,2,3,4,5)]Universitas Prima Indonesia, Medan, Indonesia,
[1)]juliansyahputratanjung@unprimdn.ac.id, [2)]mhdriofaldixtsm2@gmail.com,
[3)] haggaisitompul@gmail.com, [4)]arrafiimridho@gmail.com, [5)]jojorputria@gmail.com

Submitted : Jul 26, 2023 | **Accepted** : Aug 11, 2023 | **Published** : Oct 1, 2023

Abstract: In the digital era, facial recognition technology plays a crucial role in various applications, including gender classification. However, challenges such as variations in expressions and face positions, as well as differences in features between men and women, make this task formidable. This study aims to enhance the accuracy of gender classification using the Inception V-3 method and the Convolutional Neural Network (CNN), along with data augmentation techniques. The Inception V-3 method was chosen for its superiority in accuracy and speed. In contrast, the CNN model was selected in this study as a comparison and due to its algorithmic advantages in learning and extracting high-level features from images, including facial images, which are crucial for tasks such as gender classification. The data augmentation techniques in this study include rescaling, rotation, width and height shifts, shear range, zoom, horizontal flip, and fill method for model accuracy in gender classification with a small dataset. The study results indicate that the Inception V-3 model provides better accuracy (99.31%) in gender classification compared to the CNN model (81.31%). This conclusion underscores that the use of the Inception V-3 method with data augmentation techniques can improve the accuracy of gender classification in facial images.

Keywords: Gender Classification; Facial Recognition Technology; Data Augmentation; Inception V-3; CNN.

INTRODUCTION

In this digital age, facial recognition technology has become an integral part of many applications, such as security, user identification, and emotion recognition. One recurrent issue in facial recognition is gender classification. Gender classification based on faces holds potential benefits, like personal service optimization, market analysis, and social research. However, gender classification based on faces still poses a challenging task (Asmara et al., 2018). Several challenges need to be solved to improve gender classification accuracy. One of them is variations in facial expressions and head positions. The human face can express a multitude of emotional expressions, which can affect gender classification outcomes. Besides, differing head positions can cause variations in lighting and viewing angles, which can impact the system's ability to accurately recognize gender. Furthermore, differences in facial features between men and women also pose a challenge to gender classification (Asmara et al., 2018).One of the algorithms that can be used for gender classification of facial photos is the Naive Bayes method (Asmara et al., 2018). This algorithm can differentiate female and male faces based on features extracted

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12785
e-ISSN : 2541-2019
p-ISSN : 2541-044X

with the Principal Component Analysis method. There's also the SVM algorithm (Wahyu et al., 2019) which uses machine learning algorithms that can separate two classes by finding the best hyperplane that maximizes the margin between the two classes. Moreover, there's the CNN algorithm (Arifandi, 2022) which uses an artificial neural network consisting of multiple layers of convolution, pooling, and fully connected, which can learn high-level features on facial images.

After comparing with various algorithms, Inception V-3 (Selitskiy et al., n.d.) and CNN is decided to be used in this study. The Inception V-3 method was chosen in this study due to its advantages in accuracy and speed, while CNN serves as a comparison to determine whether Inception V-3 is superior to the standard CNN method. The Inception V-3 method (Selitskiy et al., n.d.) is a convolutional artificial neural network architecture that can recognize complex features in facial images more efficiently and effectively. This method can also address overfitting and underfitting problems by using data augmentation techniques. One way to enhance gender classification accuracy based on facial images is by using data augmentation techniques. Some examples of data augmentation that can be applied to facial images include rotation, translation, scaling, mirroring, lighting, and noise (Sabili et al., 2021).

With an increase in gender classification accuracy based on faces, facial recognition technology applications can provide more reliable and accurate results. This would allow the development of more efficient and reliable systems in various fields, such as security, user recognition, and market analysis.

LITERATURE REVIEW

Gender classification based on image deep learning classification is a subfield of computer vision that utilizes deep learning algorithms to analyze image data and predict the gender of individuals in the images. This classification includes using a face image as an input and facial feature extraction techniques such as Artificial Neural Network methods, Deep Learning, and so forth. Gender Classification specifically aims to distinguish men's and women's faces based on extracted features (Rosiani et al., 2019). A Convolutional Neural Network (CNN) is a type of Neural Network used for image processing. CNN uses the concept of convolution, which is the process of data processing by shifting filters or kernels on every part of the image. Each filter is used to capture specific features from the image, like lines, contours, or colors. This process is repeated on each layer of the CNN, making the detected features more complex and abstract on the higher layers (Ajit et al., 2020).

On the other hand, Inception V-3 is a CNN architecture developed by Google. This architecture excels in accuracy and speed in pattern recognition on images. Inception V-3 is used in this study because it can efficiently and effectively recognize complex features in facial images. The stages of Inception V-3 include the following (Jiang, 2020). Building upon the existing body of work, this study takes an innovative approach in pneumonia classification by using two similar models namely Convolutional Neural Network (CNN) and Inception V-3 model and comparing it directly with CNN model. Previous studies have exhibited strong performance utilizing either model independently but have not explored their combined use in detail. For instance, (Ali, 2023) using a Chest CT-Scan Dataset with VGG-19 model and surpassed the pre-built InceptionV3 extension, achieving an accuracy of 94.88 %. (Setiawan, 2020) used the VGG algorithm but this study conducted a comparison of CNN architectures for fundus image classification. Test results in both scenarios showed the best architecture to be VGG19 and VGG16. (Zhou et al., 2021) developed a customized VGG19 model for Pneumonia Detection and was tested using different classifiers, such as SVM-linear, SVM-RBF, KNN classifier, Random-Forest (RF), and Decision-Tree (DT) but not using CNN and Inception V-3 models like the one used in this research. Overall, our research further analyzing a pneumonia classification model through analyzing the CNN and Inception V-3 models in classifying Pneumonia based on Lung CT-Scan Images, with a clear emphasis on methodological clarity like preprocessing, hyperparameter, and thorough performance evaluation. This allows for a more understanding of the two models with similar deep learning techniques but yield different results.

METHOD

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i4.12785 p-ISSN : 2541-044X

This study employs a methodological approach that combines a range of techniques to perform gender classification on a dataset of synthetic facial images. Initially, the raw data undergoes preprocessing steps to standardize image size and apply data augmentation techniques for the diversification of the training dataset. Following this, the Inception V-3 model, a sophisticated Convolutional Neural Network (CNN), is leveraged to discern patterns in the augmented data and create a robust classifier. The overall methods of this experiment are as shown in the Fig. 1 below:

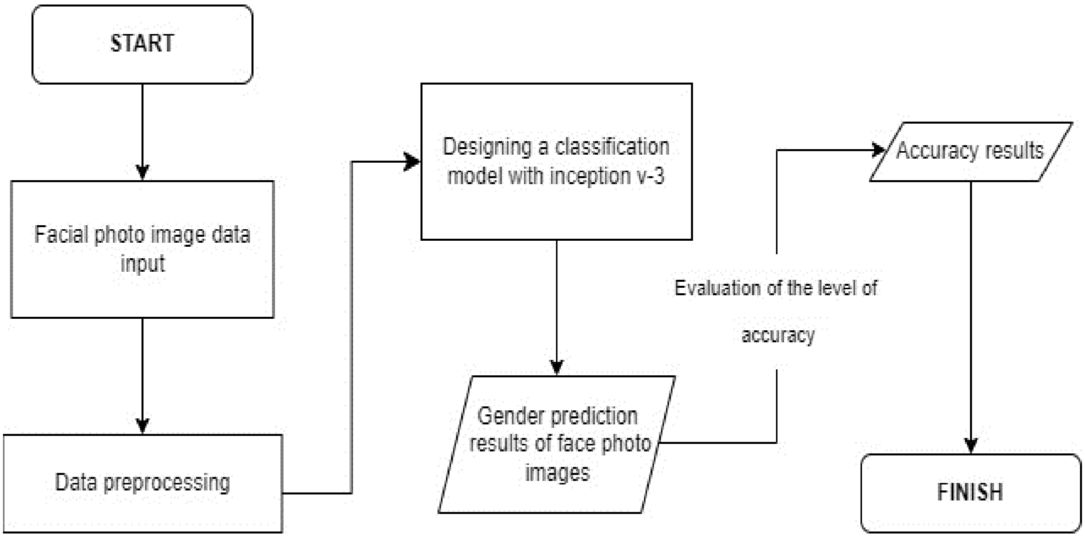

Fig 1. Data Gathering Research Stages

The stages of this research use the common Inception V-3 modeling stages used among others as follows:

Data input

The research begins with collecting a dataset consisting of facial photo images with a known gender label. The dataset used consists of 2444 male and female portraits made using a computer program called Stable Diffusion v.1.4.This dataset includes portraits of men and women, and is considered valid,the portrait must meet several criteria. Each portrait has a size of 512 x512 pixels. The following is a sample display of the dataset under study with labels gender,namely womenFig

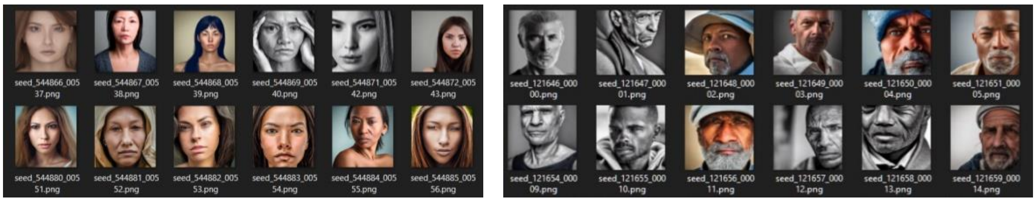

Fig 2. Face Photo Dataset

Preprocessing Data

The next step is to preprocess the data accordingly 7the size of the face image data input so that they have the same size. Then process Data Augmentation is also done by changing the length, width,rotation, zoom, and others to improve the performance of classification results accuracy.Data augmentation is the process of adding variations to the training data by how to change or modify the original data without changing the label. Data Augmentation can also improve the ability of the model to recognize complex and varied patterns in facial images

Classification Model Design with Inception V-3

After the facial photo data is collected, the next stage is designing classification model using Inception V-3 architecture. Inception V-3 is a convolutional neural network known for its capabilities in recognizing complex features in images. First of all the data will be trained with the set of datasets used for the training process, then the data set used for the testing process will be used for produce output in the form of predictive results which will calculate the level of accuracy.

Evaluation of Accuracy Level

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12785
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The evaluation stage is carried out to measure the performance of the model prediction results classification in this study and produce output in the form of accuracy. Commonly used evaluation metrics include accuracy. Measuring accuracy the extent to which the Inception V-3 model can correctly classify gender based on facial images. The results of this evaluation will provide an overview about the performance of the designed classification model. Here is the formulaused to evaluate accuracy.

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{The\ correct\ number\ of\ facial\ image\ clasifications}}{\boldsymbol{Total\ number\ of\ clasifications}} \qquad (1)$$

With four possible results namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), then the accuracy formula becomes:

$$\boldsymbol{Accuracy} = \frac{(\boldsymbol{TP+TN})}{(\boldsymbol{TP+TN+FP+FN})} \qquad \textbf{(2)}$$

RESULT

In this section, the researcher will explain the results of the research obtained. Researchers can also use images, tables, and curves to explain the results of the study. These results should present the raw data or the results after applying the techniques outlined in the methods section. The results are simply results; they do not conclude.

Table 1. Inception V-3 Confusion Matrix Test Result

	Female Prediction	**Male Prediction**
Actual Female	145	1
Actual Male	1	142

Table 2. CNN Confusion Matrix Test Result

	Female Prediction	**Male Prediction**
Actual Female	117	29
Actual Male	25	118

Table 3. Parameter Augmentation

Based on the Confusion Matrix results of both Inception V-3 CNN model in classifying the test data, we

Augmentasi Parameter	Value	Description
Rescale	1.0/255.0	Image scalleng by divinding each pixel with 255 so the value pixell are between 0 and 1
Rotation Range	45	The image will be rotate randomly at angle between 0 and 1 45 degrees
Width,Shift,Range	0.2	The image will be shifted horizontally by 20 % of the image width
Height shift Range	0.2	The image will be shifted Vertically as far as 20 % of the image heigth
Shear_Range	0.2	Shear transformation is transformation that skew the shape picture 0.2 means shear angel can be -20 to between 20 degree
Zoom_Range	0.2	The image enlarged or reduced randomly as much as 20%
Horizontal Flip	True	The image will be flipped horizontally random
Fill mode	‘Nearest’	The stategy used to fill new pixell that may appear during the transformation ‘***nearest’*** means the new pixell is filled With the nearest pixel

can conclude the accuracy of Inception V-3:

$$Inception\ V3\ Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} = \frac{(145+142)}{(145+142+1+1)} = 99.31\% \qquad (2)$$

Furthermore, we can conclude the Accuracy of CNN based on the Confusion Matrix result of the CNN

*name of corresponding author

model:

$$CNN\ Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} = \frac{(117+118)}{(117+118\ +29+25)} = 81.31\% \tag{3}$$

Based on the Confusion Matrix from the evaluation results of this study, it can be concluded that the Inception V-3 model has an accuracy of 99.31%. Meanwhile, the CNN model has an accuracy of 81.31%. Accuracy is a measure that describes how well a model is at predicting the correct class which is the gender. The higher the accuracy value, the better the model is at making predictions. In terms of accuracy, the Inception V-3 model performs better compared to the CNN model. This difference indicates that the implementation of the Inception V-3 method with data augmentation techniques produces better performance in gender classification compared to the CNN model. This not only shows that the use of data augmentation is effective in improving gender classification accuracy on face photos for both models, but also that Inception V-3 is much better than CNN model in the Gender Classification based on facial images experiment conducted in this research.

Data processing in this study only changes the size of the image to size which can be well received by the Inception V-3 model, namely 255x255 pixels.Then next, the Data Augmentation process is carried out which makes the image based on existing images in the training dataset. This technique helps increase the amount of data and reduce overfitting. Here's an explanation augmentation data parameters used in this study. All of the augmentations above apply to training data only. Meanwhile, data validation and testing are not carried out by Data Augmentation because Data Augmentation is only useful for the training process.

All of the augmentations above apply to training data only. Meanwhile, data validation and testing are not carried out by Data Augmentation because of Data Augmentation only useful for the training process.
This study uses the Inception V3 model as the basic model and conducts hyperparameter configuration or also known as fine-tuning the model to complete a special task, namely the classification of male and female photos.

Table 4. Hyperameter

hyperameter	**value**
*Input*_shape	(image_width, image_height,3)
Include_top	False
Weight	None
Layer trainable	False
Layer_Dance	1024 units'relu' activation
Layer_Dropout	0,2
Layer_Dense (*output layer*)	1 unit 'sigmoid; activation
Optimizer	RMSproot(0.000)
Loss	'Binary_Cronsetropy'
Metrics	***'Accuracy'***

Explanations of each of these hyperparameters include:
Input_shape = (image_width, image_height, 3): This hyperparameter defines the form of the input entered into the model. In this case, the input form are the image width, image height, and the number of color channels (RGB). Input shapeit is important to adjust to the dataset used.
Include_top = False: This hyperparameter is used to determine whether we want to include or not the topmost classification layer in the mode have been trained before. In this case, we're excluding layers because we're going to be adding our own custom layer.
Weights = None: Since we will load the weights from a local file, we initialize them model weight as None

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12785
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Layer.trainable = False: After loading the weights from the local file, we freeze layers of a pre-trained model. This means that the weights are on this layer will not be updated during the training process.
Layers.Dense(1024, activation='relu'): Adds a fully connected layer (Dense) with 1024 hidden units and ReLsU activation function.
Layers.Dropout(0.2): Dropout is a regularization technique in which multiple nodeshe inner layer is removed randomly with the aim of avoiding overfitting. In In this case, we use a dropout rate of 0.2.
Layers.Dense(1, activation='sigmoid'): Adds an output layer for binary classification. Since this is a binary classification, we only need one output unit with sigmoid activation function.
Optimizer = RMSprop(learning_rate=0.0001): RMSProp is an algorithm optimization used to update the weights in the model. We using a learning rate of 0.0001
Loss = 'binary_crossentropy': Since this is a binary classification problem, we using binary cross entropy as the loss function.
Metrics = ['accuracy']: The metrics used to evaluate the model

After the dataset is processed, then the next 1866 training data will be used for train the model and 289 validation data will be used to validate and 289 data testing to test the model. The division has been done from the dataset,so there is no need to do the process of dividing the dataset in this study.

Based on the test results of the confusion matrix data, the Inception V-3 model has been designed to classify gender in photos has shown that performance Very good. The confusion matrix is a tool that is often used for describe the performance of the classification model.

Table 5. Inception Model Evaluation Table V-3

	Women's Prediction	Man's Prediction
Actual Man	145	1
Actual Male	1	145

Table 6. CNN Model Evaluation Table

	Women's Prediction	Man's Prediction
Actual Man	117	29
Actual Male	25	118

Based on the prediction truth table from the evaluation results of this study, it can be conclusion, there are four prediction values of the Confusion Matrix in Training Data, including:
True Positives (TP): This is the case when the model predicts positive (in this case, 'Woman') and is actually positive. In our case, TP = 145.
True Negatives (TN): This is the case when the model predicts negative (in this case, 'Male') and is actually negative. In our case, TN = 142.
False Positives (FP): This is the case when the model predicts positive and actually negative. In our case, FP = 1.
False Negatives (FN): This is the case when the model predicts a negative and actually positive. In our case, FN = 1.
Thus the accuracy of the Inception V-3 model is as follows:

$$Akurasi = \frac{TP + TN}{(TP + TN + FP + FN} = \frac{145 + 142}{(145 + 142 + 1 + 1} = 99{,}31\% \qquad (3)$$

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12785
e-ISSN : 2541-2019
p-ISSN : 2541-044X

While the Test Data, among others:
True Positives (TP): This is the case when the model predicts positive (in this case, 'Woman') and is actually positive. In our case, TP = 117.
True Negatives (TN): This is the case when the model predicts negative (in this case, 'Male') and is actually negative. In our case, TN = 118.
False Positives (FP): This is the case when the model predicts positive and actually negative. In our case, FP = 29.
False Negatives (FN): This is the case when the model predicts a negative and actually positive. In our case, FN = 25.

Thus the accuracy of the CNN model is as follows:

$$Akurasi = \frac{TP + TN}{(TP + TN + FP + FN} = \frac{117 + 118}{(145 + 142 + 29 + 25} = 81{,}31\% \qquad (4)$$

DISCUSSIONS

In this section, the researchers can give a simple discussion related to the results of the research trials. Deep feature learning for gender classification with covered/camouflaged faces oleh Alghaili et al A Machine Learning Model for Estimation and Anemia Classification By Elkenawy E. Behavioral biometric data analysis for gender classification using feature fusion and machine learning

CONCLUSION

Based on this research, it can be concluded that the implementation of the Inception V-3 and CNN methods for gender classification in face photos is carried out with data augmentation techniques used to improve gender classification accuracy. In terms of accuracy, the Inception V-3 model performs better with an accuracy of 99.31%, while the CNN model has an accuracy of 81.31%. This difference indicates that the use of the Inception V-3 method with data augmentation techniques gives better results in gender classification in face photos compared to the CNN model. This research shows that the implementation of the Inception V-3 method with data augmentation techniques provides better performance in gender classification in face photos compared to the CNN model. The use of data augmentation is effective in improving the accuracy of gender classification in face photos and becomes a potential strategy to be applied in the development of gender recognition systems in face images. Based on this research, it is suggested to utilize other models to further test the reliability of Inception V-3 model in Gender Classification based on facial images

ACKNOWLEDGMENT

The authors would like to extend their gratitude to Bwandowando for sharing, creating and labeling the Synthetic faces dataset generated by Stable Diffusion v.1.4, which significantly contributed to this research.

REFERENCES

Ahmed, T. U., Hossain, S., Hossain, M. S., ul Islam, R., & Andersson, K. (2019). Facial expression recognition using convolutional neural network with data augmentation. *2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (IcIVPR)*, 336–341.

Ajit, A., Acharya, K., & Samanta, A. (2020). A Review of Convolutional Neural Networks. *International Conference on Emerging Trends in Information Technology and Engineering, Ic-ETITE 2020*. https://doi.org/10.1109/ic-ETITE47903.2020.049

Ali, M. S. (2023). *CHEST X-RAY BASED PNEUMONIA CLASSIFICATION USING VGG-19 CHEST X-RAY BASED PNEUMONIA CLASSIFICATION USING VGG-19*. *January*. https://doi.org/10.17605/OSF.IO/M2FPD

*name of corresponding author

Arifandi, A. (2022). Identifikasi dan Prediksi Umur serta Jenis Kelamin Berdasarkan Citra Wajah Menggunakan Algoritma Convolutional Neural Network (CNN). *RAINSTEK: Jurnal Terapan Sains & Teknologi*, *4*(2), 89–96.

Asmara, R. A., Andjani, B. S., Rosiani, U. D., & Choirina, P. (2018). Klasifikasi Jenis Kelamin Pada Citra Wajah Menggunakan Metode Naive Bayes. *Jurnal Informatika Polinema*, *4*(3), 212.

Bharadwaj, Y. S. S. (2021). Advanced Deep Learning Techniques. *Advanced Deep Learning for Engineers and Scientists: A Practical Approach*, 145–181.

Budiman, M. A. (2021). The effect of audit opinions, implementation of audit recommendations, and findings of state losses on corruption levels within ministries and institutions in the Republic of Indonesia. *Jurnal Tata Kelola Dan Akuntabilitas Keuangan Negara*, *7*(1), 113–129.

Jiang, Z. (2020). Face gender classification based on convolutional neural networks. *2020 International Conference on Computer Information and Big Data Applications (CIBDA)*, 120–123.

Muhathir, Dwi Ryandra, M. F., Syah, R. B. Y., Khairina, N., & Muliono, R. (2023). Convolutional Neural Network (CNN) of Resnet-50 with Inceptionv3 Architecture in Classification on X-Ray Image. *Computer Science On-Line Conference*, 208–221.

Punia, S. K., Kumar, M., Stephan, T., Deverajan, G. G., & Patan, R. (2021). Performance analysis of machine learning algorithms for big data classification: Ml and ai-based algorithms for big data analysis. *International Journal of E-Health and Medical Communications (IJEHMC)*, *12*(4), 60–75.

Rosiani, U. D., Asmara, R. A., & Laeily, N. (2019). Penerapan Facial Landmark Point Untuk Klasifikasi Jenis Kelamin Berdasarkan Citra Wajah. *Jurnal Informatika Polinema*, *6*(1), 55–60.

Sabili, S., Rachmadi, R. F., & Yuniarno, E. M. (2021). Verifikasi Wajah Menggunakan Deep Metric Learning pada Data Wajah dengan Disparitas Umur yang Besar. *Jurnal Teknik ITS*, *10*(2), A432–A437.

Selitskiy, S., Christou, N., & Selitskaya, N. (n.d.). *Learning Incorrect Verdict Patterns with the Meta-learning Supervisor ANN on the Established Face Recognising CNN Models*.

Setiawan, W. (2020). Perbandingan Arsitektur Convolutional Neural Network Untuk Klasifikasi Fundus. *Jurnal Simantec*, *7*(2), 48–53. https://doi.org/10.21107/simantec.v7i2.6551

Wahyu, M., Santoso, B., Wihandika, R. C., & Rahman, M. A. (2019). Ekstraksi Ciri untuk Klasifikasi Jenis Kelamin berbasis Citra Wajah menggunakan Metode Compass Local Binary Patterns. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *3*(11).

Zhou, Y., Chang, H., Lu, Y., Lu, X., & Zhou, R. (2021). Improving the Performance of VGG through Different Granularity Feature Combinations. *IEEE Access*, *9*, 26208–26220. https://doi.org/10.1109/ACCESS.2020.3031908

*name of corresponding author