International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 11, No.1 (Jan-2022)

https://dx.doi.org/10.12785/ijcds/110116

Gender Classification on Video Using FaceNet Algorithm and Supervised Machine Learning

Faisal Dharma Adhinata[1] **and Apri Junaidi**[2]

[1]*Department of Software Engineering, Faculty of Informatics, Institut Teknologi Telkom Purwokerto, Indonesia*
[2]*Department of Data Science, Faculty of Informatics, Institut Teknologi Telkom Purwokerto, Indonesia*

E-mail address: faisal@ittelkom-pwt.ac.id, apri@ittelkom-pwt.ac.id

Received 28 Mar. 2021, Revised 20 Jun. 2021, Accepted 23 Jun. 2021, Published 9 Jan. 2022

Abstract: Gender classification using human face data becomes a trending topic for researchers in the field of image processing and computer vision. The human face is biometric information that can be used to differentiate gender using a computer-aided system. Previous research utilised a local feature algorithm for extracting features on the face. However, the processing speed for one image was more than 2 seconds, making it unsuitable for real-time processing using video data. Processing video data requires a fast feature extraction algorithm because video data collects sequential images (frames). Moreover, the gender classification system's success is also measured by its accuracy, consequently it is necessary to choose the correct classification method to divide the two classes of men and women. In this research, we propose the FaceNet algorithm for feature extraction and explore several supervised machine learning methods (KNN, SVM, and Decision tree) appropriate for gender classification on video data. This study used 23,000 training data on each gender. From the experiment, combination of the FaceNet algorithm and KNN method resulted in the best accuracy of 95.75% with a processing speed of 0.059 seconds on each frame.

Keywords: Gender Classification; Real-time Processing; FaceNet Algorithm; Supervised Machine Learning

1. INTRODUCTION

Biometric is a technique that studies physical or human behaviour, often used as an input for pattern recognition [1]. Biometric is biological data that humans have since birth. People can use biological data to provide information about an individual's identity based on physical characteristics that distinguish individuals from one another [2]. The human face is an example of biometric data that is unique to each person [3]. The human face contains much information, including face shape, skin colour, eye shape, nose shape, mouth shape, and several additional attributes such as beard, moustache, hair, and eyebrows. This biometric information can be processed to obtain more information about a person.

Biometric data processing of face images can be used to determine a person's gender. In general, women's faces are more symmetrical than men's faces. The female face shape is also rounded and more petite. The male face's prominent characteristics are a wider mouth, a long upper lip, a larger nose, and a more pronounced lower forehead [4]. The human eye can recognise and distinguish these characteristics, but they cannot keep on it for specific needs. Therefore, we need the help of a computerised system to observe gender recognition with a particular purpose. Some of the uses of gender classification applications [5] include security system monitoring [6][7], marketing strategies in shopping centre [8][9], surveillance for advertising targets [10], and human-computer interaction applications [11][12]. The fields of science that process biometric data of face [13][14][15] are usually image processing and computer vision.

The main stage of computer vision that affects processing time and accuracy in gender classification is feature extraction and classification techniques. One of the feature extraction algorithms is using the local feature of the SURF algorithm [16]. This local feature algorithm has a disadvantage in applications that run in real-time processing like video data. The processing time for gender classification takes more than 2 seconds. These results make the system unable to run in real-time on the video data. This study proposes a feature extraction method for video data processing in the system of gender classification.

E-mail: faisal@ittelkom-pwt.ac.id, apri@ittelkom-pwt.ac.id*

One of the feature extraction methods is using transfer learning techniques, namely using a pre-trained model [17]. The pre-trained architectural FaceNet model identifies one's facial identity for employee attendance [18] and class attendance [19] which results in accuracy above 95%. The FaceNet can be used for real-time video data processing in face recognition cases [20]. The database used for real-time processing is 500 images with a resolution of 160 x 160. The testing result that uses a 640 x 480 resolution camera can run 19 frames per second (fps), which shows that the FaceNet algorithm is capable of running in real-time data. Our proposed research will use a database of 23,000 to test the accuracy and speed of real-time processing. Then, the stage after feature extraction is a classification based on gender features. The gender classification is male and female, so there are only two classes. Machine learning techniques known for their class use a supervised learning method [21][22]. The use of the FaceNet pre-trained model and several supervised learning classification methods was applied in this study. We study the KNN, Decision Tree, and SVM supervised learning methods to find the optimal accuracy and processing time for classifying human gender. Our hope from this research is that researchers and application developers can use appropriate processing video data methods for gender detection.

2. MATERIALS AND METHODS

The gender classification system starts by inputting a face dataset used as facial training data based on gender. This face dataset has been labelled according to human gender, namely male and female. Next, the face input is done in the pre-processing stage, namely converting the image to RGB colour and resizing the face image size. The pre-processing data is stored in an array, which feature extraction will be done using the FaceNet pre-trained model algorithm. The result of feature extraction is used for model building with some supervised learning methods. In this study, the KNN, SVM, and decision tree methods will be studied for optimal accuracy and speed. The gender classification system's architecture is visualized in Fig 1.

In the video data testing stage, the input of video data is extracted into video frames. Face detection is performed in the video frame as part of the gender classification process using facial biometric data. The detected face is pre-processed by resizing the face image size. The face image data are stored in an array for feature extraction using the FaceNet pre-trained model algorithm. The result of feature extraction is predicted using a gender model. The prediction results are in the form of the male or female gender. In the feature matching (prediction) stage, the processing time for one face data will be recorded.

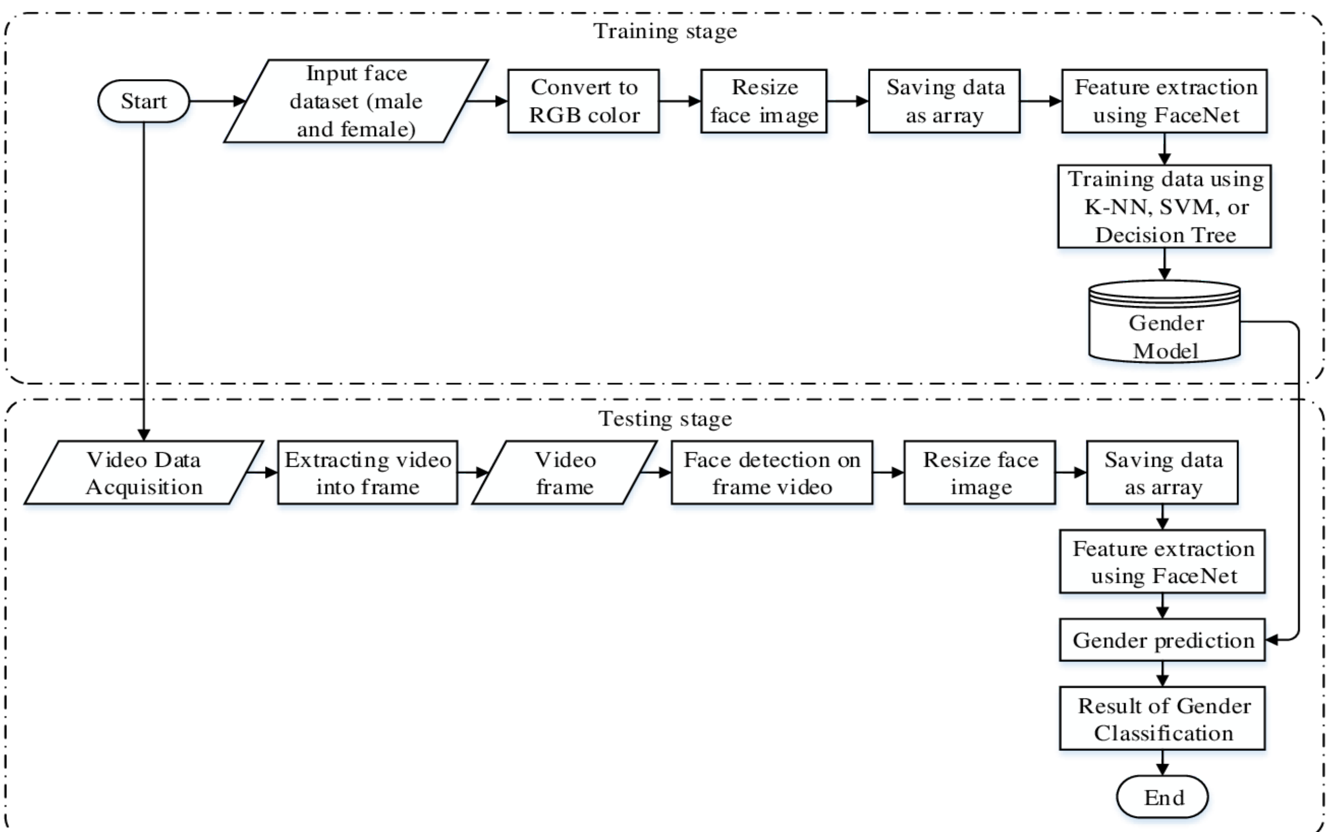


Figure 1. The architecture of gender classification system

A. Gender Classification Dataset

This paper uses a face dataset from the gender classification dataset that was cropped and saved to the male and female directory [23]. The dataset mostly takes from the IMDB dataset. In this study, each gender (male or female) used 23,000 face data for the training process. Then, using 5,500 different data for the validation. The face image data that used have been cropped for focusing on the face. Fig. 2 shows an example of a gender classification dataset.

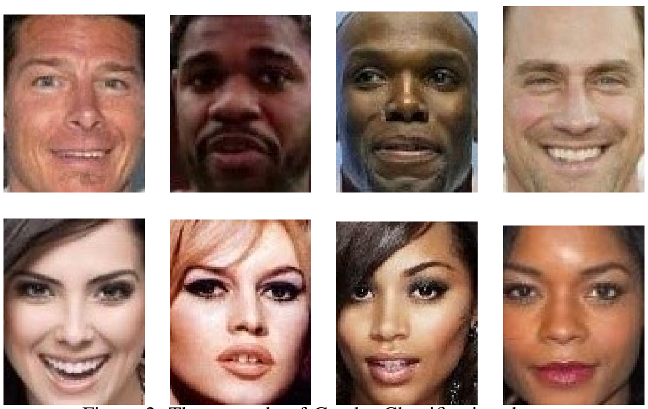


Figure 2. The example of Gender Classification dataset

B. Pre-Processing

The pre-processing stage is done on the face image before it is processed to the feature extraction stage. In this study, the gender classification dataset was converted to RGB colour and resized to a 160x160 resolution. This resizing adapts with a pre-trained FaceNet model that has been trained with 160x160 of face image resolution. In the video data testing, face images are also resized to a resolution of 160x160 before being processed to the feature extraction stage.

C. Feature Extraction

Feature extraction in this study uses the FaceNet pre-trained model algorithm, a model that has been trained by Google team researchers using deep learning Convolutional Neural Network (CNN). The FaceNet is typically used for face recognition, verification, and classification for certain purpose. It maps each face image into a euclidean space called embedding so that the space distances can be calculated the similarity of the face. Embedding is a process to obtain a level of similarity and difference of face images. If the face image has a similarity, it will get closer, and vice versa [24]. Fig. 3 illustrates the FaceNet model's structure.

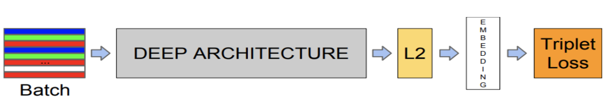


Figure 3. The structure of the FaceNet model [24]

The face image will be inputted into a deep learning architecture. Then, it will be normalised to L2. The result of this process is a feature of the face called embedding, which is followed by training using Triplet Loss [24]. Using the Triplet Loss function, Fig. 4 illustrates the feature extraction process for the FaceNet pre-trained model.

The Triplet Loss will minimise the distance between an anchor and a positive that makes a similar face image closer and maximises the distance between an anchor and a negative that makes different face images farther away [24]. This research uses the FaceNet that uses RGB images of size 160x160 with three channels of colour, and produces a face embedding vector with 128 dimensions.

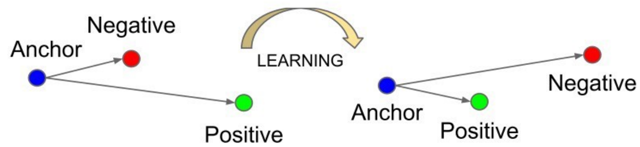


Figure 4. The Triplet Loss [24]

D. Classificaion

The classification stage is also essential in resulting in optimal accuracy and speed of time processing. In this study, three supervised learning methods will be studied: K-Nearest Neighbour (KNN), Support Vector Machine (SVM), and Decision Tree.

1) K-Nearest Neighbour (KNN)

The KNN method is an efficient lazy learning algorithm that requires labelled training data [25]. KNN compares the training data with the test data, where the training data are described in several attributes that are totalling n [26]. Each training data represents a point in n-dimensional space (vector of size n). The classification of new data is done by calculating the new data's similarity or closeness to all training data. Fig. 5 shows an illustration of the K-Nearest Neighbour method.

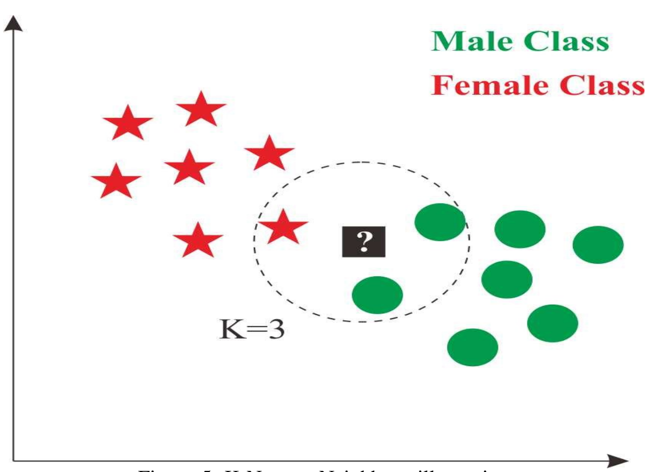


Figure 5. K-Nearest Neighbour illustration

E-mail: faisal@ittelkom-pwt.ac.id, apri@ittelkom-pwt.ac.id*

The level of similarity can be calculated using several methods, Euclidean Distance or Manhattan Distance. The Euclidean Distance is the distance between two locations on a straight line using the Pythagorean theorem [27]. Equation (1) shows the formula for Euclidean Distance. A distance of two vectors of size n, for example, $X = (X_1, X_2, \ldots, X_n)$ and $Y = (Y_1, Y_2, \ldots, Y_n)$.

$$dist(X, Y) = \sqrt{\sum_{i=1}^{n} (X_i - Y_i)^2} \tag{1}$$

The Manhattan distance between two points, X and Y, in d-space dimensions is defined in equation (2).

$$dist(X, Y) = \sum_{k=1}^{d} |X_j - Y_j| \tag{2}$$

This study will examine several K values to calculate the nearest neighbour, namely 1,3,5,7, and 9. This research will test each K value on the distance algorithm, namely Euclidean Distance and Manhattan Distance.

2) Support Vector Machine (SVM)

SVM is a supervised learning technique that is frequently used for classification. SVM creates a hyperplane in high dimensional space [28]. SVM will find good separation achieved by the most considerable distance of hyperplane to the nearest training data points of two gender classes (male and female). Margin is the distance between the hyperplane and the closest point of the male and female class called Support Vector Machine [29]. Fig. 6 illustrates the SVM technique.

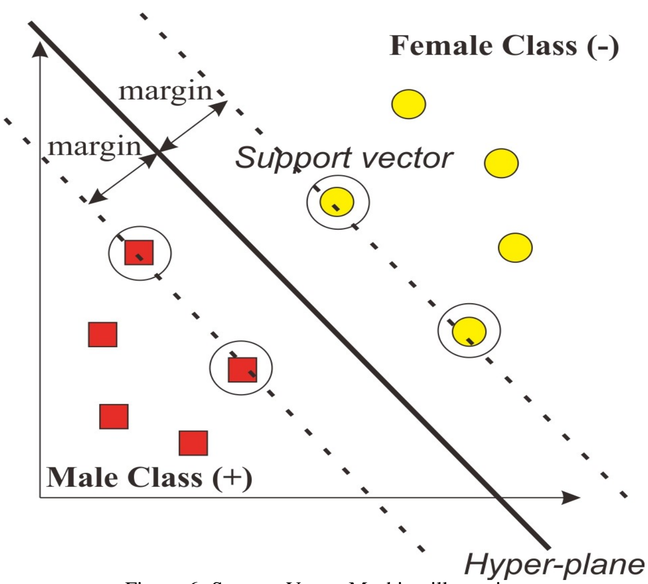


Figure 6. Support Vector Machine illustration

Data are represented as $\vec{x}_i \in \Re^d$ while the respective labels are represented as $y_i \in \{-1, +1\}$ for $i = 1, 2, \ldots, l$, where l is the numbers of data. It is assumed that both classes −1 (female class) and +1 (male sample) can be entirely separated by the hyperplane of dimension d, which is defined by equation (3).

$$\vec{w} \cdot \vec{x} + b = 0 \tag{3}$$

Pattern $\vec{x}_i$, which is included in female class can be defined as a pattern that fulfils the inequality (4).

$$\vec{w} \cdot \vec{x}_i + b \leq -1 \tag{4}$$

Meanwhile, the pattern x_i which belongs to male class fulfils the inequality (5).

$$\vec{w} \cdot \vec{x}_i + b \geq -1 \tag{5}$$

The largest margin is obtained by maximising the distance between the hyperplane and its nearest point, which is $1/\|\vec{w}\|$. It can be defined as a Quadratic Programming (QP) problem to determine the formula's minimal point (6) and taking into account the constraints of formula (7).

$$\min_{\vec{w}} \tau(w) = \frac{1}{2}\|\vec{w}\|^2 \tag{6}$$

$$y_i(\vec{x}_i \cdot \vec{w} + b) - 1 \geq 0, \quad \forall i \tag{7}$$

The SVM method has several approaches that are commonly called the kernel. In this study, 3 SVM kernels will be studied, namely the Radial Basis Function (RBF), linear, and polynomial.

3) Decision Tree

The decision tree is a prediction method for classifying data that resembled tree structure or hierarchical structure [30]. Converting data into a decision tree and decision rules is the main principle of a decision tree. The most significant advantage of using a decision tree is that it simplifies difficult decisions. This process makes decision-makers can interpret better solutions to problems.

A decision tree generates a tree that contains both decision and leaf nodes. The decision node of the tree has two or more branches. The leaf node of the tree represents the classification or assessment of a problem. The tree's root node corresponds to the best prediction [31]. Our research uses the Classification and Regression Trees (CART) algorithm. CART builds binary trees based on the function and threshold that give each node the most information gain [32]. In this research, the criterion of the decision tree will be studied, namely Gini or Entropy. The Gini impurity formula is shown in equation (8) [33].

$$GiniIndex = 1 - \sum_j p_j^2 \tag{8}$$

Where, p_j is the probability of class j. When a dataset is randomly labelled, the Gini impurity calculates the probability that some variable will be mislabeled. The Gini Index has a minimum value of 0. When the node is pure, it implies that all node's contained elements belong to the same class. Then, the Entropy criterion is calculated with equation (9) [33].

$$Entropy = -\sum_j p_j \cdot \log_2 \cdot p_j \tag{9}$$

Where, p_j is the probability of class j. Entropy is a measure of knowledge that shows how disordered the features are concerning the target. The optimum split is determined by the function with the least entropy, similar to the Gini Index.

E. Data Analysis

In this research, we discuss the accuracy of the training and testing system using a Python programming language library for predictive modelling, namely scikit-learn. One of the measurement matrices is using $accuracy_score$. The formula for calculating the mathematical accuracy is shown in equation (10). The accuracy score function calculates the accuracy of correct predictions as a fraction.

$$Accuracy(y, y') = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} 1(y'_i = y_i) \tag{10}$$

Where, y'_i is the predicted value of the i-th sample and y_i is the actual value, then the fraction of true predictions over $n_{samples}$.

After obtaining the model, we tested the face through real-time video to get data on the accuracy and time required to process one video frame. The formula for calculating accuracy is shown in equation (11) [34].

$$Accuracy(y, y') = \frac{TP + TN}{TP + TN + FP + FN} \tag{11}$$

True Positive (TP) refers to the amount of positive data that the system accepts as positive, True Negative (TN) refers to the amount of negative data that the system accepts as negative, False Positive (FP) refers to the amount of negative data that the system accepts as positive, and False negative (FN) refers to the amount of positive data that the system accepts as negative.

3. EXPERIMENTAL RESULT AND ANALYSIS

In this research, some experiments were done on the gender classification method. The classification method to be tested uses the K-Nearest Neighbour, Support Vector Machine, and Decision Tree methods. Then, testing is also done on the speed of video data processing. The computer specification and operating system are shown in Table 1.

TABLE I. COMPUTER CONFIGURATION

Criteria	Configuration
Processor	Intel Core i3-9100F CPU @ 3.60 GHz
Memory	8192 MB
Display	Radeon RX550
Hard disk	500 GB
CCTV Camera	2.0 MP with ten fps
Operating System	Windows 10 Pro 64-bit

A. Training Results

1) Learning using K-Nearest Neighbour

In the gender classification experiment using the KNN method, the closest distance calculation uses the Manhattan distance and the Euclidean distance. The closest distance computation is performed to determine the number of similarities based on the facial image features. Then, the characteristics or features of the tested data are compared with each original data. Fig. 7 shows the variation in the K value and the distance of the KNN algorithm.

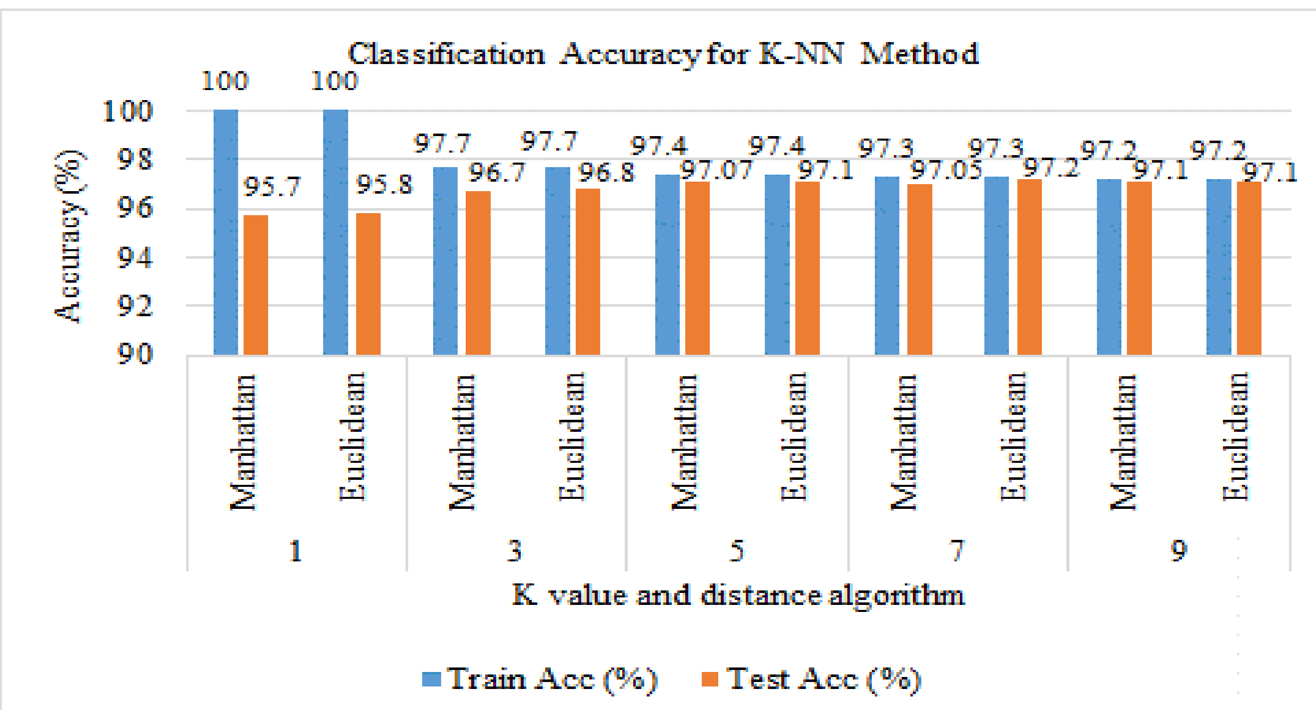


Figure 7. The results of variations in K values and the distance algorithm on the KNN method

E-mail: faisal@ittelkom-pwt.ac.id*, apri@ittelkom-pwt.ac.id

The value of K is the number of reference points to be compared with the test point. At $K = 1$, the closest distance to the test point value will be sought. For K values that are more than 1, voting is done based on the majority of the class. Therefore, the value of K is an odd number, so that voting can be done based on the majority of its existence in a class. In the experimental results in Fig. 7, the training accuracy at $K = 1$ is 100% because the feature image data are very close to each class so that the greater of the K value causes a decrease in the system's ability to determine male or female classes. However, the best test accuracy value in this experiment is at $K = 7$ 0f 97.2%, and the train accuracy value is 97.3% It is better than $K = 1$, which is slightly overfitting. Therefore, in the experiment, $K = 7$ with Euclidean distance will be used to test the video data.

2) Learning Using Support Vector Machine

In the learning process using SVM, the input space's data are transformed into a feature space using a kernel trick. The three kernels hold up an important role in gender classification. Fig. 8 shows the comparison results of kernel use in SVM.

Based on Fig. 8, the test accuracy's best result uses the polynomial kernel, which is 97.2%. This polynomial kernel classification will be used for the testing of the video data. The linear kernel accuracy result is below 85% because the linear kernel draws a straight line to separate the two classes (male and female). It contrasts with the RBF kernel and polynomial, which is suitable for image data where the image features are scattered.

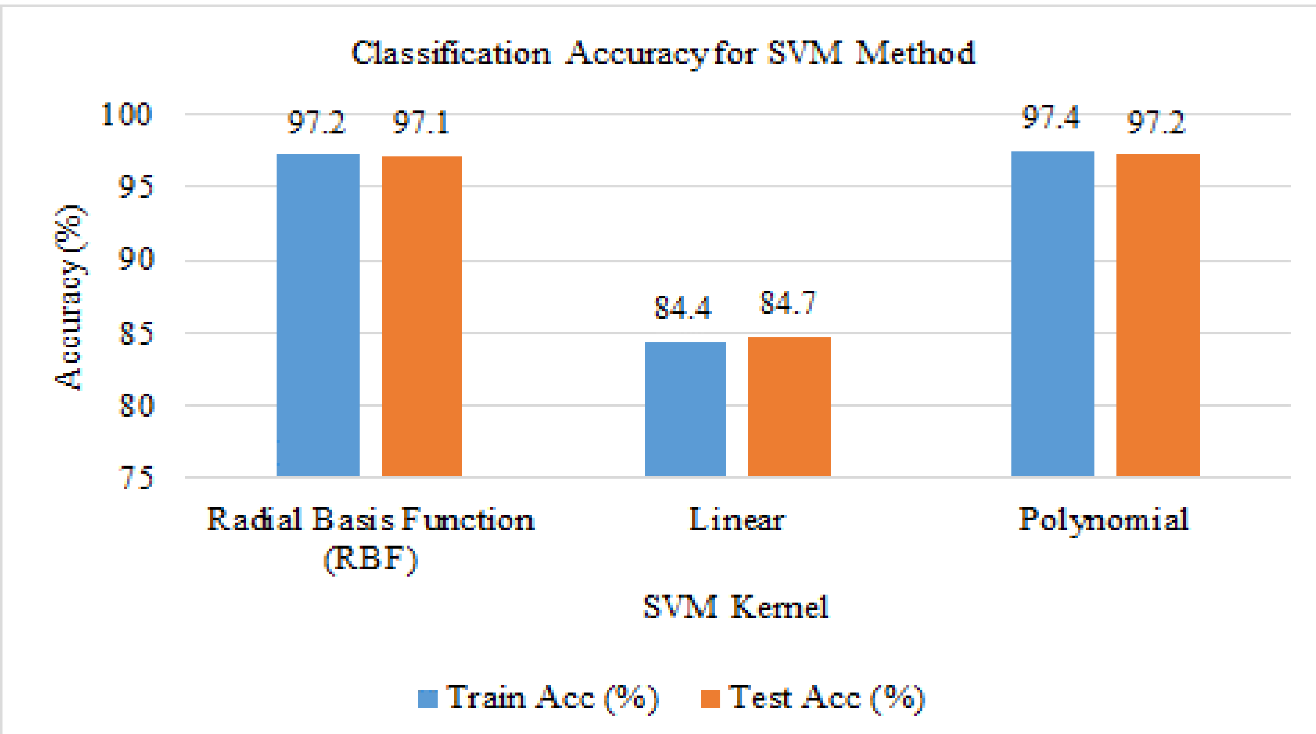


Figure 8. The results of variations in kernels of SVM method

3) Learning using Decision Tree

The decision tree method concept helps select suitable features for splitting the tree into subparts. The depth of the tree will be less or more dependent on the decision tree building's fitness. Entropy criterion helps build a suitable decision tree for selecting the best splitter. Then, the criterion of gini impurity is also similar to the entropy criterion in the decision tree. Fig. 9 shows the experimental results of the depth and the criteria of the decision tree.

In general, the classification results using the decision tree shown in Fig. 9, the decision tree criteria do not significantly affect the accuracy results. The tree depth is the most influence on train accuracy results. However, the higher value of maximum depth causes overfitting, namely train accuracy far higher than the test accuracy. For example, at a depth of 15, the training accuracy value is above 95% but the testing accuracy value is around 80%. The greater the depth value creates a larger gap in the accuracy of training and testing. Therefore, using the decision tree method for processing features on the image is not quite right. Based on the results shown in Fig. 9, the decision tree classification method with a depth of 5 and gini criterion will be used for testing on video data because the result is not overfitting and its accuracy is higher than entropy criterion.

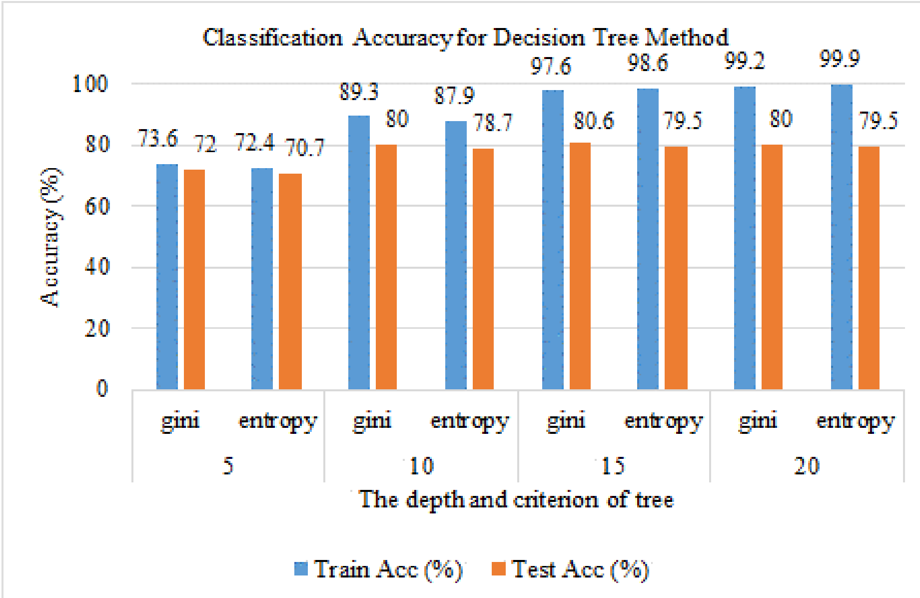
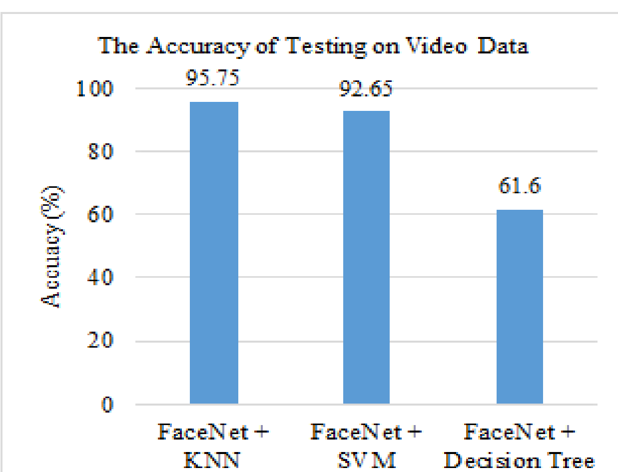


Figure 9. The results of variations in depth and criterion on the Decision tree method

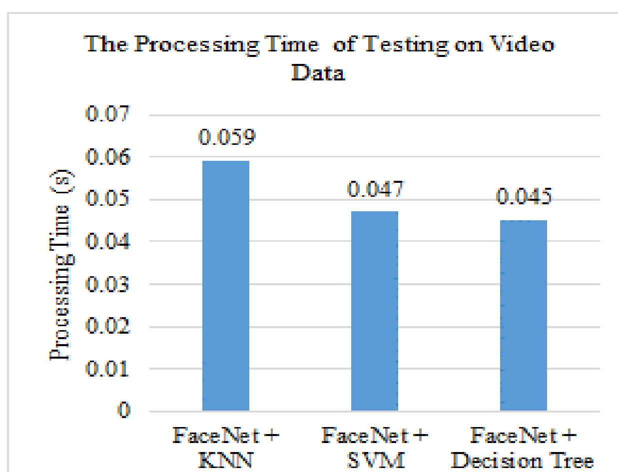
B. Testing on Video Data

Experiments on the video data use ten fps video data. This dataset consists of data on men and women that have not been cropped in their faces. The video resolution is Full HD 1920 x 1080. Experiments are carried out on each of the supervised learning methods (FaceNet + KNN, FaceNet + SVM, and FaceNet + Decision Tree) that produced the best accuracy value. Fig. 10 shows the experimental results of the combination of feature extraction algorithm and classification methods.

Based on Fig. 10, the best classification method for gender classification uses the KNN method. Video data processing requires fast feature extraction in each video frame. This research uses ten fps video data means every 1 second, there are ten frames. The FaceNet algorithm for processing one video frame only takes 0.059 seconds for classifying the data using the KNN method. This result is a little bit longer than the SVM and decision tree algorithms because the KNN compares the nearest neighbour's testing data in the 7 sample data. However, the experimental accuracy of male and female video data gives the best result of 95.75%. In this study's classification methods, the combination methods are faster than the previous study [16], requiring more than 2 seconds to process a single image. It proves that the FaceNet algorithm can be used as a feature extraction for video data.



(a)



(b)

Figure 10. The result of testing on video data, a) Accuracy b) Processing Time

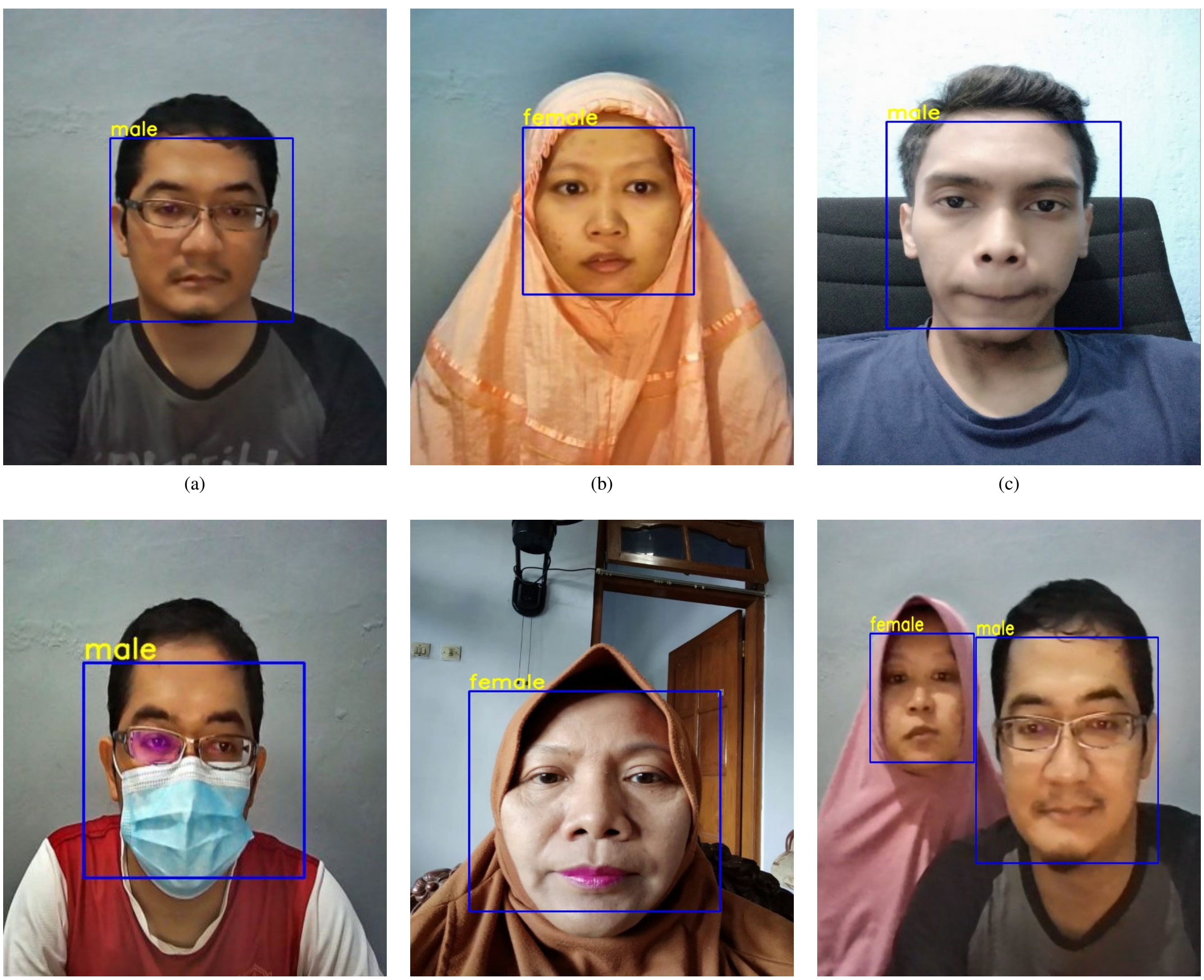
E-mail: faisal@ittelkom-pwt.ac.id*, apri@ittelkom-pwt.ac.id

C. Discussion

The results of the gender classification using the KNN method classifier, as shown in Fig. 7, found the use of variations in the distance method (Euclidean and Manhattan) has no significant effect on accuracy results. Likewise, the use of criteria in the decision tree method classifier, as shown in Fig. 9, between gini and entropy also has no significant effect on the accuracy results. It is different when using the SVM classifier, as shown in Fig. 8, the use of kernel variations dramatically affects the accuracy results. The optimal accuracy results use the KNN and SVM classifiers. However, optimal results are obtained using the KNN classifier because the classification process is looking for the nearest neighbour. Unlike the SVM case, which is separated using the kernel in each of its class regions, which usually has an incorrect image feature in its class division.

The use of the FaceNet algorithm and the KNN classification method provides optimal results for gender classification. The feature extraction using the FaceNet algorithm is 35 times faster than feature extraction in the previous research [16]. Fig. 11 shows an example of the results of gender classification on real-time video data. In Figure 11d, the face covered by the mask can still be detected correctly.

In this research, there is still a weakness in face detection. The detected faces still have parts that are not face as shown in Fig. 11. There is a background area that is captured along with the face. Future research can be modified on the face detection algorithms, so that face detection only focuses on the face. If the focus is only on the face, the accuracy of gender classification will likely increase.



(a) (b) (c)

(d) (e) (f)

Figure 11. Example of gender classification results: a, c, d) male person, b, e) female person f) male and female person

4. CONCLUSIONS AND FUTURE WORK

We introduce a gender classification system based on human biometric data, that is faces. Speed is the main focus of video data processing. The feature extraction algorithm determines the speed of the system. Then the classification method plays a role in achieving the best accuracy. In this study, the combination of the FaceNet on feature extraction algorithm and the KNN classification method results in an accuracy of 95.75% with an average speed of 0.059 seconds for processing on each frame.

This research provides an overview of methods for processing real-time video data in cases of human gender classification. For example, entrepreneurs who want to know their customer segments can use the FaceNet method with the KNN for gender classification. The researchers can find out the development of gender classification using several supervised machine learning methods. However, this research still needs better future research. Our approach has a weakness in face detection. The following study can be modified on the face detection algorithm to only focus on the face.

References

[1] A. Shama, P. Chaturvedi, and S. Arya, "Human Recognition Methods based on Biometric Technologies," *Int. J. Comput. App*, vol. 120, no. 17, pp. 1–7, 2015. doi: 10.5120/21316-4312.

[2] Y. Lin and H. Xie, "Face Gender Recognition based on Face Recognition Feature Vectors," in *Proc. 2020 IEEE 3rd Int. Conf. Inf. Syst. Comput. Aided Educ. ICISCAE 2020*, pp. 162–166, 2020. doi: 10.1109/ICISCAE51034.2020.9236905.

[3] I. Candradewi, B. N. Prastowo, and D. Latihef, "Gender Classification from Facial Images Using Support Vector Machine," *J. Theor. Appl. Inf. Technol*, vol. 97, pp. 2684–2692, 2019.

[4] Z. Skomina, M. Verdenik, and N. I. Hren, "Effect of aging and body characteristics on facial sexual dimorphism in the Caucasian Population," *PLoS One*, vol. 15, no. 5, pp. 1–15, 2020. doi: 10.1371/journal.pone.0231983.

[5] C. Shan, "Learning local binary patterns for gender classification on real-world face images," *Pattern Recognit. Lett*, vol. 33, no. 4, p. 431–437, 2012. doi: 10.1016/j.patrec.2011.05.016.

[6] S. A. Khan, M. Nazir, S. Akram, and N. Riaz, "Gender classification using image processing techniques: A survey," in *Proc. 14th IEEE Int. Multitopic Conf. 2011, INMIC 2011*, pp. 25–30, 2011. doi: 10.1109/INMIC.2011.6151483.

[7] M. Raza, M. Sharif, M. Yasmin, M. A. Khan, T. Saba, and S. L. Fernandes, "Appearance based pedestrians' gender recognition by employing stacked auto encoders in deep learning," *Futur. Gener. Comput. Syst*, vol. 88, no. 4, pp. 28–39, 2018. doi: 10.1016/j.future.2018.05.002.

[8] Ö. Özbudak, M. Kirci, Y. Çakir, and O. Güneş, "Effects of the facial and racial features on gender classification," in *Proc. Mediterr. Electrotech. Conf. - MELECON*, no. December 2014, pp. 26–29, 2010. doi: 10.1109/MELCON.2010.5476346.

[9] J. Dong, Y. Du, and Z. Cai, "Gender recognition using motion data from multiple smart devices," *Expert Syst. Appl*, vol. 147, 2020. doi: 10.1016/j.eswa.2020.113195.

[10] I. K. Timotius and I. Setyawan, "Using edge orientation histograms in face-based gender classification," in *2014 Int. Conf. Inf. Technol. Syst. Innov. ICITSI 2014 - Proc.*, no. November, pp. 93–98, 2014. doi: 10.1109/ICITSI.2014.7048244.

[11] R. Sarkar, S. Bakshi, and P. K. Sa, "A Real-time Model for Multiple Human Face Tracking from Low-resolution Surveillance Videos," *Procedia Technol*, vol. 6, pp. 1004–1010, 2012. doi: 10.1016/j.protcy.2012.10.122.

[12] A. Kaur and B. V. Kranthi, "Comparison between YCbCr Color Space and CIELab Color Space for Skin Color Segmentation," *Int. J. Appl. Inf. Syst*, vol. 3, no. 4, pp. 30–33, 2012. doi: 10.1016/j.protcy.2012.10.122.

[13] G. D. K. Kishore and B. Mukamlla, "Detecting human and classification of gender using facial images msift features based gsvm," *Int. J. Recent Technol. Eng*, vol. 8, no. 3, pp. 1466–1471, 2019. doi: 10.35940/ijrte.B3782.098319.

[14] D. K. K. Galla and B. Mukamalla, *Real Time Gender Classification Based on Facial Features Using EBGM*. Springer International Publishing, 2020.

[15] M. Afifi and A. Abdelhamed, "Afif4: Deep gender classification based on AdaBoost-based fusion of isolated facial features and foggy faces," *J. Vis. Commun. Image Represent.*, vol. 62, no. 3, pp. 77–86, 2019. doi: 10.1016/j.jvcir.2019.05.001.

[16] B. Hapitoglu and C. Kose, "A gender recognition system from facial images using surf based bow method," *2nd Int. Conf. Comput. Sci. Eng. UBMK 2017*, vol. 5, no. 5, pp. 988–993, 2017. doi: 10.1109/UBMK.2017.8093405.

[17] M. M. Khayyat and L. A. Elrefaei, "Towards Author Recognition of Ancient Arabic Manuscripts Using Deep Learning: A Transfer Learning Approach," *Int. J. Comput. Digit. Syst*, vol. 90, no. 5, pp. 783–799, 2020. doi: 10.12785/ijcds/090502.

[18] F. Cahyono, W. Wirawaran, and R. F. Rachmadi, "Face Recognition System using Facenet Algorithm for Employee Presence," vol. 90, no. 5, pp. 57–62, 2020. doi: 10.1109/icovet50258.2020.9229888.

[19] T. Nyein and A. N. Oo, "University Classroom Attendance System Using FaceNet and Support Vector Machine," in *2019 International Conference on Advanced Information Technologies (ICAIT)*, pp. 171–176, Nov. 2019. doi: 10.1109/AITC.2019.8921316.

[20] S. Huang and H. Luo, "Attendance System Based on Dynamic Face Recognition," in *Proc. - 2020 Int. Conf. Commun. Inf. Syst. Comput. Eng. CISCE 2020*, pp. 368–371, 2020. doi: 10.1109/CISCE50729.2020.00081.

[21] I. Muhammad and Z. Yan, "Supervised Machine Learning Approaches: a survey," *ICTACT J. Soft Comput*, vol. 05, no. 03, pp. 946–952, 2015. doi: 10.21917/ijsc.2015.0133.

[22] I. Waspada, A. Wibowo, and N. S. Meraz, "Supervised Machine Learning Model for Microrna Expression Data in Cancer," *IJ. Ilmu Komput. dan Inf.*, vol. 10, no. 2, p. 108, 2017. doi: 10.21609/jiki.v10i2.481.

[23] A. Chauhan, "Gender Classification Dataset — Kaggle," 2019. https://www.kaggle.com/cashutosh/gender-classification-dataset (accessed Dec. 08, 2020).

[24] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 815–823, June 2015. doi: 10.1109/CVPR.2015.7298682.

[25] Z. Deng, X. Zhu, D. Cheng, M. Zong, and S. Zhang, "Efficient knn classification algorithm for big data," *Neurocomputing*, vol. 195, no. 2, pp. 143–148, 2016. doi: 10.1016/j.neucom.2015.08.112.

[26] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques*. United States of America: Elsevier, 3rd ed., 2012.

[27] R. Suwanda, S. Z, and E. M. Zamzami, "Analysis of Euclidean Distance and Manhattan Distance in the K-Means Algorithm for Variations Number of Centroid K," *J. Phys. Conf. Ser.*, vol. 1566, no. 1, pp. 143–148, 2020. doi: 10.1088/1742-6596/1566/1/012058.

[28] B. Ghaddar and J. Naoum-Sawaya, "High dimensional data classification and feature selection using support vector machines," *J. Oper. Res.*, vol. 265, no. 3, pp. 993–1004, 2018. doi: https://doi.org/10.1016/j.ejor.2017.08.040.

[29] A. S. Nugroho, A. B. Witarto, and D. Handoko, "Support Vector Machine-Teori dan Aplikasinya dalam Bioinformatika," 2003.

E-mail: faisal@ittelkom-pwt.ac.id, apri@ittelkom-pwt.ac.id*

[30] S. B. Kotsiantis, "Decision trees: A recent overview," *Artif. Intell. Rev.*, vol. 39, no. 4, p. 261–283, 2013. doi: 10.1007/s10462-011-9272-4.

[31] N. B. Noor, M. S. Anwar, and M. Dey, "Comparative Study between Decision Tree, svm and knn to Predict Anaemic Condition," in *BECITHCON 2019 - 2019 IEEE Int. Conf. Biomed. Eng. Comput. Inf. Technol. Heal.*, no. November, pp. 24–28, 2019. doi: 10.1109/BECITHCON48839.2019.9063188.

[32] P. Thariga, I. S. Sitanggang, and L. Syaufina, "Comparative Analysis of Spatial Decision Tree Algorithms for Burned Area of Peatland in Rokan Hilir Riau," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 14, no. 2, pp. 684–691, 2016. doi: 10.12928/TELKOMNIKA.v14i1.3540.

[33] P. Gulati, A. Sharma, and M. Gupta, "Theoretical Study of Decision Tree Algorithms to Identify Pivotal Factors for Performance Improvement: A Review," *Int. J. Comput. Appl.*, vol. 141, no. 14, pp. 19–25, 2016. doi: 10.5120/ijca2016909926.

[34] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006. doi: 10.1016/j.patrec.2005.10.010.

Faisal Dharma Adhinata obtained a Master of Computer Science (M.Cs) degree in Computer Science from Universitas Gadjah Mada, Indonesia, in 2020. Currently, he is a lecturer at the Department of Software Engineering, Institut Teknologi Telkom Purwokerto. His research interests are Artificial Intelligence, machine learning techniques, image processing, and computer vision.

Apri Junaidi obtained a Master of Computer Science (M.C.S) degree from Universiti Teknikal Malaysia Melaka, Malaysia, in 2012. He joined Institut Teknologi Telkom Purwokerto as head of expertise group "Rekayasa Data". Currently, he is a lecturer at the Department of Data Science. His research interests include machine learning techniques, data analytic, and voice recognition.