



Enhancing Face Authentication for Online Examination Systems Using Median Filtering and MobileNetV2

Dadang Sudrajat¹, Dian Ade Kurnia^{2*}, Rudi Kurniawan³, Othman bin Mohd⁴, Maulana Sujarwadi⁵,
Salman Alfarizi⁶

^{1,5}Program Studi Teknik Informatika, STMIK IKMI, Cirebon, Indonesia

²Program Studi Manajemen Informatika, STMIK IKMI, Cirebon, Indonesia

^{3,5}Program Studi Rekayasa Perangkat Lunak, STMIK IKMI, Cirebon, Indonesia

⁴Fakulti Teknologi Maklumat dan Komunikasi, Universiti Teknikal Malaysia, Melaka, Malaysia

¹dianade2014@gmail.com, ²d42ngsudrajat2012@gmail.com, ³rudi226@gmail.com, ⁴mothman@utem.edu.my,

⁵maulanasujarwadi876@gmail.com, ⁶salmanalfariziapollo@gmail.com

Abstract

Digital transformation in higher education is driving the uptake of online tests, which require academic integrity, security, and robust user experience. In the context of authentication of users, deep learning based face recognition, in particular the Convolutional Neural Network (CNN) architectures, such as MobileNetV2, combined with intermediate filter, promises to deliver a consistent performance across a wide range of devices and imaging environments. However, there are limited comprehensive studies evaluating the final integration of the median filter and MobileNetV2 in high-value test scenarios. This study contributes by proposing an effective end-to-end Face Authentication Pipeline, assessing the median impact of filtering on MobileNetV2 performance, and validating it with a prototype application. The authentic face dataset was collected using the Teachable Machine, preprocessed with cropping, resizing, and median filtering, and then augmented through rotation, shift, shear, zoom, reversal, and brightness adjustment. The MobileNetV2 model was trained with Adam in a stepwise manner, starting with 0.001 and then 0.0001 for 20 epochs in a batch size of 32, and was evaluated for accuracy, precision, recall, and F1 score. Results show that the accuracy curve has remained stable at almost 95 percent during the 20th epoch; most grades achieved 1.00 in both precis, recall and F1, with some classings showing a limited decrease due to facial similarity or expression differences. These findings confirm that MobileNetV2 median filtering can be the basis for an effective, accurate and ready to integrate face recognition in online testing applications on a wide range of devices.

Keywords: digital transformation; face authentication; median filtering; MobilNetV2; online assesment

How to Cite: Dadang Sudrajat, Dian Ade Kurnia, R. Kurniawan, Othman bin Mohd, Maulana Sujarwadi, and Salman Alfarizi, "Enhancing Face Authentication for Online Examination Systems Using Median Filtering and MobileNetV2", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 9, no. 6, pp. 1526 - 1537, Jan. 2026.

Permalink/DOI: <https://doi.org/10.29207/resti.v9i6.7185>

Received: September 17, 2025

Accepted: December 28, 2025

Available Online: January 8, 2026

This is an open-access article under the CC BY 4.0 License

Published by Ikatan Ahli Informatika Indonesia

1. Introduction

The digital transformation in higher education is driving changes in the way institutions design, deliver and evaluate learning [1], [2]. An important part of this effort is the online examination system, as it can improve the efficiency, accessibility and innovation of the examination process [2],[3]. Various technologies, ranging from online testing platforms to offline-online integration and biometrics [4], are now being widely explored to preserve academic integrity without

compromising the learning experience of students [5], [6]. Among these technologies, face recognition has attracted attention because it promises practical, real-time identity verification in scenarios where no proximate is physically present [7]. Student perception studies also confirm that the reliability and security of the technology is key to building trust in online learning [8], [9]. This trust is important as it influences the students' motivation and performance in the assessment.

Accordingly, studies in distance learning show that the use of face recognition can help to ensure the identity of learners and to increase trust in the evaluation process [10]. Integrative studies on the Internet of Things and biometrics also highlight the need for a reliable data infrastructure for identity-verification services to operate efficiently in the educational environment [5]. On the other hand, a discourse about supervisory ethics is still emerging [11]. Technology on-line proctors requires a balance between academic integrity and respect for students' autonomy and privacy [6]. Therefore, the development of face recognition in online examination systems needs to combine technical reliability with ethical considerations, ease of integration and computational power [3].

In reality, however, remote authentication still faces problems in various real-world environments: lighting is inconsistent, viewing angles vary, camera quality varies and computing devices are limited [12]. This condition may degrade the accuracy of the identification, cause authentication to fail, and prevent the performance of the test from being. Challenges increase in high value examinations, as the risk of impersonation and attacks remains if face-to-face supervision is omitted [10], [13]. At the same time, there is still no consensus on standard practices for combining image pre-processing and lightweight model architectures, which will consistently be the best in end devices with limited power and computing.

Attempts to use deep neural network models often rely on high-power computing. However, online examination systems running on home computers or personal devices require model efficiency to maintain a fast, resource efficient and stable validation process [14]. Pre-processing techniques such as median filtering offer opportunities to reduce noise and preserve edges, but their use for face recognition in online tests is still rarely reported in a systematic and quantifiable way. Moreover, security issues related to counterfeiting attacks, such as video and photo playback, require features to detect life signs. Without detection of life, the system risks accepting false identities, which compromises the integrity of the assessment [15]. In other words, solutions that rely solely on static recognition are no guarantee strong enough to deal with evolving attack scenarios.

In other words, solutions based solely on static recognition are not guaranteed to be robust enough to cope with evolving attack scenarios [13]. In the field of distance learning, research shows that deep learning helps in registration and validation of participants [10]. On the preprocessing side, median filtering has been proven to be effective in reducing noise without blurring the edges, so that the quality of facial features is maintained for the recognition process [16]. Studies on improving image quality in the medical field also show that using median filtering prior to the detection or classification phase improves the accuracy of the system when the input is linear [17] - [19]. Although

the context is different, the principle of signal-to-noise ratio improvement is relevant for facial authentication on online exams that are full of variations in imaging conditions.

On the other hand, MobileNetV2 is an effective choice for the architecture of the model, as it uses depth-separable convolutions and linear constraints to reduce computational complexity without sacrificing accuracy. Recent studies show MobileNetV2 is better than its predecessor and competitive against heavier architectures, and is suitable for devices with limited resources [20] - [22]. The implementation of transfer learning strategies and gradual adaptation also showed good performance in various real-world application domains, signaling the robustness that is essential for facial authentication on devices with limited capabilities [22], [23]. However, the systematic integration between median filtering and MobileNetV2 in online exam scenarios equipped with liveness detection has not been widely discussed with rigorous evaluation protocols.

As a transition, the pre-process literature enriches the methodological basis. Median filtering effectively reduces noise and preserves edge details, which is important for maintaining facial shape and texture information [17], [24]. The combination of median filtering with contrast-enhancing techniques such as equalization reportedly improves noise resistance and exposure variation in facial recognition [25]. Although some of the references are not directly in the domain of the online exam, the results give a strong indication that proper preprocessing can improve the accuracy of lightweight models. At the same time, the ethical framework and user perceptions should not be ignored; The literature shows that a sense of security, trust, and perception of control influence the acceptance of technology on online exams [6], [8], [9].

Various literature reviews show that there is a common point in educational institutions requiring reliable, lightweight, and ethical authentication for online exams; median filtering has the potential to improve image quality; MobileNetV2 provides an efficient computing foundation; and liveness detection reduces the risk of attack [5], [10], [15], [20]. However, there are still no comprehensive studies available that test the median blend of MobileNetV2 filtering in high-value online exam scenarios with an evaluation of accuracy, speed, and resistance to variations in lighting and angles, while observing aspects of user acceptance and ethical considerations of supervision. This gap leads to the need for studies that design, implement, and assess realistic end-to-end authentication.

The study aims to design and evaluate an online face recognition system combining median filtering as the primary pre-processing and MobileNetV2 as the primary architecture. The contribution offered is threefold in size. First, a structured test of the effect of the median filter on the accuracy and stability of authentication under different lighting conditions and

viewing angles. Secondly, MobileNetV2 implementation optimized for end devices, so that the verification process runs fast and is resource efficient. Thirdly, integration of life detection to reduce the risk of spoofing, while still paying attention to ethical considerations and user perception, to make the solution widely acceptable [6], [8]. The scope includes pipeline design, comparative testing, and performance analysis, which weighs accuracy, time to reach a conclusion, and resilience to simple attack. This framework will bridge the gap between the practical need to conduct on-line tests and the scientific evidence of efficient pre-treatment alloys and lightweight models.

2. Methods

This research method uses a systematic experimental quantitative approach [26]. The stages of the research are described in Figure 1, which starts from data acquisition, preprocessing, data training, model development, model evaluation, to the application of models in facial authentication applications in online exam systems.

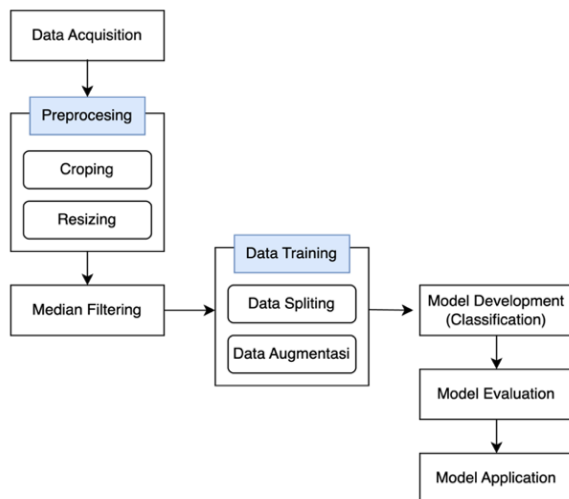


Figure 1. Stages of Research Method

Based on Figure 2, the initial stage of facial data is obtained through the *data acquisition* process which is carried out using the Teachable Machine application [27].

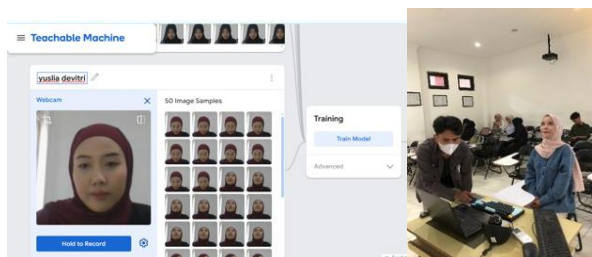


Figure 2. Data Collection Process via Teacable Machine

Images of students' faces studying are recorded live to ensure that the data used is authentic and relevant to the context of online exam authentication. The dataset was collected and tested on a sample of 31 students, which

was taken with a 720-pixel HD webcam type camera. Thus, it can represent real conditions in the implementation of online system-based exams, as well as the basis for the next stage of processing, as shown in Figure 3.



Figure 3. Sample Data generated from Teachable Machine

Furthermore, the facial data obtained is processed at the preprocessing stage to improve the consistency and uniformity of the dataset. The cropping process is done so that the face area can be separated from the background, so that the main focus remains on the object that is the target of recognition, as shown in Figure 4. It can also be visually seen on the histogram as shown in Figure 5. After that, the face image is resized into uniform dimensions, so as to ensure the fit of the input to the architectural needs of the model to be used.

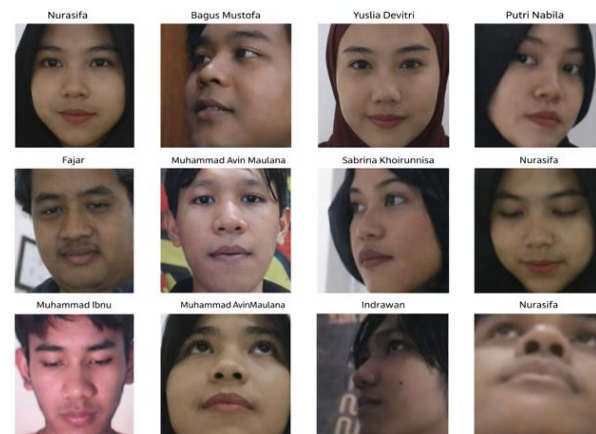


Figure 4. Original Data Sample after Cropping

In the median filtering stage, the filtering process is applied through the median filtering technique. This technique is used to reduce noise in facial images, so that the visual quality of the image can be improved without sacrificing important details [28]. With the use of median filtering, the sharpness of the facial contours and the clarity of the facial feature patterns are maintained, which ultimately contributes to the improved performance of the facial recognition model, as shown in Figure 6.

The *data training* stage was carried out by dividing the dataset using the data splitting technique [29]. The data

is divided into two subsets, namely 80% for the training data and 20% for the test data as shown in Figure 8. Thus the distribution of data can be balanced between the learning stage and the model testing stage. With this

strategy, the reliability of the model can be ensured through the evaluation of data that was not previously involved in the training process [30].

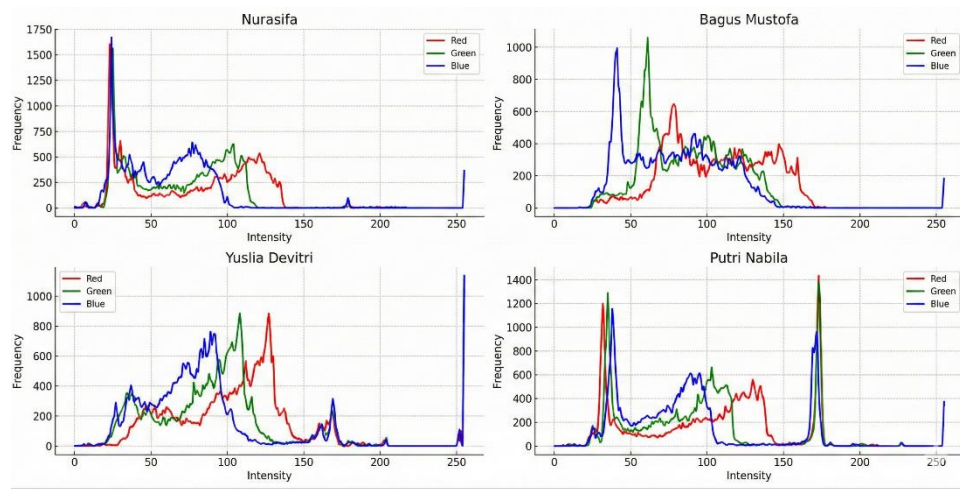


Figure 5. Histogram Sample Data Original

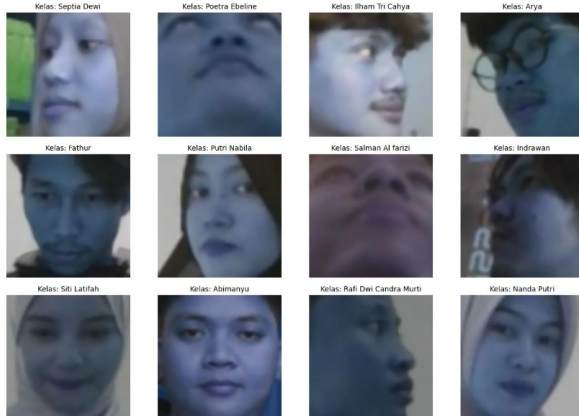


Figure 6. Sample Data after Filtering Using Median Technique

In addition to data sharing, the augmentation process is also applied at the data training stage, as shown in

Figure 9. Augmentation techniques are used to expand the diversity of data by generating new variations of existing facial imagery, i.e. through rotation, flipping, or lighting modification [31]. With this additional variation, the model can be trained on a more representative dataset, so that the generalization ability of new data can be significantly improved as presented in Table 1. The augmentation process that has been defined in Table 1 is then applied to the facial image dataset. The augmentation results are shown in Figure 9, where the facial image undergoes a variety of transformations such as rotation, shift, reversal, and lighting changes. These variations are generated to enrich the diversity of the dataset so that the model can learn from facial conditions with different viewing angles and lighting intensities.

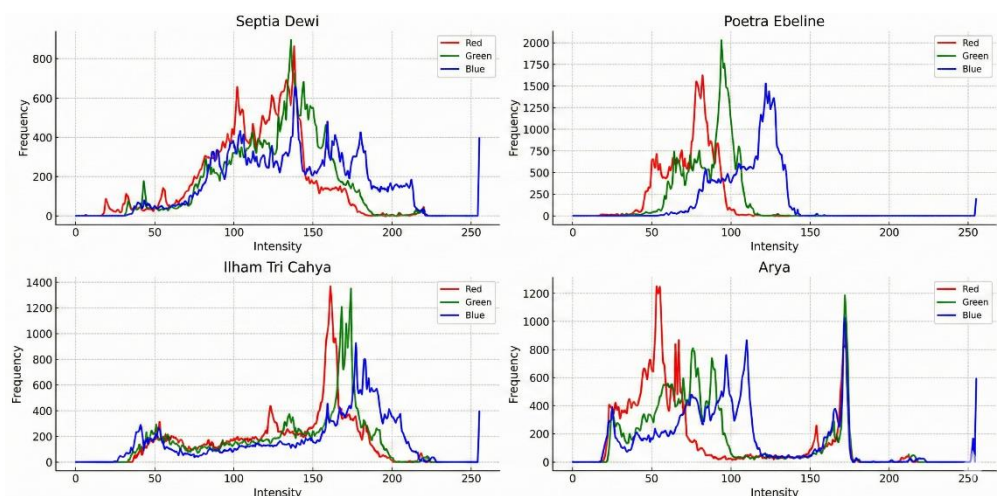


Figure 7. Histogram after Filter using the Median Technique

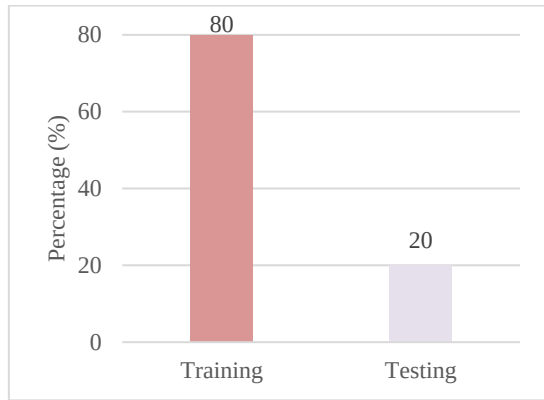


Figure 8. Data Splitting (Proportional Data Training and Testing)

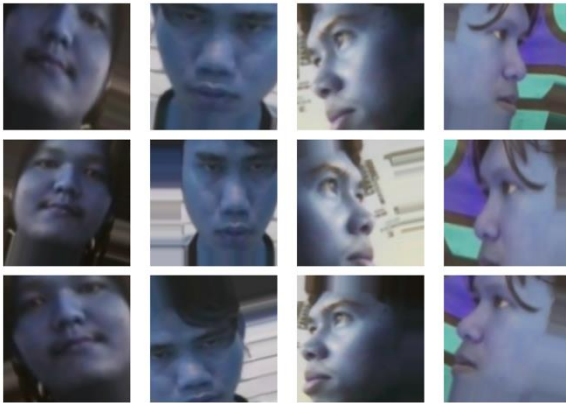


Figure 9. Augmentation Results

In addition, the color intensity distribution of the augmented image is analyzed through the RGB histogram graph shown in Figure 10. The histogram depicts the distribution of the intensity values of the red (R), green (G), and blue (B) channels on four different faces. With this analysis, it can be ascertained that the augmentation process not only produces geometric variations but also affects the characteristics of color distribution, which is important to support the model's performance in recognizing faces more robustly.

Table 1. Data Augmentation Parameters

Parameter	Description
rotation_range=30	Randomly rotate the image up to 30 degrees.
width_shift_range=0.2	Randomly swipes horizontal images up to 0.2 of their width.
height_shift_range=0.2	Randomly shift vertical images up to 0.2 of their height.
shear_range=0.2	Apply shear transformation with an intensity of 0.2.
zoom_range=0.2	Randomly zoom (zoom in/out) up to 0.2%.
horizontal_flip=True	Flip images horizontally.
brightness_range=[0.7, 1.3]	Randomly changes the brightness of the image between 0.7 and 1.3.
fill_mode="nearest"	Fill in empty pixels after transformation with the nearest pixel value.

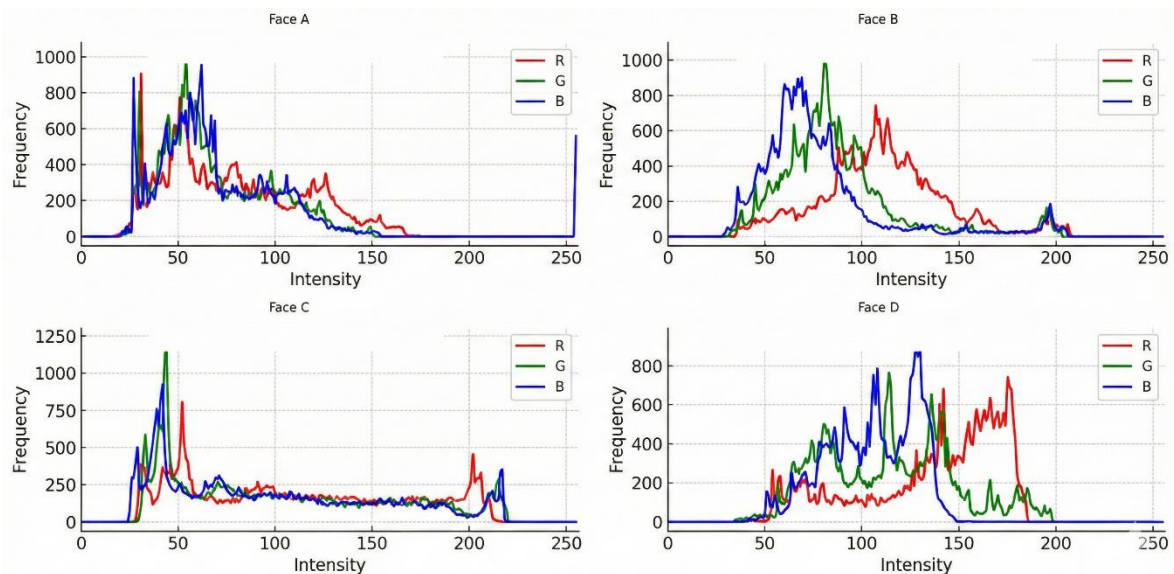


Figure 10. Histogram Augmentation

The model is built to perform classification based on pre-processed facial features. Through this architecture, the facial authentication process can be performed with a fast response time while maintaining a high level of accuracy.

Model *development* is carried out using the *MobileNetV2* architecture as presented in Figure 11, which is known to be efficient in computing and suitable for implementation on mobile devices [32].

Based on Figure 11, the process starts from the input layer, where a 224×224 pixel color image with three color channels (RGB) is fed into the network. This input size was chosen because it provides a balance between adequate image detail and computational efficiency. The image is then processed through initial convolution with the 3×3 kernel, resulting in 32 feature maps measuring 112×112 pixels. This stage serves to capture basic visual patterns such as edges, corners, and simple textures.

Furthermore, the network passes through a series of bottleneck blocks that are the hallmarks of MobileNetV2. In the first to second blocks, the spatial size of the image is reduced to 56×56 pixels. This reduction in resolution was followed by the next bottleneck block (blocks 3–6) which reduced the size to 28×28 pixels. At this stage, the network not only shrinks the spatial dimension, but also expands the number of channels so that it can store a more complex representation of features. A similar process is continued on blocks 7–10 resulting in a representation with a size of 14×14 pixels, followed by blocks 11–13 which maintain a spatial size of 14×14 but enrich the number of canals. Next, in blocks 14–16, the size is reduced again to 7×7 pixels, which is a dense representation of all the visual information that has been studied previously.

After going through the bottleneck stage, the network enters the 1×1 convolution layer which expands the number of channels to 1280, without changing the spatial dimensions. The main function of this layer is to strengthen the representation capabilities of features to be more expressive before heading towards the final stage of classification. Then, global average pooling is applied, where each feature channel of 7×7 is averaged to one value, resulting in a vector along 1280. This mechanism is very important because it is able to significantly reduce the dimensions of the data while maintaining the semantic meaning of the extracted features.

The final stage is the fully connected layer, which connects 1024 hidden units with 31 target classes. It is in this part that the classification process is carried out, so that the system can determine which class best matches the input image based on the patterns that have been studied throughout the network.

To support the training and development process of the MobileNetV2 architecture in this study, a number of key parameters were used that play a role in regulating model performance. These parameters include optimization algorithms, learning rate values, number of epochs, batch sizes, loss functions, and evaluation metrics [33]. Full details of the parameters used can be seen in Table 2.

Table 2. Model Parameter

Parameter	Description
Optimasi	Adam
Learning Rate	$1e-3$ (stage1) dan $1e-5$ (stage2)
Epoch	20
Batch Size	32
Loss Function	Categorical Crossentropy
Metrics	Accuracy

2.5.2 Model Evaluation

The final stage involves the process of evaluation and application of the model. Evaluation is carried out by measuring model performance using accuracy, precision, recall, and F1-Score metrics[34], so that the quality of the prediction can be assessed comprehensively.

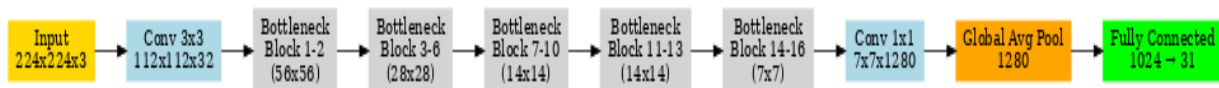


Figure 11. MobilenetV2 Architecture Mode

Once the results of the adequate evaluation are obtained, the model is then integrated into the mobile application for use in facial authentication on the online exam system. Thus, the results of this study can make a real contribution to improving the security and reliability of digital-based exam systems.

3. Results and Discussions

3.1 Model Performance

Figure 12 illustrates the performance of the MobileNetV2 model enhanced with median filtering. Subfigure (a) displays the progression of training and validation accuracy, while subfigure (b) shows training and validation loss across 20 epochs. These visualizations highlight how well the model learns from training data and sustains generalization on validation data.

Based on Figure 12, it can be seen that the accuracy of training and validation increases consistently as the number of epochs increases. At the beginning of the

epoch, accuracy was still relatively low, but it increased sharply to over 90% by the 5th epoch, and close to 95% by the 20th epoch. This shows that the model is able to learn data patterns well and the generalizations on the validation data are also quite stable.

On the loss chart, it can be seen that the value of training and validation losses has decreased significantly from the beginning of training. The initial loss value was in the range of >2.0 for training and >1.0 for validation, but gradually decreased to close to 0.05–0.1 in the last epoch. This decrease indicates that the model is getting better at minimizing prediction errors.

The small difference between the training and validation curves on both graphs shows that overfitting is relatively low [35], as validation accuracy remains high and validation losses remain stable, even lower than training losses in some epoches. This indicates that the application of median filtering as preprocessing has succeeded in improving the quality of image input [35], so that the MobileNetV2 model can perform feature

extraction more effectively and produce good facial authentication performance on online exam systems.

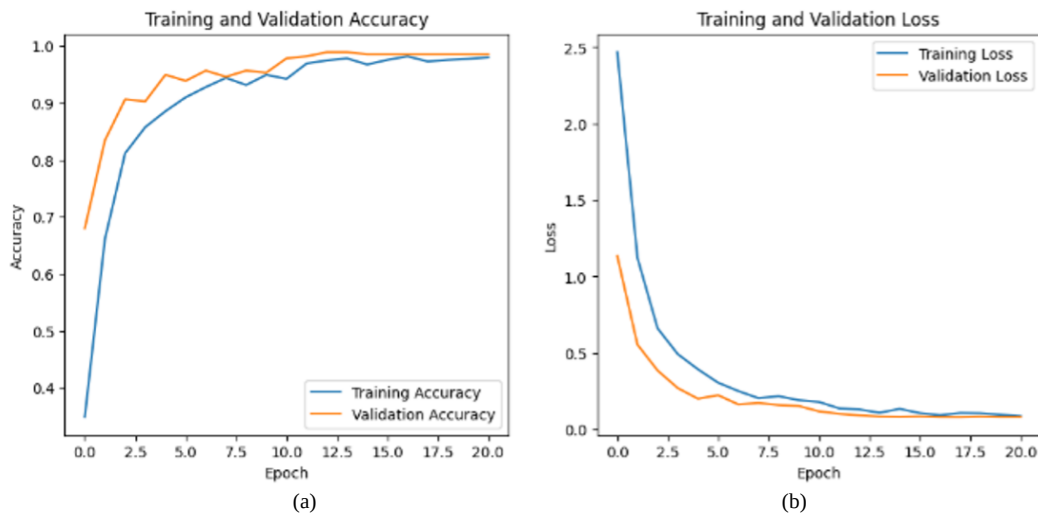


Figure 12. Model Performance

3.2 Model Evaluation Results

Table 3 illustrates the results of the performance evaluation of the MobileNetV2 model with median filtering preprocessing on the face authentication task in the online exam system. Evaluation was carried out using *precision*, *recall*, *F1-score*, and *support* metrics for each class (individual). Table 3 aims to illustrate the extent to which the model is able to recognize each individual's face accurately and consistently.

Table 3. Model Evaluation

No	Name of Student	Precision	Recall	F1-Score	Support
1	Abimanyu	1.00	1.00	1.00	8
2	Afan	1.00	1.00	1.00	8
3	Arya	1.00	1.00	1.00	10
4	Bagus Mustofa	1.00	1.00	1.00	8
5	Dimas	1.00	1.00	1.00	7
6	Dimas Eka	1.00	1.00	1.00	10
7	Fairuzzahran Bachtiar	1.00	1.00	1.00	8
8	Fajar	1.00	1.00	1.00	10
9	Fathur	1.00	1.00	1.00	10
10	Haris Burhanudin	1.00	1.00	1.00	7
11	Ilham Tri Cahya	1.00	1.00	1.00	10
12	Iman	1.00	1.00	1.00	10
13	Indrawan	1.00	1.00	1.00	10
14	Maulana Sujarwadi	1.00	1.00	1.00	8
15	Muhammad Avin Maulana	1.00	0.90	0.95	10
16	Muhammad Ibnu	1.00	1.00	1.00	8
17	Nanda Putri	1.00	1.00	1.00	10
18	Nurasifa	1.00	0.80	0.89	10
19	Poetra Ebeline	1.00	1.00	1.00	10
20	Putri Nabila	1.00	1.00	1.00	10
21	Rafi Dwi Candra Murti	1.00	1.00	1.00	10
22	Rika	1.00	1.00	1.00	9
23	Riyan Setiawan	1.00	1.00	1.00	6
24	Sabrina Khoirunnisa	1.00	0.90	0.95	10
25	Salman Al farizi	1.00	1.00	1.00	9
26	Septia Dewi	1.00	1.00	1.00	8
27	Sindi Amelia	0.71	1.00	0.83	10
28	Siti Latifah	1.00	1.00	1.00	10
29	Siti Nikmat	1.00	1.00	1.00	9
30	Wandi	1.00	1.00	1.00	5
31	Yuslia Devitri	1.00	1.00	1.00	10

Based on the results shown in Table 3, it can be observed that in general the model provides very high performance. Almost all individuals obtained precision, recall, and F1-score scores of 1.00, indicating that the model was able to classify faces with perfect accuracy in most classes. This indicates that the combination of the MobileNetV2 architecture with median filtering as the pre-processing stage successfully improves the quality of feature extraction, so that the system is able to perform optimal facial authentication.

However, there are some exceptions to look out for. Seperti pada kelas Muhammad Avin Maulana dan Sabrina Khoirunnisa, nilai recall hanya mencapai 0.90, sehingga F1-score menurun menjadi 0.95. This condition indicates the presence of false negatives, i.e. the model fails to recognize a small part of the image of the individual. In the Nurasifa class, the recall value even only reached 0.80 with an F1-score of 0.89, which indicates that there are still a number of prediction errors in the validation data. Meanwhile, in the Sindi Amelia class, although the recall value is 1.00, the precision value drops to 0.71, so the F1-score is only 0.83. This indicates a false positive case where the model mistakenly recognizes the face of another individual as Sindi Amelia.

The results of this evaluation show that the model has an excellent generalization rate, with only a few individuals showing a slight decline in performance on certain metrics. These findings also show that although median filtering is successful in reducing image noise, in some classes there are still challenges in the form of visual similarities between individuals or variations in facial expressions that can affect recognition accuracy. Thus, further research can be directed at strengthening data augmentation strategies or applying additional regulatory mechanisms to the model to improve performance consistency across all classes.

3.3 Model Implementation for Face Authentication

To allow the model to be used by the system, a backend structure is built that functions to read and load the model from the storage file. At this stage, the generated model files (*label_final_encoder.pkl* and *mobilenetv2_final_model.h5*) are placed in a dedicated directory and called when the server is first run.

This calling process ensures that the model is ready to use when the system is up without the need for retraining. Once the model is successfully loaded, initial testing is carried out using the Postman application, which is a commonly used tool to verify server and API endpoint functionality. In the test stage, the endpoint in charge of processing faces through a preloaded model is called to ensure system performance. If the server returns a corresponding response, such as face label information or a successful message, then the model load is declared successful and the server is ready to receive requests from the client application.

Once the model is ready for use, the next stage is the development of a mechanism so that facial data can be sent by users to the server for processing through the recognition system. This process begins when a facial image is sent from a client application, i.e. a mobile device-based application, to the backend via a pre-provisioned URL endpoint, such as */auth/recognize* or */api/detect-face*.

When the image is received, the system performs a pre-processing stage consisting of several main steps. First, the image data is read and converted into a numerical array format (pixel matrix) so that it can be processed by a computer. This is done because computers are unable to understand images visually without numerical representations. Second, the area of the face in the image is detected by utilizing the Haar Cascade method to determine the location of the face. Third, the detected part of the face is cut from the original image and adjusted in size to 160×160 pixels, according to the needs of the model. This pre-processing stage is very important because the model can only work optimally when it receives input of a consistent size and format.

Once the facial image has been successfully processed, the cropping and resizing data is fed into the FaceNet model to be converted into embedding, which is a unique numerical representation that reflects the characteristics of the face. Furthermore, embedding is given to the classifier to predict the user's identity. If the confidence score of the predicted result exceeds the predetermined threshold, then the face is recognized and the authentication process is successful.

Conversely, if the confidence value is below the threshold, the system will respond that the face is not recognized so that the authentication process is declared to have failed.

Each time the system performs the facial recognition process, the prediction results are not only sent back to the user, but also recorded in the form of logs for further

documentation and analysis purposes. The logs contain a variety of important information, including the name or label of the face that the system successfully predicted, the time of the recognition process, the accuracy value of the prediction, the validation status that indicates the success or failure of the process, and the face embedding generated by the model.

All of this information is stored in a structured format in two file formats, namely CSV and JSON, so that it can support flexibility in data processing and integration with other systems.

Figure 13, shows the HomeScreen Application which is designed so that the user's focus is directly on the choice of the main feature, namely the authentication process through taking facial images using the camera and access to all logs that have been stored. At the top of the screen, the STMIK IKMI Cirebon logo and the application title are displayed, which serves as a visual identity as well as giving users an initial impression of the purpose and context of the application.

Under the identity section, several interactive menu buttons are provided that are the main access to the app's features. The Camera Start button (in blue) serves to direct the user to the camera features so that the face capture process can be done as the initial stage of authentication. Meanwhile, the View All Reports button (in green) redirects users to the ReportScreen page, which displays a list of authentication results logs with pre-saved success statuses. With the design of the interface in Figure 13, the flow of application usage becomes more intuitive, while ensuring that users can easily access the main functions available.



Figure 13. Homescreen Applications



Figure 14. Screen Authentication

The shooting workflow starts when the user presses the "Start Camera" button on the app's home screen. Once the button is activated, the system directs the device to perform the camera function. The camera is then used to capture facial imagery in a supported format, such as JPEG, so that the capture can be further processed at the pre-processing and facial recognition stages.

Figure 14 is Screen Authentication where this facility is a core component of the application that runs the facial authentication process. The main function of this page is to capture the user's facial image directly through the device's camera and send it to the server for the face recognition process. The success of authentication at this stage is greatly influenced by the quality of the images captured, the stability of the network connection, and the performance of the facial recognition model running on the server.

When the page is opened, the device's camera is automatically activated and a preview of the facial image is displayed on the screen. Users can adjust the position of the face in real-time to be in the center of the frame, ensure adequate lighting, and choose the appropriate shooting angle. This process is also equipped with a manual cropping feature that allows users to make adjustments to the area of the face to be taken is shown in Figure 15.

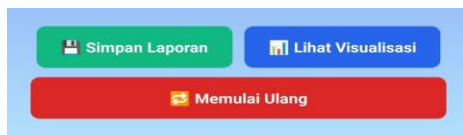


Figure 15. Control Menu Interface

The workflow starts when the user presses the "Start Camera" button on the HomeScreen page. Once the command is executed, the device's camera is activated to capture the user's facial image. The image obtained is then displayed on the screen so that the user as presented in Figure 16, so that they can make adjustments, namely cropping or adjusting the position of the face. This stage aims to ensure that the face is in the appropriate position before the data is sent to the system for further processing.



Figure 16. Face Cropping Process by Users on Authentication Systems



Figure 17. Face Detection Process at the Authentication Stage

Figure 17 shows the face authentication stage, where the system is performing a face detection process on the images that have been captured from the device's camera. At this stage, the app displays a status of

"Detecting faces..." as an indicator that facial images are being processed. Once the face is detected, the image is then packaged in FormData format and sent to the server via the API endpoint /auth/recognize using the POST method. This submission process is carried out by the recognizeFace() function found in the api.js file, which is in charge of ensuring that the data is sent in accordance with the backend standards so that it can be further processed for user identity verification.

Figure 18 and Figure 19 present an overview of the processing stages, where the response begins when the server receives the facial image that has been transmitted from the application interface. At this initial stage, the system ensures that the image input is valid and properly formatted to be processed further. Next, the server executes a facial recognition model that has been trained to perform the identification process by comparing the input image with the existing database of registered users. Through this stage, the model produces predictive outputs that indicate the probability of a match, reflecting both the accuracy and reliability of the system. Based on these processing results, the system generates an immediate response in the form of a pop-up notification that clearly shows the authentication status, whether the process has been successful or has failed. The notification is also designed to display additional information such as the corresponding username and the predictive accuracy level achieved by the model, thereby increasing transparency for users. The results of this authentication are displayed directly on the device's screen so that users can instantly verify their identity status without delay. Furthermore, to strengthen accountability, all outcomes of the authentication process are automatically recorded and stored in system logs, creating a digital archive that can later be used for monitoring activities, security audits, as well as evaluation and system improvement needs in the future.

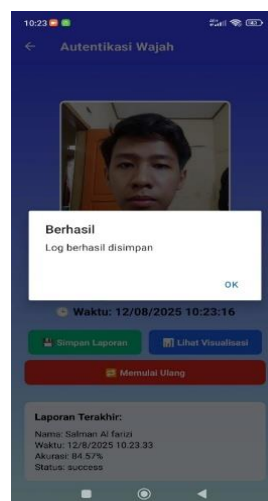


Figure 18. Face Authentication Process Success Notification

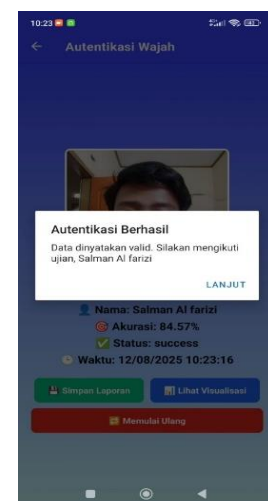


Figure 19. Valid Face Authentication Results

The Evaluation Screen page is an important part of the application that functions to display the results of the

performance evaluation of the facial recognition model as presented in Figure 20. On this page, users can get an objective picture of how well the model is performing based on the evaluation data sent from the backend. When a page is accessed, the system automatically sends a request to the server via an evaluation-specific API endpoint . Next, the backend processes the request and returns data in the form of the results of the visualization of the model performance that has been generated as show in Figure 21.

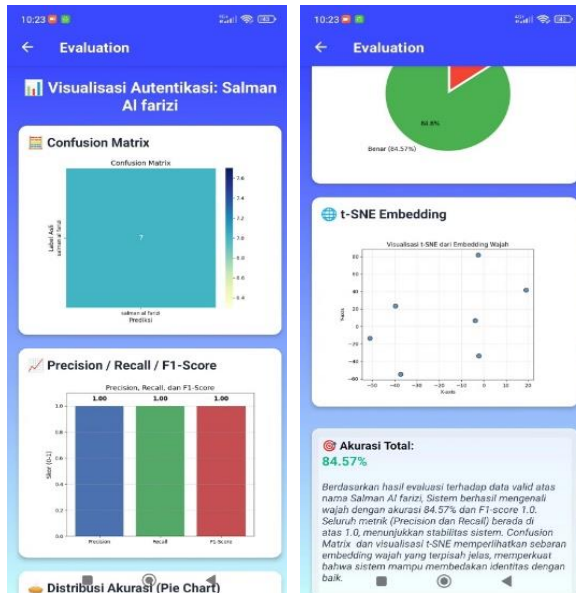


Figure 20. Visualization of Face Authentication Evaluation Results

Figure 21. Visualization of Accuracy Distribution and t-SNE Embedding on Face Authentication Evaluation

Some of the visualization forms displayed include the confusion matrix, which shows the number of correct predictions for each recognized face class, making it easier to identify the level of accuracy per class. Furthermore, precision, recall, and F1-score metrics are also displayed to provide an understanding of the accuracy, completeness, and balance between the two in the predictions made by the system. Another visualization in the form of pie charts is used to show the distribution of the proportion of data or prediction results in the form of a circle, both correctly recognized and incorrectly recognized. In addition, t-SNE plots are also presented to project the representation of facial data into a two-dimensional field. Through t-SNE, the similarities between faces can be visualized intuitively, where adjacent points show a high degree of similarity, while distant points represent significant differences according to the model.

3.3 Discussion

The first problem that is important to discuss is the lack of consistency of approach that combines image preprocessing techniques with lightweight models for facial authentication in online exams. Previous studies have confirmed the effectiveness of median filtering in reducing noise without obscuring important details of the face [17]. However, its application in the context of

online exams is still rarely reported systematically. At the same time, MobileNetV2 proves to be efficient and suitable for devices with computing limitations [20], [22]. This gap shows that the integration of the two approaches has not been comprehensively studied, so this study contributes by testing the effect of median filtering on MobileNetV2 performance directly in an online exam scenario.

The second problem is related to the limited generalization power of facial authentication systems in real conditions. The literature shows that the variability of lighting, the angle of the face capture, as well as the quality of the camera on the end device greatly affect the accuracy of recognition [5], [10]. However, most previous studies have focused on controlled environments with standardized datasets, so they have not fully represented the varied online exam situation [36]. This research tries to fill this gap by conducting experiments using authentic facial data taken directly from students, as well as strengthening the dataset with augmentation techniques so that the system is more resistant to changing conditions.

The third problem is the security and ethical aspects of using facial authentication in online exams. Previous studies have highlighted the risk of spoofing via photos or videos if the system is not equipped with a liveness detection mechanism [15], [37], and raise ethical debates about student privacy in online exam proctoring [6]. However, the integration between technical solutions in the form of liveness detection and privacy ethical considerations is still rarely discussed within the same framework [36]. This study tries to close the gap by designing an authentication system that not only emphasizes technical accuracy, but also pays attention to biometric security aspects and ethical issues to make it more acceptable in higher education practice [38].

4. Conclusions

Based on the results of the study that has been conducted, it can be concluded that the study successfully designed and evaluated a facial authentication system for online exams by integrating median filtering as the pre-processing stage and the MobileNetV2 architecture as the core of the classification model. The results of the experiments showed that this combination was able to increase authentication accuracy by close to 95% with high generalization stability. Median filtering has proven to be effective in reducing noise while maintaining facial details, making it easier for MobileNetV2 to extract relevant features. The application of data augmentation also supports the model's resistance to variations in exposure, viewing angles, and image quality.

Although the overall performance of the system is very good, there are still some cases of individual classes that experience a decrease in precision or recall values due to similarities between faces or variations in expressions. This condition indicates the potential for

false positives and false negatives that need to be overcome in follow-up research. In addition, this study has not in-depth discussed the integration of liveness detection mechanisms to deal with the threat of spoofing, and has not tested user acceptance regarding ethical and privacy aspects in online exam supervision.

Thus, this research contributes to providing efficient, lightweight, and accurate technical solutions to support online exam security. However, further development is needed, among others through strengthening anti-spoofing mechanisms, exploring more diverse data augmentation techniques, and assessing the ethical dimensions and user acceptance. These steps are expected to be able to produce a facial authentication system that is not only technically reliable, but also adaptive, secure, and widely accepted in digital-based higher education practices.

Acknowledgements

The author would like to thank the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia, through the Directorate General of Research and Development, for the funds provided within the framework of the 2025 Basic Fundamental Research Fund (PFD) Program. Special thanks were also conveyed to STMIK IKMI Cirebon for the support, assistance, and facilities provided throughout the research process.

References

- [1] M. Hammad, Y. Liu, and K. Wang, "Multimodal Biometric Authentication Systems Using Convolution Neural Network Based on Different Level Fusion of ECG and Fingerprint," *Ieee Access*, vol. 7, pp. 26527–26542, 2019, doi: 10.1109/access.2018.2886573.
- [2] A. F. Teixeira and M. Jos, "How Higher Education Institutions Are Driving to Digital Transformation: A Case Study," *Educ. Sci.*, vol. 11, no. 10, pp. 1–14, 2021.
- [3] D. Saha, S. Das, and D. Acharjee, "Exploring the Efficacy and Implementation of an Online Examination System : A Comprehensive Study on eXamPro in Modern Education," no. December, 2023.
- [4] S. H. Himel, M. A. Iqbal, M. I. Rana, and T. Mozumder, "Face-App: A Real-Time Face Recognition E-Attendance System for Digital Learning," *Glob. J. Eng. Technol. Adv.*, vol. 11, no. 1, pp. 13–24, 2022, doi: 10.30574/gjeta.2022.11.1.0049.
- [5] M. Rukhiran, S. Wong-In, and P. Netinant, "IoT-Based Biometric Recognition Systems in Education for Identity Verification Services: Quality Assessment Approach," *Ieee Access*, vol. 11, pp. 22767–22787, 2023, doi: 10.1109/access.2023.3253024.
- [6] S. Coghlan, T. Miller, and J. M. Paterson, "Good Proctor or 'Big Brother'? Ethics of Online Exam Supervision Technologies," *Philos. Technol.*, vol. 34, no. 4, pp. 1581–1606, 2021, doi: 10.1007/s13347-021-00476-1.
- [7] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: A review," *Electron.*, vol. 9, no. 8, pp. 1–53, 2020, doi: 10.3390/electronics9081188.
- [8] S. Xu, K. I. Khan, and M. F. Shahzad, "Examining the Influence of Technological Self-Efficacy, Perceived Trust, Security, and Electronic Word of Mouth on ICT Usage in the Education Sector," *Sci. Rep.*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-66689-4.
- [9] S. Raje and S. E. Stitzel, "Strategies for Effective Assessments While Ensuring Academic Integrity in General Chemistry Courses During COVID-19," *J. Chem. Educ.*, vol. 97, no. 9, pp. 3436–3440, 2020, doi: 10.1021/acs.jchemed.0c00797.
- [10] A. B. S. Salamh and H. İ. Akyüz, "New Deep Learning Model for Face Recognition and Registration in Distance Learning," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 12, pp. 29–41, 2022, doi: 10.3991/ijet.v17i12.30377.
- [11] F. Kharbat and A. A. Daabes, "E-Proctored Exams During the COVID-19 Pandemic: A Close Understanding," *Educ. Inf. Technol.*, vol. 26, no. 6, pp. 6589–6605, 2021, doi: 10.1007/s10639-021-10458-7.
- [12] N. E. L. Fadel, "Facial Recognition Algorithms : A Systematic Literature Review," 2025.
- [13] A. Ullah, H. Xiao, and T. Barker, "A Dynamic Profile Questions Approach to Mitigate Impersonation in Online Examinations," *J. Grid Comput.*, vol. 17, no. 2, pp. 209–223, 2018, doi: 10.1007/s10723-018-9442-6.
- [14] M. Garlinska, M. Osial, and K. Proniewska, "The Influence of Emerging Technologies on Distance Education," 2023.
- [15] S. Liu, Y. Song, M. Zhang, J. Zhao, S. Yang, and K. Hou, "An Identity Authentication Method Combining Liveness Detection and Face Recognition," *Sensors*, vol. 19, no. 21, p. 4733, 2019, doi: 10.3390/s19214733.
- [16] R. Mishra, N. Mittal, and S. K. Khatri, "Digital Image Restoration using Image Filtering Techniques," *2019 Int. Conf. Autom. Comput. Technol. Manag. ICACTM 2019*, pp. 268–272, 2019, doi: 10.1109/ICACTM.2019.8776813.
- [17] D. Mishra, N. Roy, M. Bapat, and A. Gudipalli, "Performance Evaluation of Brain Tumor Detection Using Watershed Segmentation and Thresholding," *Int. J. Smart Sens. Intell. Syst.*, vol. 14, no. 1, pp. 1–12, 2021, doi: 10.21307/ijssis-2021-020.
- [18] B. r. c. Joseph, I. J. Jebadurai, G. J. L. Paulraj, J. Jebadurai, and M. M. VARUVEL, "Deep Vein Net: Deep Vein Thrombosis Identification via Sooty Tern Optimized Deep Learning Network," *RRST-Ee*, vol. 69, no. 1, pp. 115–120, 2024, doi: 10.59277/rrst-ee.2024.1.20.
- [19] Y. Huang, W. Dai, and S. He, "Strategies for Marine Remote Sensing Target Detection Based on Deep Learning Model and Internal Wave Simulation Imaging Technology," *Appl. Math. Nonlinear Sci.*, vol. 9, no. 1, 2023, doi: 10.2478/amns.2023.1.00125.
- [20] S. Dou, L. Wang, D. Fan, L. Miao, J. Yan, and H. He, "Classification of Citrus Huanglongbing Degree Based on CBAM-MobileNetV2 and Transfer Learning," *Sensors*, vol. 23, no. 12, pp. 1–18, 2023, doi: 10.3390/s23125587.
- [21] B. Liu, D. Zou, L. Feng, S. Feng, P. Fu, and J. Li, "An FPGA-based CNN accelerator integrating depthwise separable convolution," *Electron.*, vol. 8, no. 3, 2019, doi: 10.3390/electronics8030281.
- [22] M. Ye and J. Zhang, "Mobip: a lightweight model for driving perception using MobileNet," *Front. Neurobot.*, vol. 17, 2023, doi: 10.3389/fnbot.2023.1291875.
- [23] K. Guo et al., "Efficient Adaptive Incremental Learning for Fruit and Vegetable Classification," *Agronomy*, vol. 14, no. 6, 2024, doi: 10.3390/agronomy14061275.
- [24] D. A. Kurnia, S. Risnanto, Sugiyatno, A. N. Kusuma, Y. Naufal, and O. Bin Mohd, "Enhance Low-Resolution Face Recognition Accuracy Using Median and Mean Filtering Techniques Based on Convolutional Neural Networks (CNN)," in *2024 18th International Conference on Telecommunication Systems, Services, and Applications (TSSA)*, IEEE, Oct. 2024, pp. 1–7. doi: 10.1109/TSSA63730.2024.10864112.
- [25] G. M. Kosgiker, A. Deshpande, and K. Anjum, "An Efficient Network-based Skin Lesion Segmentation in Dermoscopic Images," *Int. J. Imaging Syst. Technol.*, vol. 31, no. 2, pp. 874–894, 2021, doi: 10.1002/ima.22545.
- [26] A. Ghanad, "An Overview of Quantitative Research Methods," *Int. J. Multidiscip. Res. Anal.*, vol. 06, no. 08, pp. 3794–3803, 2023, doi: 10.47191/ijmra/v6-i8-52.
- [27] J. Siswanto, J. Rahmadiarto, and M. F. Naufal, "Facial Expression Recognition to Detect Student Engagement in Online Lectures," *Teknika*, vol. 13, no. 2, pp. 226–232, 2024, doi: 10.34148/teknika.v13i2.853.
- [28] M. H. Lee and S. Y. Shin, "Modified Pixels Based Fast

- Median Filter in Impulse Noise Environments,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 9, no. 3, p. 731, 2018, doi: 10.11591/ijeecs.v9.i3.pp731-741.
- [29] Q. H. Nguyen *et al.*, “Influence of data splitting on performance of machine learning models in prediction of shear strength of soil,” *Math. Probl. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/4832864.
- [30] E. Lopez, J. Etxebarria-Elezgarai, J. M. Amigo, and A. Seifert, “The importance of choosing a proper validation strategy in predictive models. A tutorial with real examples,” *Anal. Chim. Acta*, vol. 1275, no. February, p. 341532, 2023, doi: 10.1016/j.aca.2023.341532.
- [31] A. Mumuni and F. Mumuni, “Data augmentation: A comprehensive survey of modern approaches,” *Array*, vol. 16, no. August, p. 100258, 2022, doi: 10.1016/j.array.2022.100258.
- [32] M. J. Adamu *et al.*, “Efficient and Accurate Brain Tumor Classification Using Hybrid MobileNetV2–Support Vector Machine for Magnetic Resonance Imaging Diagnostics in Neoplasms,” *Brain Sci.*, vol. 14, no. 12, 2024, doi: 10.3390/brainsci14121178.
- [33] S. Ahmad, “An Efficient Optimization Technique for Training Deep Neural Networks,” 2023.
- [34] J. H. Cabot and E. G. Ross, “Evaluating prediction model performance,” *Surgery*, vol. 174, no. 3, pp. 723–726, 2023, doi: 10.1016/j.surg.2023.05.023.
- [35] A. Rayeni and H. Naderi, “Predicting the Canadian Yield Curve Using Machine Learning Techniques,” pp. 1–30, 2025.
- [36] E. E. B. Adam and S. Sathesh, “Evaluation of Fingerprint Liveness Detection by Machine Learning Approach - A Systematic View,” *J. Ismac*, vol. 3, no. 1, pp. 16–30, 2021, doi: 10.36548/jismac.2021.1.002.
- [37] Z. Akhtar and G. L. Foresti, “Face Spoof Attack Recognition Using Discriminative Image Patches,” *J. Electr. Comput. Eng.*, vol. 2016, pp. 1–14, 2016, doi: 10.1155/2016/4721849.
- [38] A. M. Elkhataat, “Practical Randomly Selected Question Exam Design to Address Replicated and Sequential Questions in Online Examinations,” *Int. J. Educ. Integr.*, vol. 18, no. 1, 2022, doi: 10.1007/s40979-022-00103-2.