1st Proceeding of International Conference on Science and Technology UISU (ICST)
"Contribution of Engineering Disciplines to Achieving Sustainable Development Goals Research Realizing the Vision of Golden Indonesia 2045"
ISSN: 3063-7929 (Online)
DOI: https://doi.org/10.30743/icst

Comparison of the Use of CNN Methods with SVM for Eye Detection in Car Drivers

M. Khairul Ikhsan[1], Roslina[2], Rika Rosnelly[3]
[1]Master of Computer Science, Faculty of Engineering and Computer Science, Universitas Potensi Utama, Medan. Indonesia
[2]Department of Computer and Information Systems, Politeknik Negeri Medan. Indonesia
[3]Department of Computer Science, Faculty of Engineering and Computer Science, Universitas Potensi Utama, Medan. Indonesia
[*]Corresponding author: ikhsantkj0406@gmail.com

Abstract -This study aims to compare the two Convolutional Neural Network (CNN) methods and the Support Vector Machine (SVM) method using Haar Casecade to detect eye patterns, CNN and SVM function to detect eyes in open or close conditions. SVM in the study uses HOG to process images from Haar Casecade so that they can be read by SVM and uses kernels, namely RBF, poly and sigmoid. The case study taken is the detection of sleepy eyes in car drivers using 250 open eye images and 250 closed eye images. The experimental results show that CNN is superior by showing an accuracy value of 98%, precision closed 96% precision open 100%, recall closed 100% and recall open 96%, for F1-score closed and open 98%, while SVM in this study shows an accuracy result of 95% precision close 98% and open 92%, recall close 93% and open 98%, F1-score 95%. This study confirms that the selection of the CNN method is the most appropriate to be used in combination with Haar Casecade because it gets higher accuracy results. These findings can be used as a basis for further development in the drowsiness detection system for car drivers.

Keywords: ***Comparison; Haar Casecade; CNN; SVM; HOG; Kernel; Sleepiness.***

1. INTRODUCTION

Rapid developments in the field of Deep learning have brought new innovations in object detection. Convolutional Neural Networks (CNN) is one type of Deep learning that is very effective in processing image data. CNN is able to automatically extract features from data without the need for a complex manual feature extraction process[1]. This advantage makes CNN very effective in tasks such as object detection and pattern recognition, including eye detection in car drivers. Before discussing further the comparison between traditional methods such as Haar Cascade and modern methods such as CNN, it is important to understand the background and evolution of these two approaches.

Advances in machine learning have also resulted in systems for analyzing data and acquiring knowledge without explicit programming. This technology has seen a huge surge in its application to solve various problems and is capable of automation in various fields, mainly due to the explosion of data, improvements in ML techniques, and advances in computing capabilities [2].Machine learning techniques such as SVM, logistic regression, and decision trees are relevant and effective in various image classification applications. These techniques offer advantages in processing speed, interpretability, and lower cost, while still providing high accuracy. The application of data augmentation and preprocessing techniques also helps improve the performance of these models in image classification tasks[3]

Driver's eye detection is a technology used to monitor and analyze the driver's eye condition to determine the level of alertness and detect signs of driver fatigue [4]. This technology is essential to improve road safety by providing real-time feedback and warnings to prevent accidents caused by driver fatigue[5]. In the context of driver's eye detection, accuracy and speed of detection are important factors to consider. Accurate detection can help the system to provide early warnings when the driver shows signs of fatigue or drowsiness. Speed of detection is also important because the system must be able to work in real-time to provide a fast and accurate response. The comparison between traditional Machine learning methods such as Haar Cascade and SVM with Deep learning methods such as CNN becomes relevant because each has its own advantages [6].

This study aims to evaluate the performance of the eye detection method using Haar Cascade by comparing deep learning versus Machine learning models in detecting driver eye conditions. It is expected that an optimal comparison can be found to achieve high classification accuracy. Haar Cascade is a classification method that uses Haar features, first introduced by Viola and Jones. This method is very popular for object detection because of its simple structure, high detection rate, and fast detection speed [7]. The Haar Cascade method has a long history and has been widely used in various applications. Haar Cascade uses Haar-like features calculated based on the difference in intensity between pixels in a certain area of the image. These features are then combined in a cascade classifier that is able to detect objects quickly and efficiently[8]. Haar Cascade is used to detect frontal human faces with high accuracy. This method utilizes Haar features to identify facial parts such as the eyes, nose, and mouth based on differences in brightness in those areas [7].

SVM is a Machine learning algorithm used for classification and regression. SVM works by finding the best hyperplane that separates the data into different classes [9]. SVM operates by finding the optimal hyperplane that separates the data classes by the largest margin. This approach helps in maximizing the distance between different classes, which is important for improving accuracy and reliability in image classification [10].In the context of eye detection, SVM can be used to classify areas of an image that contain eyes and those that do not [11]. SVM is very effective in handling high-dimensional data, which

1st Proceeding of International Conference on Science and Technology UISU (ICST)
"Contribution of Engineering Disciplines to Achieving Sustainable Development Goals Research Realizing the Vision of Golden Indonesia 2045"
ISSN: 3063-7929 (Online)
DOI: https://doi.org/10.30743/icst

is common in image classification where each pixel can be considered a different feature. This capability allows SVM to handle the complexity of image data well, resulting in accurate classification [12].

On the other hand, CNN is a type of artificial neural network designed to process data in the form of grids, such as images [13]. CNN consists of several convolutional layers that extract features from the input image and several pooling layers that reduce the dimensionality of the features. CNN has the ability to automatically learn features from data without the need for manual feature extraction, which makes it very effective in tasks such as object detection [14]. With this ability, CNN is able to overcome various challenges faced by traditional methods, such as variations in lighting, viewing angles, and environmental conditions[15]. Data normalization is an important step for numerical stabilization of CNN which changes the range of pixel intensity values and transforms the input image into a range of pixel values that are more familiar or normal to the senses[16].

In the context of this study, it is important to evaluate both methods based on several key criteria, including accuracy, speed, and robustness to environmental variations. Detection accuracy is a key criterion that an eye detection system must meet to ensure that it can provide the right early warning to the driver. Detection speed is also important because the system must work in real-time to provide a quick response. Robustness to environmental variations is also important to ensure that the system can perform well under various conditions, such as changes in lighting, viewing angles, and facial expressions.

This research is expected to be used as a basis for the development of a more efficient and accurate eye detection system, which can ultimately help reduce traffic accidents caused by fatigue or drowsiness. Thus, this research not only has academic value but also has significant practical impact in improving road safety.

2. METHODS

This research method adopts the Knowledge Discovery in Databases (KDD) approach, which is a systematic methodology to obtain knowledge from relevant, useful, and understandable data[17]. The KDD process involves several interconnected stages, as shown in Figure 1.

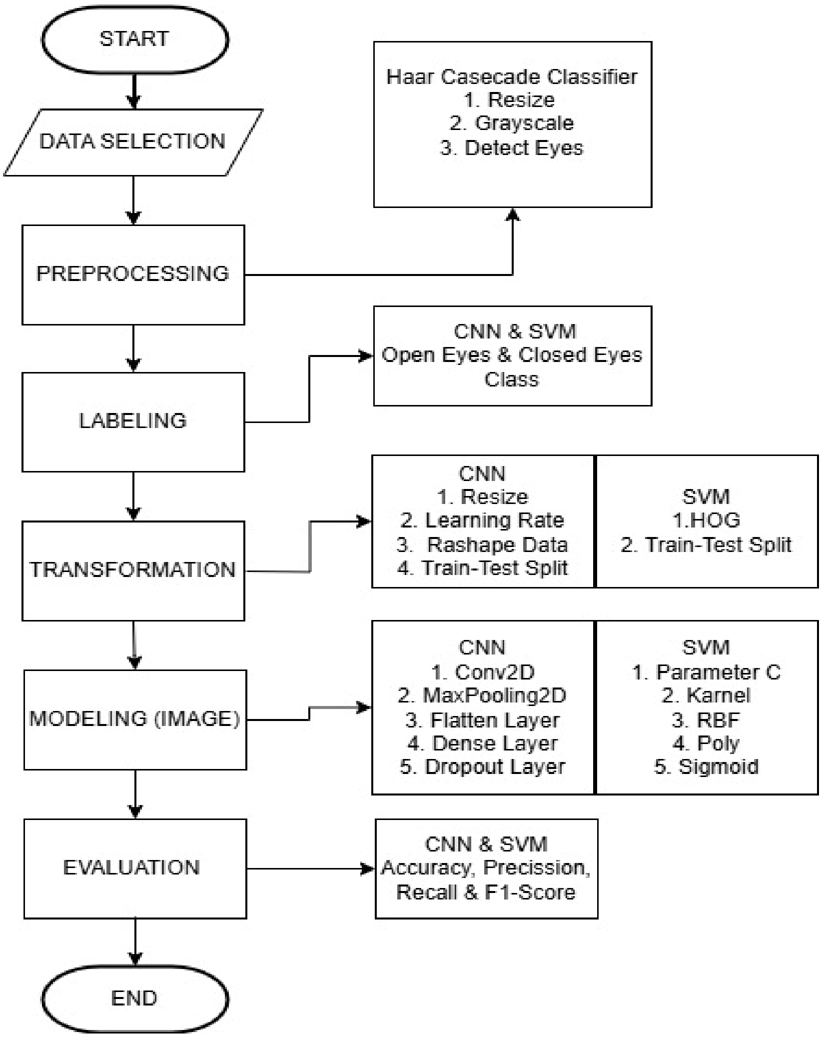

Figure 1. Research Stage

ISSN: 3063-7929 (Online)
DOI: https://doi.org/10.30743/icst

2.1. Data Selection

The dataset used in this study was obtained from the Kaggle platform, which provides various quality datasets for data analysis and machine learning purposes. This dataset contains eye images with two main categories: eyes in open condition and eyes in closed condition. This dataset is designed to support various applications, such as fatigue detection, biometrics, and pattern recognition.

2.2. Preprocessing

Preprocessing is an important step in preparing data for comparative analysis of methods. In this study, we use Haar Casecade to detect patterns in the eyes and resize and convert images in the dataset to grayscale while CNN and SVM function to detect open or closed eye image data. Here is an explanation of each step:

Haar cascade classification haar casecade functions to detect eye patterns and resize and convert images to grayscale, this stage is very important so that CNN works easier and smoother to get maximum results. The following are the results of the preprocessing of the Haar Casecade Classification.

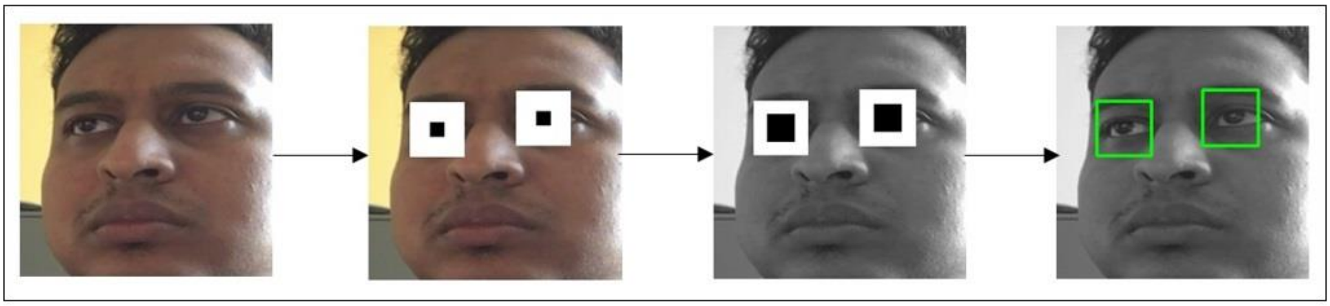

Figure 2. Haar Casecade Classification

The image above explains the stages of Classification using Haar Casede where the first image before classification and the second image after classification will display a box on the eye, indicating the progress of the eye detection process in the eye, then an image of the eye that has been grayscaled and the final stage of Haar Casede Classification displays a box on the eye transparently so that the eye can be seen in an open or closed condition.

2.3. Labeling

The dataset used for sleepiness detection consists of images that have been labeled into two main categories, namely open (eyes open) and close (eyes closed). Labeling is done by separating images into two separate folders based on the subject's eye condition. This labeling process aims to facilitate model learning in recognizing visual patterns of open and closed eyes.

2.4. Transformation

The Transformation stage is very important so that CNN and SVM can produce accurate readings. At this stage, CNN has four stages of accuracy, namely: Resize, Learning Rate, Reshape Data, Train-Test Split and for SVM it has two Transformation stages, namely: HOG and Train-Test Split.

2.5. Modelling

The transformed data is then modeled using the CNN and SVM methods. The next modeling process is preprocessing and transformation using the Python programming language. A total of 500 image data were used for this study. The data will be divided into training data and testing data, the data consists of 250 open eye data and 250 close eye data. The following are the modeling results from CNN and SVM.

a. CNN Modeling

The following shows the architecture of the Convolutional Neural Network (CNN) model designed for image classification with 2 classes, namely open and close. This model consists of several main layers, each of which has a specific function in extracting features, reducing dimensions, and processing data to produce the final output in the form of predictions. This model is designed with a combination of convolutional and dense layers, which allows for pattern recognition in images while providing flexibility for classification. However, with a large number of parameters, regularization such as dropout becomes very important to prevent overfitting. Here are the details of each layer:

1st Proceeding of International Conference on Science and Technology UISU (ICST)
"Contribution of Engineering Disciplines to Achieving Sustainable Development Goals Research Realizing the Vision of Golden Indonesia 2045"
ISSN: 3063-7929 (Online)
DOI: https://doi.org/10.30743/icst

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 222, 222, 32)	320
maxpool2d_1 (MaxPooling2D)	(None, 111, 111, 32)	0
conv2d_2 (Conv2D)	(None, 109, 109, 64)	18,496
maxpool2d_2 (MaxPooling2D)	(None, 54, 54, 64)	0
conv2d_3 (Conv2D)	(None, 52, 52, 128)	73,856
maxpool2d_3 (MaxPooling2D)	(None, 26, 26, 128)	0
flatten (Flatten)	(None, 86528)	0
dense_1 (Dense)	(None, 128)	11,075,712
dropout (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 2)	258

Total params: 11,168,642 (42.60 MB)
Trainable params: 11,168,642 (42.60 MB)
Non-trainable params: 0 (0.00 B)

Figure 3. CNN Modeling

The image above shows the modeling results of CNN Total Parameters totaling 11,168,642 total parameters used in the model, including trainable and non-trainable parameters. Trainable Parameters totaling 11,168,642 and Non-Trainable Parameters totaling 0 because there are no fixed parameters that are not trained in this model.

b. SVM Modeling

The following is an image of the results of Support Vector Machine (SVM) modeling with an evaluation of the combination of C parameters and kernel types (RBF, Polynomial, and Sigmoid). This table shows the accuracy values obtained from each combination of parameters.

	C	Kernel	Accuracy
0	0.1	rbf	0.85
1	1.0	rbf	0.92
2	10.0	rbf	0.88
3	100.0	rbf	0.89
4	0.1	poly	0.78
5	1.0	poly	0.85
6	10.0	poly	0.80
7	100.0	poly	0.82
8	0.1	sigmoid	0.75
9	1.0	sigmoid	0.81
10	10.0	sigmoid	0.79
11	100.0	sigmoid	0.80

Figure 4. SVM Modeling

The image above shows the results of the RBF Kernel giving the best accuracy results at C = 1.0 (0.92) and fairly stable results at other C values. The Polynomial (poly) Kernel has lower performance than RBF. The best value occurs at C = 1.0 (0.85), but decreases at C = 100.0 (0.82) and the Sigmoid Kernel shows lower overall performance than RBF and Polynomial, with the best accuracy at C = 1.0 (0.81).

2.6. Evaluation

The evaluation stage represents the final phase in the KDD process, where the model's outcomes are assessed to ensure they align with the existing facts. It is crucial to test the research findings to determine the accuracy of the results obtained. In this study, the data testing process involves calculating accuracy, precision, recall, and F1-Score metrics.

3. RESULTS AND DISCUSSION

The following are the results of the Accuracy of the comparison of the two methods that have been carried out to determine the most optimal method in detecting drowsiness in car drivers. This Accuracy process aims to identify the advantages of each method, so that it can produce the most effective and accurate approach in the application of drowsiness detection.

3.1. Accuracy Results Using CNN Method

In this sub-chapter, the accuracy results obtained from the implementation of the Convolutional Neural Network (CNN) method in detecting driver drowsiness will be presented. The CNN method is used because of its advantages in processing image data, especially in recognizing relevant visual patterns and features. The following is a table of accuracy results:

1st Proceeding of International Conference on Science and Technology UISU (ICST)
"Contribution of Engineering Disciplines to Achieving Sustainable Development Goals Research Realizing the Vision of Golden Indonesia 2045"
ISSN: 3063-7929 (Online)
DOI: https://doi.org/10.30743/icst

Table 1. CNN Accuracy Results

	Precision	Recall	F1-Score
Accuracy			0.98
Closed	0.96	1.00	0.98
Open	1.00	0.96	0.98

The model has an overall accuracy of 98%, meaning the model is able to classify 98% of the 100% images correctly. For the closed eye category, the model has a precision of 96%, indicating that out of all the closed eye predictions, 96% were correct. Recall for this category reaches 100%, meaning the model did not miss a single image of a truly closed eye. The combination of precision and recall results in an F1-Score of 98% in this class.

Meanwhile, for the open eye category, the model has a perfect precision of 100%, meaning that all open eye predictions are correct. However, the recall is slightly lower at 96%, indicating that some open eye images were not detected properly. Even so, the F1-Score remains at 98%. Overall, the average precision, recall, and F1-Score for both categories are each at 98%, indicating excellent performance. This model is quite reliable in detecting eye conditions, with a small chance of prediction errors. The following is a graphical display of the CNN model evaluation showing the performance of accuracy, precision, recall, and F1-Score. This graph illustrates that the model has an overall accuracy of 98%, with high precision, recall, and F1-Score on both classes (open and close), indicating consistent and reliable performance in the classification task.

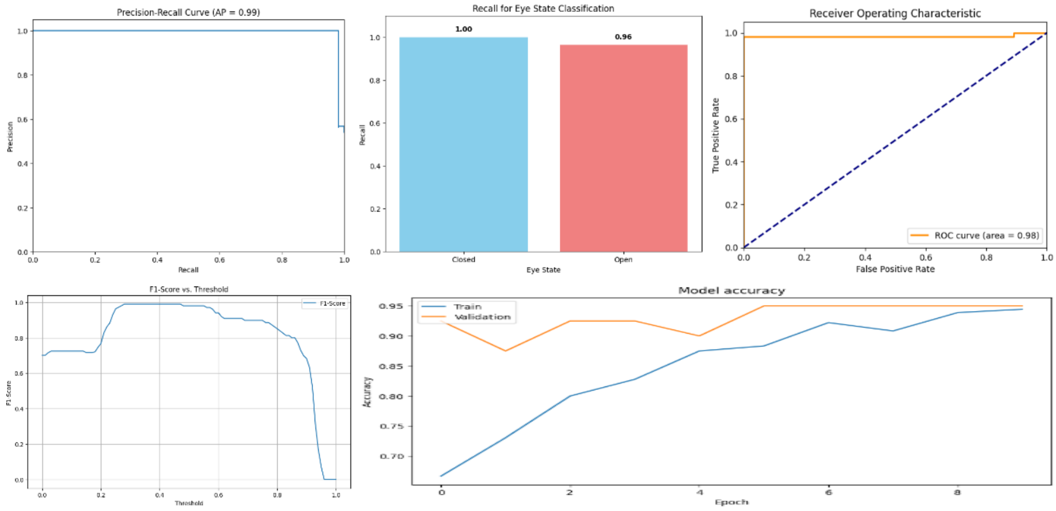

Figure 5. Accuracy Graph

3.2. Accuracy Results Using SVM Method

In this sub-chapter, the accuracy results obtained from the implementation of the Support Vector Machine (SVM) method in detecting driver drowsiness will be presented. The SVM method is used because of its advantages in processing image data, especially in recognizing relevant visual patterns and features. The following is a table of accuracy results:

Table 2. SVM Accuracy Results

	Precision	Recall	F1-Score
Accuracy			0.95
Closed	0.98	0.93	0.95
Open	0.92	0.98	0.95

The model has an overall accuracy of 95%, meaning the model is able to classify 95% of 100% of the images correctly. For the closed eye category, the model has a precision of 98%, indicating that out of all the closed eye predictions, 98% were correct. Recall for this category reaches 93%, meaning that 93% of the actual closed eyes were successfully detected. The combination of precision and recall results in an F1-Score of 95% in this class.

Meanwhile, for the open eye category, the model has a precision of 92%, meaning that the open eye prediction is 92% correct. However, the recall is higher, at 98%, indicating that there are some open eye images that are detected well. Thus, the F1-Score remains at 95%. Overall, the average precision, recall, and F1-Score for both categories are each at 95%, indicating very good performance. This model is quite reliable in detecting eye conditions, with a small chance of prediction errors.

ISSN: 3063-7929 (Online)
DOI: https://doi.org/10.30743/icst

The following is a graphical display of the SVM model evaluation showing the performance of accuracy, precision, recall, and F1-Score. This graph illustrates that the model has an overall accuracy of 95%, with high precision, recall, and F1-Score on both classes (open and close), indicating consistent and reliable performance in the classification task.

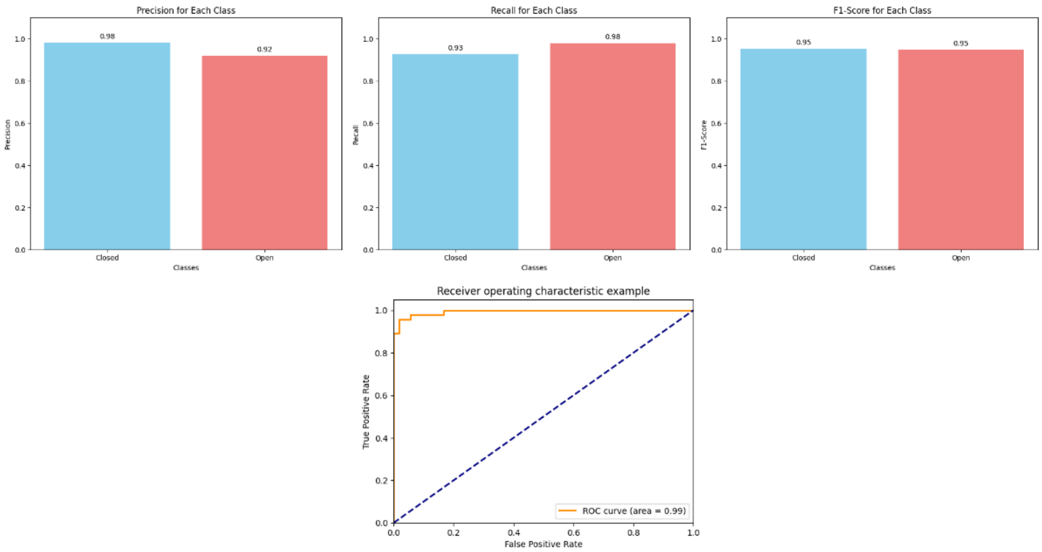

Figure 6. Accuracy Graph

4. CONCLUSION

Based on the results of the comparative study conducted, it can be concluded that the application of the Convolutional Neural Network (CNN) method for image detection in car drivers in detecting drowsiness shows a very good level of accuracy, reaching 98%. The use of deep learning methods applied to CNN has proven to be an effective solution, because it is able to produce high accuracy values. The process of implementing this method involves several main stages. The first stage is collecting image data using the Kaggle platform, which provides a dataset in the form of images of eyes in open and closed conditions. The next stage is modeling, which is designed to identify and describe the characteristics of open and closed eyes accurately. At the evaluation stage, this model showed optimal performance, with an accuracy of 98%. The results of this study strengthen the effectiveness of the CNN method in detecting drowsiness in drivers, which has the potential to be applied more widely to driving safety applications. These findings also confirm that deep learning technology can make a significant contribution to overcoming real-world problems, especially in supporting early detection systems in vehicles.

ACKNOWLEDGMENT

The author would like to express his deepest gratitude to Prof. Dr. Ir. Roslina, M.I.T and Prof. Dr. Rika Rosnelly, S. Kom., M. Kom, as Supervisor I and Supervisor II who have provided direction and guidance in completing this research. As well as all related parties that I cannot mention one by one.

REFERENCES

[1] S. N. Himawan, R. Sohiburroyan, N. B. Nugraha, J. Teknik Informatika, and P. Negeri Indramayu, "Deteksi Kantuk Pengemudi Menggunakan Deep Learning," *Semin. Nas. Ind. dan Teknol. (SNIT), Politek. Negeri Bengkalis*, no. November, pp. 1–8, 2022.
[2] R. Boutaba *et al.*, "Open Access A comprehensive survey on machine learning for networking : evolution , applications and research opportunities," 2018.
[3] T. Zhang *et al.*, "Evaluation of different machine learning models and novel deep learning-based algorithm for landslide susceptibility mapping," *Geosci. Lett.*, vol. 9, no. 1, 2022.
[4] R. Florez, F. Palomino-Quispe, R. J. Coaquira-Castillo, J. C. Herrera-Levano, T. Paixão, and A. B. Alvarez, "A CNN-Based Approach for Driver Drowsiness Detection by Real-Time Eye State Identification," *Appl. Sci.*, vol. 13, no. 13, 2023.

ISSN: 3063-7929 (Online)
DOI: https://doi.org/10.30743/icst

[5] R. Andonie and A. C. Florea, "Weighted random search for CNN hyperparameter optimization," *Int. J. Comput. Commun. Control*, vol. 15, no. 2, pp. 1–11, 2020.
[6] E. Karypidis, S. G. Mouslech, K. Skoulariki, and A. Gazis, "Comparison Analysis of Traditional Machine Learning and Deep Learning Techniques for Data and Image Classification," vol. 21, pp. 122–130, 2022.
[7] C. Choi, J. Kim, J. Hyun, Y. Kim, and B. Moon, "Face Detection Using Haar Cascade Classifiers Based on Vertical Component Calibration".
[8] S. Abidin, "Deteksi Wajah Menggunakan Metode Haar Cascade Classifier Berbasis Webcam Pada Matlab," *J. Teknol. Elekterika*, vol. 2, no. 1, p. 21, 2018.
[9] I. Monika Parapat and M. Tanzil Furqon, "Penerapan Metode Support Vector Machine (SVM) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 10, pp. 3163–3169, 2018, [Online]. Available: http://j-ptiik.ub.ac.id
[10] P. Benchawattananon *et al.*, "Automatic classification of spread - F types in ionogram images using support vector machine and convolutional neural network," *Earth, Planets Sp.*, 2024.
[11] Y. Albadawi, M. Takruri, and M. Awad, "A Review of Recent Developments in Driver Drowsiness Detection Systems," *Sensors*, vol. 22, no. 5, pp. 1–41, 2022.
[12] Y. Ding, "Learning from label proportions for SAR image classification," 2017, doi: 10.1186/s13634-017-0478-8.
[13] H. Naeem, B. M. Alshammari, and F. Ullah, "Explainable Artificial Intelligence-Based IoT Device Malware Detection Mechanism Using Image Visualization and Fine-Tuned CNN-Based Transfer Learning Model," *Comput. Intell. Neurosci.*, vol. 2022, 2022.
[14] H. J. Jie and P. Wanda, "Runpool: A dynamic pooling layer for convolution neural network," *Int. J. Comput. Intell. Syst.*, vol. 13, no. 1, pp. 66–76, 2020.
[15] H. Darmanto, D. Learning, T. Learning, and G. Descent, "PENGENALAN SPESIES IKAN BERDASARKAN KONTUR OTOLITH," vol. 2, 2019.
[16] Y. Chen *et al.*, "Deep Learning – Based Brain Computed Tomography Image Classification with Hyperparameter Optimization through Transfer Learning for Stroke," 2022.
[17] M. Azzam Al Fauzie and J. Akhir Putra, "Clustering Data Menggunakan Metode K-Means untuk Rekomendasikan Pembelajaran Akademik bagi Siswa Aktif dalam Ekstrakurikuler," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 1, pp. 642–648, 2023.