Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)
Vol. 8, No. 2, July 2025, pp. 309 – 317
p-ISSN: 2614-3372 | e-ISSN: 2614-6150

Ideal Temperature Classification of Meeting Rooms Using You Only Look Once Architecture Version 8 and Multilayer Perceptron Based on Human Density Image Data

[1]Naufal Taufiq Ridwan, [2]Winda Yulita, [3]Rahman Indra Kesuma, [4]Uri Arta Ramadhani, [5]Radhinka Bagaskara
[1,2,3,5]Program Studi Teknik Informatika, Institut Teknologi Sumatera, Indonesia
[4]Program Studi Teknik Elektro, Institut Teknologi Sumatera, Indonesia
Email: [1]naufal.120140044@student.itera.ac.id, [2]winda.yulita@if.itera.ac.id, [3]rahman.indra@if.itera.ac.id, [4]uri.ramadhani@el.itera.ac.id, [5]radhinka.bagaskara@if.itera.ac.id

Article Info
Article history:
Received Dec 13th, 2024
Revised Feb 20th, 2025
Accepted Mar 2nd, 2025

Keyword:
AC Temperature Classification
Human Density
Multi-Layer Perceptron
YOLOv8

ABSTRACT
Indonesia, located along the equator, experiences a tropical climate that results in high indoor temperatures. Elevated temperatures can affect health, making Air Conditioning (AC) necessary to regulate indoor environments. However, improper use of AC systems, such as leaving them on even when a room is unoccupied, can lead to significant energy waste. This research focuses on the efficient use of AC systems through the integration of sensors and cameras, combining two distinct technologies. The first technology is object detection using You Only Look Once (YOLOv8), which was chosen for its superior performance in terms of speed, accuracy, and computational efficiency. The second is the classification of optimal AC temperatures using the Multilayer Perceptron (MLP) algorithm, selected for its high performance in accuracy, sensitivity, and speed. In addition, the study takes into account human density in the room to optimize temperature regulation. The integration of object detection and temperature classification technologies enables the system to operate in real time and automatically adjust temperature settings based on dynamic room conditions. The research successfully implemented YOLOv8 for object detection and Multilayer Perceptron for optimal room temperature classification. Test results showed precision, recall, and F1-score values of 0.82, 0.92, and 0.86, respectively.

Corresponding Author:
Winda Yulita,
Teknik Informatika,
Institut Teknologi Sumatera,
Jl. Terusan Ryacudu, Way Huwi, Kec. Jati Agung, Kabupaten Lampung Selatan, Lampung
Email: winda.yulita@if.itera.ac.id
DOI: http://dx.doi.org/10.24014/ijaidm.v8i2.34230

1. INTRODUCTION

Indonesia is located on the equator so it receives a lot of sunlight which makes Indonesia have a tropical climate [1]. This tropical climate causes the air temperature in certain areas, especially indoors, to be high and uncomfortable. In addition, the density of people in meeting rooms has a direct effect on the increase in room temperature. According to research by Lee and Park, each individual in a closed room can contribute to an average temperature increase of 0.5°C per hour due to body heat produced [2]. This causes the need for cooling to increase significantly as the number of people in the room increases. In addition, similar research states that high-density meeting rooms without good temperature control can cause a decrease in productivity of up to 25% due to room discomfort [3].

To maintain room comfort, the use of air conditioning technology such as Air Conditioners (AC) is a common solution. However, the use of AC is often inefficient. For example, many users tend to leave the

AC on even though the room is empty or only used briefly [4]. According to Randazzo et al., AC consumes around 35%-42% of the total electricity usage in an office, making it one of the largest contributors to electricity waste [5].

Many previous studies have been conducted for AC electricity savings, including detecting the presence of people using various sensors and cameras [6][7][8], but temperature determination is still using rule-based. The rule-based approach tends to require manual adjustment and maintenance by humans, which can be impractical or inefficient in dynamic environments [9]. Therefore, an ideal AC temperature classification technique based on machine learning based on human density is needed.

One way is to use YOLO and Multilayer Perceptron (MLP) which has proven to be ideal for detecting and classifying objects. This is in accordance with several studies conducted on different objects, such as research related to the Integration of MLP with YOLO for object detection in medical images in the form of brain tumors by combining YOLO for early detection and MLP for tumor type classification [10]. Other research is related to the Combination of YOLO and MLP for vehicle detection and classification in intelligent transportation systems. This study uses YOLO to detect vehicles on the highway and MLP to classify vehicle types [11]. Another study is the prediction of visibility, namely the distance of an object's visibility with the naked eye during the day which is used for air traffic purposes to determine landings, takeoffs, and others. The prediction was carried out using YOLO and MLP. Based on this study, the MLP algorithm showed better overall performance results. Additionally, the use of the MLP algorithm helps mitigate issues related to the limited number of objects detected in shorter ranges [12].

The research we conducted used the YOLOv8 Algorithm to detect human head objects and the Multilayer Perceptron (MLP) algorithm for ideal temperature classification. The YOLOv8 algorithm has better performance in terms of speed, accuracy and computational load required compared to other object detection algorithms [13][14][15]. Meanwhile, the Multilayer Perceptron (MLP) algorithm was chosen because the algorithm has good performance in terms of accuracy, sensitivity and speed [16]. This MLP algorithm will be used for the ideal AC temperature classification process from the extraction of head features generated from YOLOv8. This combination of algorithms has never been done before to determine the ideal AC temperature based on room density.

2. RESEARCH METHOD

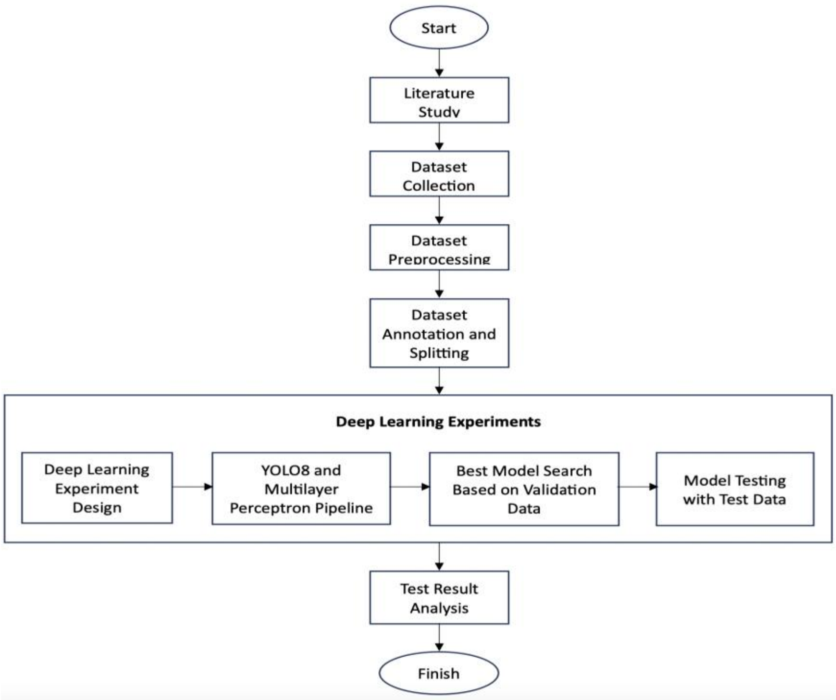


Figure 1. Research flow

This research was done by designing, implementing, and testing the performance of both software and hardware. After the performance results of that sub-system are confirmed, then the hardware and software units are integrated. After the integration was completed, the whole system was tested carefully and comprehensively to ensure proper functionality and seamless operation of the integrated system. In this work, we implement the algorithm YOLOv8 and Multilayer Perceptron (MLP) for AC-based temperature classification room image with data pipelining process from the YOLOv8 model and MLP, where YOLOv8 is responsible for the head detection process and MLP is responsible for the temperature classification process. The research flow carried out is shown in Figure 1.

2.1 Data and Software

2.1.1 Dataset Collection

In this research, two datasets were used to train YOLOv8 and MLP, respectively, to detect head objects and classify AC temperature recommendations. The first dataset is sourced from Roboflow and will be used to train YOLOv8, while the second dataset was collected directly from a meeting room in Building A of the Sumatra Institute of Technology. Both datasets consist of images of rooms containing human objects. The first dataset contains 2,390 images, and the second dataset consists of 415 images.

2.1.2 Preprocessing

To build a model, image preprocessing is a crucial step before conducting deep learning experiments [17]. In this phase, the images will undergo a resizing process. The dataset obtained contains images of relatively large size, so resizing is necessary to reduce and standardize the size of all images in the dataset to 320 x 320 pixels. This also aims to achieve good efficiency during the model training process.

2.1.3 Annotation

In the first dataset, the data used is a collection of images containing head images sourced from Roboflow. These images have been labeled for each head part in the form of a .txt file, which contains information related to the position and relative size of the head's bounding box. Each line represents one bounding box, consisting of several data attributes: class, x_{min}, y_{min}, height, and width. The x and y coordinates are presented in a relative scale ranging from [0, 1].

Then, the annotation process was also performed on the second dataset by labeling each head object in every image and assigning labels for the recommended temperature class. The labeling of temperature classes was conducted through a survey involving several respondents, including experts in the field of architecture who study thermal comfort and general individuals who spend their daily activities in air-conditioned rooms. Each image was given three options representing each class: cool, moderate, and cold.

2.1.4 YOLOv8 Implementation

1) YOLOv8 Architecture

The primary function of the YOLOv8 model in this research is to detect the head area in an image. In this phase, there are several steps in the detect_image() function. First, the model is loaded using Torch. Then, the image is read from the image parameter. Since the OpenCV library returns the image in BGR format, it is then transformed into RGB format and resized to 320 x 320. Next, the model performs the prediction. The model output is then processed through the non_max_supression() function to prevent duplicate detections of the same object [19].

2) YOLOv8 Hyperparameter

The range of values or types of hyperparameters tested for YOLOv8 is shown in Table 1.

Table 1. Hyperparameters tested for YOLOv8

	Parameter
Optimizer	SGD, Adam W, RMSProp
Epoch	min = 0, max = 100
Batch	8, 16
Learning Rate	0.01, 0.001

Having the above configuration, there are 12 hyperparameter combinations which will be then processed in the model training phase.

3) YOLOv8 Model Training

Having the hyperparameters combination from the previous phase, by doing a training model, the combination of number 8 showed the best result since having the highest mAP@50-95 (0.328), excellent performance at detecting heads with varying degrees of overlap, which is important for the consistency of detection in real conditions. Despite having Validation Loss slightly higher (4,712) and mAP@50 slightly lower (0.411) compared to the 2nd hyperparameter combination, this combination has a higher value of mAP@75 (0.137). So the 8th combination is chosen to be the most suitable hyperparameter configuration for the model and dataset in this research.

2.1.5 MLP Implementation

1) MLP Architecture

In this step, the MLP-feed forward architecture was created using 3 main layers, namely the input layer, hidden layer, and output layer.

2) MLP Hyperparameters

The range of values or types of hyperparameters tested for MLP is shown in Table 2.

Table 2. Hyperparameters tested for MLP

	Parameter
Hidden Neuron	[64, 32], [128 64]
Epoch	min = 0, max = 100
Batch	8, 16
Activation Function	ReLu, Sigmoid
Optimizer	Adam, SGD, RMSProp

By considering the above parameter, there are 24 hyperparameter combinations obtained which will be then processed in the model training phase.

3) MLP Model Training

Next, the training process is carried out for each hyperparameter combination which will be tested. This process is done to get the best model for the process AC temperature classification. Nevertheless, the second dataset is used to generate a feature map before training begins. This is done because the MLP model will be used as a classification process of the feature map resulting from the detection carried out by YOLOv8.

2.1.6 Pipelining YOLOv8 and MLP

The system will start by detecting the head area using YOLOv8. From the prediction results, feature extraction will then be carried out. Later it will be used by MLP for the air conditioner temperature classification process. The resulting feature map is a head area obtained from detection results by YOLOv8. The background of the feature map is blue, and other colors besides the background color indicate the presence of a head with a certain area. Then the multilayer perceptron will carry out classification based on the feature map.

2.2 Measurement Method

To assess whether a classification model has good performance or not, we can refer to its performance measurement parameters, namely precision, recall, and accuracy. Calculating these parameters involves using a matrix called a confusion matrix. Multiclass Confusion Matrix contains four important parameters, i.e. True Postitive (TP), False Negative (FN), False Positive (FP), and True Negative (TN) [20]. The accuracy and recall are calculated as equation 1 and 2.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

The precision and F1 Score are obtained with equation in the equation 3 and 4.

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$F1\ Score = \frac{2}{recall^{-1} + precision^{-1}} \tag{4}$$

3. RESULTS AND ANALYSIS

3.1 Dataset

This research uses two kinds of datasets: A dataset obtained from Roboflow and A dataset taken from ITERA’s meeting room at Building A. Both datasets are images of rooms filled with people as shown in Figure 2.

Figure 2. Sample of both datasets

Both datasets are then annotated by drawing a bounding box around each head from photos and saving each of the bounding boxes’ coordinates in a .txt file as shown in Figure 3. The results of these preprocessed datasets are then used for the model’s learning processes: Roboflow datasets for the YOLOv8 model, and the ITERA datasets for the MLP model.

0 0.2504 0.4480 0.0597 0.1051
0 0.2760 0.7411 0.1203 0.1723
0 0.5591 0.7768 0.1286 0.1910
0 0.7487 0.7249 0.1035 0.1726
0 0.9064 0.8253 0.1341 0.1898

Figure 3. Bounding box and its coordinates

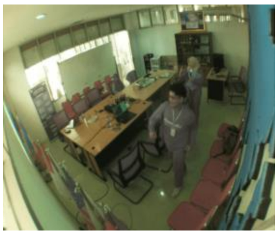
Other than the picture datasets, it’s required to label each type of room temperature for ITERA’s dataset. There are three categories for temperatures: cool, medium, and cold, as seen in Table 3. Labeling is done by conducting a survey in which respondents are experts on architecture, with the results seen in Table 4. Preprocessed datasets are then divided into three categories: training, validation, and testing, using a 75:15:10 ratio respectively. For the Roboflow dataset, it means 1630 pictures for training, 466 pictures for validation, and 233 pictures for testing. The ITERA dataset is divided into 310 pictures for training, 63 pictures for validation, and 42 pictures for testing.

Table 3. Dataset structure

Folder	Image
Cool	image_001.jpg image_002.jpg …
Medium	image_004.jpg image_005.jpg …
Cold	image_006.jpg image_007.jpg …

Table 4. Training dataset for temperature classification

Image	Survey Results			End Label
	Cool	Medium	Cold	
	0	2	8	cold

Image	Survey Results			End Label
	Cool	Medium	Cold	
	10	0	0	cool
…				

3.2 YOLOv8

YOLOv8 will be used for detecting how many human heads from a picture. Its architecture is built by PyTorch. Using PyTorch, the first step is running the load model. The dataset contains BGR pictures, in which they will be transformed into 320x320 pixels of RGB pictures. These RGB pictures are the ones that will be predicted by using YOLOv8. YOLOv8 is capable of running 64.5 billion operations per second for predictions. In order to do predictions optimally, YOLOv8 requires an optimal hyperparameter combination. Using 12 hyperparameter combinations, we can obtain the best hyperparameter by doing the training process as shown in Table 5.

Table 5. YOLOv8's hyperparameter combination

Combinations	Epoch	Optimizer	Learning Rate	Batch Size	Val Loss	mAP@50:95
1	[0, 100]	SGD	0.01	8	5.162	0.299
2	[0, 100]	AdamW	0.01	8	4.707	0.308
3	[0, 100]	RMSProp	0.01	8	6.172	0.155
4	[0, 100]	SGD	0.001	8	4.875	0.296
5	[0, 100]	AdamW	0.001	8	4.758	0.301
6	[0, 100]	RMSProp	0.001	8	5.938	0.142
7	[0, 100]	SGD	0.01	16	4.909	0.286
8	[0, 100]	AdamW	0.01	16	4.712	0.328
9	[0, 100]	RMSProp	0.01	16	5.971	0.149
10	[0, 100]	SGD	0.001	16	4.856	0.289
11	[0, 100]	AdamW	0.001	16	4.720	0.316
12	[0, 100]	RMSProp	0.001	16	5.727	0.139

The experiments showed that the 8th combination resulted in the most optimal configuration, with the highest mAP@50:95 score of 0.328. This means that the combination can provide a solid performance and consistency for predicting.

3.3 MLP

MLP will be used for classifying the temperature of the air conditioner, using the previous YOLOv8's prediction result. MLP is implemented using MLP-feed forward with 3 main layers: input layer, hidden layer, and output layer. Much like YOLOv8, MLP also requires the most optimal hyperparameter configuration [18]. Before finding the optimal configuration via training, however, both datasets went through a process called feature map generation as shown in Figure 4, using the results from YOLOv8. These feature maps will be used in the classification process by MLP.

The most optimal hyperparameter configuration for MLP is the hyperparameter with hidden neurons 128, and 64, activation function ReLu, batch size 16, and optimizer Adam. This hyperparameter scored a relatively high accuracy of 88.24%, resulting in a satisfactory hyperparameter.

3.4 YOLOv8 and MLP's Pipeline

After designing the YOLOv8 and MLP's architecture, and obtaining both models' optimal hyperparameters, the next step is the pipelining process. The input data will be run through the YOLOv8 model to predict how many heads are from each picture. The prediction then continued into the feature extraction process, resulting in a feature map that will be classified using the MLP model as shown in Figure 5. The result is a classification for air conditioner temperature.

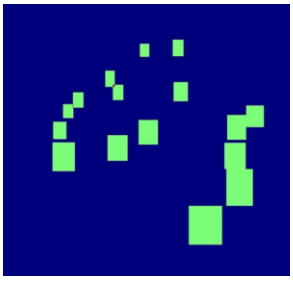


Figure 4. Feature map

Figure 5. YOLOv8 prediction and feature extraction

Afterward, the testing will be conducted to evaluate the performance of the pipeline. For YOLOv8's test will be using 233 pictures, while for the MLP & pipeline test will be using 42 pictures. Each test will result in confusion matrices.

3.5 Results

In general, the results have been satisfactory. Table 6 shows that YOLOv8 testing produces good performance, with a 0.71 precision score, but the score is still acceptable. YOLOv8's results overall gave a solid and consistent performance for object detection.

Table 6. Test results

	Tests	Precision	Recall	F1-Score
YOLOv8	Positive Class	0.82	0.89	0.85
	Negative Class	0.71	0.89	0.78
MLP	Positive Class	0.87	0.98	0.92
	Negative Class	0.82	0.98	0.90
Pipeline YOLOv8 & MLP	Positive Class	0.82	0.92	0.86
	Negative Class	0.81	0.81	0.81

Meanwhile, MLP also produced good results. The point of MLP's test is to evaluate how good MLP model is in predicting temperature class. Both positive and negative classes resulted above 0.80 in every category, and some even reached as high as 0.98 for recall score. This means that the MLP model is having a great performance for temperature classification.

The pipeline also resulted in good scores, with positive class resulting in 0.82, 0.92, and 0.82 for precision, recall, and F1-score respectively. The balance between precision and recall results on positive class compared to the negative class' results, may show that the pipeline model is better at predicting positive cases. On top of this, pipelining YOLOv8 and MLP resulted in a fast prediction with an average of ± 1 second of execution time.

4. CONCLUSION

In this research, we managed to implement temperature classification of rooms using YOLOv8 and MLP. YOLOv8 has the role of detecting how many humans are in the room, while MLP's role is to classify the ideal room temperature based on how many heads are detected using YOLOv8. The result of the pipeline is 0.82, 0.92, and 0.82 for precision, recall, and F1-score. The optimal configuration is hyperparameter optimizer AdamW, batch size 16 and learning rate 0.001 for YOLOv8, and hyperparameter hidden size (128, 64), activation function ReLu, batch size 16, and optimizer Adam for MLP. Pipelining YOLOv8 and MLP requires an average of ± 1 second of execution time to provide a classification prediction.

For future research, there are many aspects that can be done to improve the prediction precision, recall, and F1-score respectively. It requires further experiments using many untested hyperparameter configurations that may produce better results. The addition of parameters such as average room temperature and humidity can improve the model performance. Furthermore, the image dataset can be improved by using new images from multiple angles.

ACKNOWLEDGEMENTS

This research is supported by Direktorat Riset, Teknologi, dan Pengabdian kepada Masyarakat; Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi; Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia; 2024.

REFERENCES

[1] Boer, R., & Faqih, A. Climate Change Impacts in Indonesia: Challenges and Opportunities for Adaptation. Climatic Change. 2018; 151(2), 245-258. DOI: 10.1007/s10584-018-2267-4
[2] Lee, H., & Park, S. (2021). Analysis of Energy Wastage in Office Air Conditioning Systems. Energy and Buildings. 2021;150, 456–470.
[3] Johnson, K., & Miller, D. Health Impacts of Prolonged Exposure to High Temperatures Indoors. International Journal of Public Health. 2020; 65(2), 123–132.
[4] Pathan, M. M., Shaikh, Z. A., & Gupta, S. A survey on occupancy-based energy management and control strategies for heating, ventilating, and air conditioning systems. Sustainable Cities and Society, 2020; 53, 101994
[5] Randazzo, T., et al. Energy Consumption Patterns in Air Conditioning Systems. Renewable Energy Journal. 2020; 95, 85–98.
[6] Chen, P., et al. Human Presence Detection Using Motion Sensors for Energy Saving. Sensors and Actuators A: Physical. 2019; 305, 112–125.
[7] Huang, Z., & Lin, J. Optimizing Indoor Cooling with Smart Sensors. Automation in Construction. 2020; 120, 103395.
[8] Sun, W., & Liu, H. Rule-Based Systems for Indoor Climate Management. Journal of Artificial Intelligence Research; 2021; 48, 78–90.
[9] Redmon, J., & Farhadi, A. YOLOv3: An Incremental Improvement. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020; 1–10.
[10] Zhang, Y., Li, X., & Wang, J. Brain Tumor Detection Based on YOLOv3 and Multilayer Perceptron. IEEE Access, 2020; 8, 83361-83370.
[11] Li, S., Zhao, M., & Liu, Y. Vehicle Detection and Classification in Complex Scenarios Using YOLO and Multilayer Perceptron. IEEE Transactions on Intelligent Transportation Systems. 2021; 22(7), 4238-4249.
[12] M. Janok, R. Forgáč and B. Badidová. A Contribution to the Prevailing Visibility Prediction Based on Neural Network Approach, IEEE, 2024 New Trends in Signal Processing (NTSP), Demanovska Dolina, Slovakia, 2024; pp. 1-6, doi: 10.23919/NTSP61680.2024.10726309.
[13] R. Varghese and S. M. YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), Chennai, India, 2024; pp. 1-6, doi: 10.1109/ADICS58448.2024.10533619.
[14] Y. Fang, W. Huang, C. Zheng, X. Huang, M. -T. Lin and J. -K. Yang, Comparative Analysis of Object Detection Algorithms for Wood Defect Detection, 2024 IEEE 7th International Conference on Electronic Information and Communication Technology (ICEICT), Xi'an, China, 2024; pp. 991-996, doi: 10.1109/ICEICT61637.2024.10670842.
[15] Sohan, M., Sai Ram, T., Rami Reddy, C.V. A Review on YOLOv8 and Its Advancements. Springer. 2023; Singapore. https://doi.org/10.1007/978-981-99-7962-2_39
[16] Bikku, T. Multi-layered deep learning perceptron approach for health risk prediction. J Big Data **7**, 50 2020. https://doi.org/10.1186/s40537-020-00316-7
[17] Koresh, H.J.D. Impact of the Preprocessing Steps in Deep Learning-Based Image Classifications. Natl. Acad. Sci. Lett. 2024; 47, 645–647. https://doi.org/10.1007/s40009-023-01372-2
[18] L. Puurunen, J. Chaudhary, R. Kanth and J. Heikkonen. Human Presence Detection with Thermal Sensor using Multilayer Perceptron Algorithm, 2021 IEEE 7th World Forum on Internet of Things (WF-IoT), New Orleans, LA, USA, 2021; pp. 669-673, doi: 10.1109/WF-IoT51360.2021.9595529.
[19] M. Ivašić-Kos, M. Krišto and M. Pobar, Human detection in thermal imaging using YOLO, Proceedings of the 2019 5th International Conference on Computer and Technology Applications, 2019.
[20] M. Rizk and I. Bayad. Human Detection in Thermal Images Using YOLOv8 for Search and Rescue Missions. 2023 Seventh International Conference on Advances in Biomedical Engineering (ICABME), Beirut, Lebanon. 2023; pp. 210-215, doi: 10.1109/ICABME59496.2023.10293139.

BIBLIOGRAPHY OF AUTHORS

Naufal Taufiq Ridwan is an Informatics Engineering student at the Sumatra Institute of Technology.

Winda Yulita as a lecturer in Informatics Engineering at the Sumatra Institute of Technology. I am a member of the Artificial Intelligence and Data Engineering (AIDE) Research Group. My research is related to Artificial Intelligence (AI) such as text mining.

Rahman Indra Kesuma is a lecturer at Institut Teknologi Sumatera (ITERA) in the Informatics Engineering Undergraduate Program. Competence and mastery in artificial intelligence, expert systems, decision support systems, machine learning, and deep learning. Currently, has focusing on dissertation research in video anomaly detection and prediction with self-supervised learning at the School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia.

Uri Arta Ramadhani as a lecturer in Electrical Engineering at the Sumatra Institute of Technology. I am part of the Telecommunication, Electronics, and Computer Systems Research Group. My research is related to radio channel propagation, intelligent surfaces, signal processing, and cellular networks.

Radhinka Bagaskara as a lecturer in Informatics Engineering at the Sumatra Institute of Technology. I am a member of the Artificial Intelligence and Data Engineering (AIDE) Research Group. My research is related to Artificial Intelligence (AI) such as gaming.