Jurnal Teknologi Informasi Dan Terapan (J-TIT) Vol. 11 No. 2 Tahun 2024 ISSN: 2580-2291

Received November 15th,2024; accepted December 19th, 2024. Date of publication December 30th, 2024
Digital Object Identifier: https://doi.org/10/25047/jtit.v11i2.5643

Improving Online Exam Verification with Class-Weighted and Augmented CNN Models

ILHAM FANANI[1], RIANTO RIANTO[2]

[1]Universitas Teknologi Yogyakarta, Siliwangi, St., Indonesia
[2] Universitas Teknologi Yogyakarta, Siliwangi, St., Indonesia

CORESPONDING AUTHOR: RIANTO RIANTO (email:rianto@staff.uty.ac.id)

ABSTRACT The COVID-19 pandemic has shifted interactions to virtual platforms, significantly impacting education, particularly online exams. However, these online exams have vulnerabilities, including exam jockeys. This study proposes a face classification model using a Convolutional Neural Network (CNN) to verify online exam takers. The model uses preprocessing techniques, i.e. normalization, data augmentation, and class weighting, to balance data and enhance generalization utilizing TensorFlow. The results show an overall accuracy of 85%, with a precision of 86.34%, a recall of 84.24%, an F1-score of 85.28% for legal takers, and a precision of 83.65%, recall of 85.81%, and an F1-score of 84.71% for illegal takers. These results indicate the model's balanced performance between legal and illegal classes. By integrating CNN with tailored preprocessing and training strategies, this study addresses gaps in existing authentication methods, offering a robust approach to online exam verification. The proposed model shows a chance for practical applications. However, further optimization through larger datasets and advanced augmentation techniques is recommended to improve its accuracy and adaptability to diverse real-world contexts.

KEYWORDS: classification, CNN, exam jockeys, online exam

I.INTRODUCTION

The COVID-19 pandemic has drastically changed various aspects of daily life, including how individuals work and think. Before the pandemic, many jobs were done face-to-face, but with social restrictions, most activities have shifted to virtual platforms [1] [2]. This change has affected the economic and social sectors and how people carry out their daily activities, such as shopping, making payments, consulting, and taking exams [3]. This shifting phenomenon requires adapting digital technology to meet the needs of activities previously carried out in person [4].

One of the activities that attracts attention is online exams. Before the pandemic, exams were conducted in classrooms or special places with direct supervision, but now exams are online [5]. Although it offers flexibility for participants, online exams also bring new challenges, namely the validity and legitimacy of exam participants [6]. One of the problems that arises is the existence of illegal practices known as exam jockeys. Exam jockeys are tasked with helping exam participants work on exam questions so they can threaten the integrity of the exam [7].

Exam jockeys are increasingly worrying because they can damage the fairness and credibility of online exam results. Unauthorized exam participants or those who receive third-party assistance to answer questions can disrupt the academic and professional evaluation system. Therefore, effective measures are needed to prevent this practice and ensure that every participant who takes an online exam is legitimate.

A face recognition method is used to verify the legitimacy of online exam takers. This approach is expected to distinguish between legal and illegal online exam takers using facial data [8]. The participation of illegal exam participants can be minimized using the classification model to ensure the integrity of the online exam.

This study proposes advanced techniques, i.e. class weighting and parameter tuning, to optimize the CNN architecture for face recognition. These techniques have practical implications, as class weighting addresses potential data imbalance, ensuring that the model maintains balanced accuracy across legal and illegal participants. Additionally, parameter tuning optimizes batch size, dropout rate, and learning rate to avoid overfitting and underfitting, improving the model's generalization ability. Unlike previous studies that primarily rely on unbalanced datasets or generic CNN models, this research emphasizes robust training strategies to enhance classification performance.

The approach used to solve the problem of exam jockeys is to apply the Convolutional Neural Network (CNN) architecture to classify facial images [9] [10]. CNN has been proven effective in various face recognition tasks and can extract essential features from facial images [11]. The classifier model is expected to identify one individual's face from another so that the exam takers can be verified accurately.

This study uses secondary data [12] to build a model to detect legal and illegal exam takers. The Convolutional Neural Network (CNN) architecture is applied with parameter tuning [13], especially on image size, batch size, drop out [14] and learning rate [15], to produce a model with optimal accuracy and avoid overfitting and underfitting [16].

This study contributes to the face recognition method and provides broad implications for improving the quality and integrity of education or professional certification. This approach can be further developed to cover various contexts and other verification needs in the future so that this technology can provide broader benefits to society.

Integrating class weighting and parameter tuning represents a novel approach to applying CNNs for face recognition in online exam verification. By addressing the limitations of previous methods, this study offers a more balanced and robust solution for distinguishing legal and illegal exam participants. This contribution enhances the technical reliability of face recognition systems and provides practical implications for maintaining fairness and credibility in remote evaluation systems.

II.METHOD

This This study uses secondary data obtained from the site https://www.kaggle.com/datasets/vishesh1412/celebrity-face-image-dataset/data. This dataset consists of two classes, each containing 800 images, for a total of 1,600 images. The images are categorized into two classes: "legal" and "illegal." The image size is standardized to 224 x 224 pixels to maintain consistency as input for the CNN model, and each image is normalized to a range of [0, 1] to accelerate model convergence during training and mitigate the impact of extreme numerical values.

The use of secondary data was motivated by several practical and ethical considerations. Collecting primary data, such as real online exam participant images, presents significant challenges due to privacy regulations, the need for institutional permissions, and logistical constraints. Secondary data provides a diverse and accessible resource, suitable for proof-of-concept development. The selected dataset offers sufficient variations in facial expressions, lighting conditions, and viewing angles, which enrich the model's ability to identify legal and illegal participants while approximating real-world online exam scenarios.

This variation adds value by simulating actual exam conditions, where non-uniform lighting and varied angles are common. Example images from each class are shown in Fig. 1.

FIGURE 1. Example of data set

Each face image in this data set is a treasure trove of visual attributes, including a variety of facial expressions (neutral, smiling, serious), lighting conditions (bright, dim, mixed), and shooting angles (frontal, oblique, or top-down). These attributes are not just features but the very essence of the data that enhances the model's learning process, thereby improving its generalization ability in online exam scenarios with diverse visual conditions.

The Kaggle data set was selected based on the primary considerations, namely reliability, diversity, data quality, and its wide use as a standard in similar studies. This data set used facial features covering various expressions and lighting conditions as essential elements for building a robust and generalizable classification model. In general, the research stages are shown in Fig. 2.

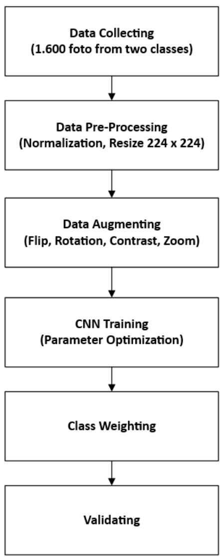

FIGURE 2. The research stages

Details regarding each stage in the research are explained as follows:

1. Data Pre-Processing

The first stage in model training is data pre-processing. At this stage, all images are resized to 224 x 224 pixels to maintain the consistency of the input in the model. In addition, image normalization is also carried out by changing the pixel value from the range [0, 255] to [0, 1]. This normalization process is carried out using the following equation:

$$X_{norm} = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (1)$$

X_{norm} is the normalized pixel value, X is the original pixel value, and min(X) and max(X) are the image's minimum and maximum pixel values , respectively. The normalization process is critical because it can help speed up the convergence of the model during training. It also prevents the model from being disturbed when faced with extreme numerical values [17].

2. Data Augmenting

Data augmentation is done to increase the variety of training data without increasing the available data. This augmentation can be described by a transformation operation applied to the image, for example, a rotation of the image by an angle θ defined by the rotation matrix:

$$\begin{pmatrix} x^1 \\ y^1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (2)$$

The coordinates (x,y) are the original coordinates of the pixel, and (x^1,y^2) are the pixel coordinates after rotation. In addition to rotation, augmentation includes horizontal and vertical flips, zooming, and contrast adjustments. The purpose of augmentation is to improve the model's generalization ability to variations in the position, orientation, and scale of objects in the image so that the model can cope with potentially different situations in the data that it has never seen before [18].

3. CNN Model Training

This study uses a Convolutional Neural Network (CNN) model designed explicitly for two-class classification of facial image data. In the early stages of training, the CNN model is initialized with random weights and trained end-to-end using available data in two classes, namely "legal" and "illegal". The Categorical Cross entropy loss function is used to optimize the model parameters, which are defined as follows:

$$\text{Loss} = -\sum_{i=1}^{n} y_i \log(\hat{y}i) \quad (3)$$

Label y_i is the actual label, $\hat{y}i$ is the probability predicted by the model, and n is the number of classes. During the training process, hyperparameters such as batch size, learning rate, and number of epochs are adjusted so that the model can achieve good accuracy without experiencing overfitting or underfitting problems. This strategy helps the model adapt to the characteristics of the data to improve prediction accuracy on the test data.

4. Class Weighting

The class weighting is applied during training to ensure that the model is not biased towards one class. Class weighting is calculated using the equation:

$$w_c = \frac{N}{n_c \; x \; C} \quad (4)$$

The weight w_c is the weight for class c, N is the total number of samples, n_c is the number of samples in class c, and C is the total number of classes. Even though the data set used is balanced, with the same number of images in each class, class weighting is still done to maintain balance and ensure that the model is not more biased towards one class [19].

5. Model Validation and Evaluating

Following the training process is complete, the model is evaluated using validation data. This evaluation is done by calculating several important metrics such as accuracy, precision, recall, and F1-score. One of the main tools used to analyze model performance is the Confusion Matrix, which provides an overview of each class's correct and incorrect predictions. The formula used to calculate precision, recall, and F1-score is as follows:

$$\text{Akurasi} = \frac{TP + FP}{TP + FP + TN + FN} \quad (5)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (7)$$

$$F1 - \text{Score} = \frac{2 \text{ x Precision x Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

TP is True Positive, TN is True Negative, FP is False Positive, and FN is False Negative. Overall, the methods used in this study are data pre-processing, augmentation, transfer learning, fine-tuning and data analysis using Confusion Matrix [20].

III.RESULT AND DISCUSSION

This study focuses on developing an image classification model using a Convolutional Neural Network (CNN) architecture trained from scratch. The model is expected to classify images into two predetermined classes through training stages: data pre-processing, augmentation, model training, and hyperparameter optimization. In this section, the results of model training and validation will be presented in detail, along with an in-depth discussion of model performance based on evaluation metrics, i.e. accuracy, precision, recall, and F1-score. Analysis using the Confusion Matrix is also carried out to provide a more complete picture of the model's ability to classify and identify areas that can potentially cause prediction errors.

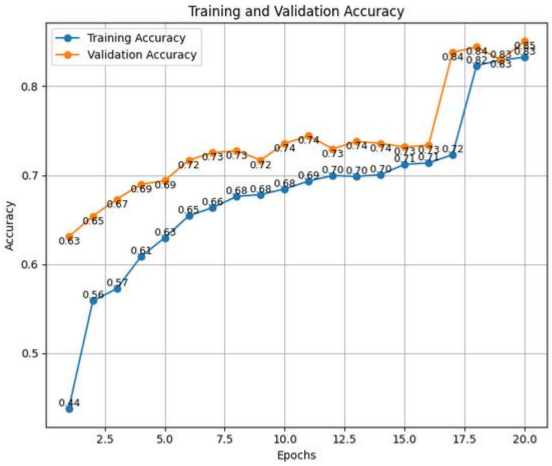

FIGURE 3. The accuracy results

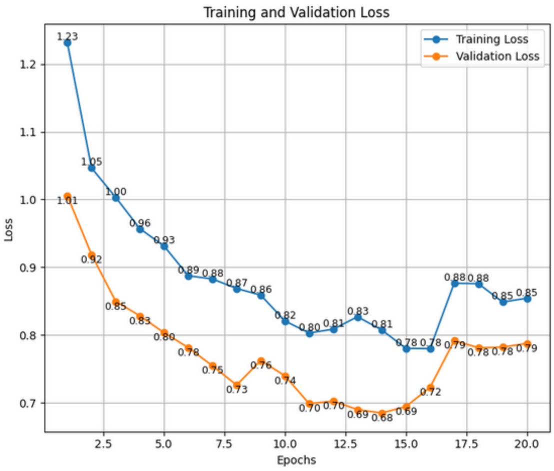

FIGURE 4. The loss results

The graph in Fig. 3 showcases the model's ability to recognize patterns, as evidenced by the development of accuracy, while Fig. 4 shows loss during the training and validation process. The model's training accuracy, which starts at 44 percent in the first epoch, gradually increases. As the number of epochs increases, the model's pattern recognition capabilities become more pronounced, leading to a consistent increase in accuracy to around 83 percent in the 16th epoch. This impressive pattern recognition culminates in a significant spike in accuracy, indicating the model's stability at around 84-85 percent at the end of training.

On the other hand, validation accuracy also demonstrates a steady upward trend since the start of training. Beginning at 63 percent in the first epoch, the validation accuracy gradually rises to 84 percent in the 17th epoch and remains stable until the last epoch. This stability is a clear indication of the model's capacity to learn from the training data and maintain its performance on previously unseen validation data, without any significant signs of overfitting.

In addition to accuracy, the loss graph provides further insight into the model's learning process. At the beginning of training, the training loss was at 1.23 and continued to decrease, eventually reaching around 0.7 in the last epoch. This decrease indicates that the model is improving at minimizing prediction errors on the training data, indicating that the model has succeeded in learning effectively.

Validation loss also shows a similar decreasing pattern, which is a sign of the model's learning process. It starts from 1.01 in the first epoch and decreases steadily to around 0.79. Although there have been slight fluctuations in the last few epochs, validation loss has not shown a significant increase, indicating that the model remains stable and does not experience overfitting to the validation data.

The The accuracy and loss graphs consistently demonstrate the model's performance and its ability to maintain good generalization to new data. These results confirm the reliability of the trained model for classifying images under various conditions, ensuring its secure performance.

Several studies have explored the application of machine learning and deep learning methods in facial recognition. For instance, [21] examined the performance of Support Vector Machines (SVM), Multi-Layer Perceptron (MLP), and Convolutional Neural Networks (CNN) in real-time face recognition tasks. Their findings demonstrated that CNN outperformed traditional methods like SVM and MLP in terms of accuracy and adaptability, particularly in handling diverse conditions.

Similarly, Paul and Acharya [22] conducted a comparative analysis of various facial recognition algorithms, including PCA with Eigenfaces, SVM, KNN, and CNN. Their results highlighted CNN as the most effective algorithm, delivering superior accuracy and robustness, particularly in complex scenarios with variations in lighting, facial angles, and expressions. This study emphasized the limitations of traditional methods like SVM in dealing with real-world variations, further reinforcing the advantages of deep learning architectures.

Building on these works, this study leverages CNN to address the challenges in facial classification, particularly in distinguishing between legal and illegal participants in online exams. Unlike traditional methods, the CNN architecture in this study is designed to handle diverse facial variations and conditions, ensuring higher adaptability and accuracy in real-world scenarios.

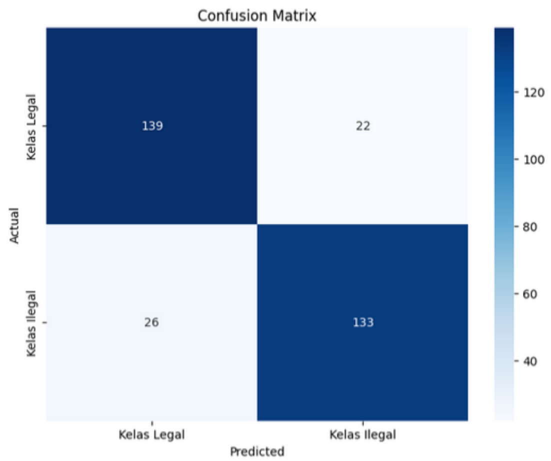

FIGURE 5. The confusion matrix results

The Confusion Matrix in Fig. 5 provides a more detailed picture of the model's performance in classifying images. Out of 161 images that are actually "legal," the model successfully classifies 139 images correctly, while 22 images are incorrectly classified as "illegal." Conversely, out of 159 images that are actually "illegal," the model successfully classifies 133 images correctly, while 26 images are incorrectly classified as "legal." Based on the Confusion Matrix, both classes' precision, recall, and F1-score values are obtained, as shown in Table 1.

TABLE 1. Evaluation results

	Legal (%)	Illegal (%)
Precision	86.34	83.65
Recall	84.24	85.81
F1-Score	85.28	84.71

Table 1 shows the model's performance in classifying images into legal and illegal classes based on the precision, recall, and F1-score values calculated from the Confusion Matrix. The model has a precision value of 86.34 percent in the legal class, meaning that of all predictions classified as legal, 86.34 percent are truly legal class images. This high precision value indicates that the model accurately predicts legal images with relatively low errors. However, some wrong predictions still exist, where images from the illegal class are classified as legal, which is reflected in the False Positives value.

For a Recall on the legal class, the model achieved a value of 84.24 percent. This value shows that the model can detect 84.24 percent of all legal class images. The meaning is the model's ability to recognize the legal class well, but around 15.76 percent of legal images still fail to be recognized and classified as illegal (False Negatives). Although this Recall value is relatively high, the misclassification shows that the model still needs help recognizing all legal classes perfectly.

In the illegal class, the model has a precision of 83.65 percent, meaning that of all the predictions classified as illegal, 83.65 percent are illegal images. Although the Precision in the illegal class is slightly lower than in the legal class, it is still high, indicating that the model is quite good at predicting illegal images with an acceptable error rate.

The Recall for the illegal class is 85.81 percent, indicating that the model can detect 85.81 percent of all illegal images. The higher Recall value compared to Precision in this class indicates that the model is slightly better at recognizing most illegal images. However, this value also indicates that about 14.19 percent of illegal images are incorrectly classified as legal, which is reflected in the number of False Negatives for the illegal class.

The F1-score for the illegal class is 84.71 percent, indicating that the model has a reasonably good balance between Precision and Recall in detecting the illegal class. Although slightly lower than the F1-score for the legal class, this value is still close to the previous two metrics, indicating that the model has stable performance in identifying the illegal class quite well. The small differences between Precision, Recall, and F1-score indicate that the model has a reasonably consistent ability. However, there is room for further improvement, especially in reducing the misclassification between legal and illegal classes.

Overall, the model performs reasonably well in classifying legal and illegal images, with an overall accuracy of 85 percent. The balanced values of Precision, Recall, and F1-score indicate that the model can handle both classes well, although there are some prediction errors. These errors are most likely caused by the similarity of features between the legal and illegal images or by limitations in the preprocessing and data augmentation stages. To improve model performance, further optimization is needed, such as increasing the amount of training data or using more varied augmentation techniques, which can help the model recognize more subtle differences between the two classes.

Class weighting during training was critical in achieving these balanced performance metrics. By addressing potential biases toward the majority class, class weighting ensured that the model treated both classes equitably, even in scenarios where feature overlap or visual similarities between classes could lead to misclassification. This approach highlights the importance of incorporating advanced training techniques in CNN architectures to improve the robustness and fairness of classification systems. The results emphasize that even with balanced datasets, class weighting significantly reduces prediction bias and enhances overall model reliability.

The Confusion Matrix in Fig. 5 shows several classification errors occur, especially in the illegal class. Of the 159 images that should be classified as the illegal class, 26 are incorrectly classified as the legal class. Similarly, in the legal class, out of 161 images, 22 images are incorrectly classified as the illegal class. These errors can be caused by several factors, including [23]:

1. Visual similarity between classes: If there are similar visual features between images from class 1 and class 2, the model may have difficulty distinguishing them, resulting in classification errors.
2. Limited features extracted: Although the ResNet50 model is very good at extracting features, it is possible that some features that are crucial for distinguishing between the two classes are not sufficiently represented in the training process, especially if the amount of training data is relatively small.
3. Suboptimal data augmentation: The data augmentation techniques used may need to be revised to create the variation needed for the model to recognize finer differences between the two classes.

The training parameters used in this model, i.e. batch size, learning rate, number of epochs, and optimizer, significantly influence the model's accuracy and stability. In detail, the contribution of each training parameter is as follows:

1. Batch Size (32)
   In this study, the batch size was set at 32, which provides a balance between training speed and prediction accuracy. This size allows for a faster process without sacrificing generalization performance.
2. Learning Rate (0.0001)
   Using a learning rate of 0.0001, the model achieves stable convergence without the risk of jumping over the optimal point or converging too quickly. This value was achieved after several experiments with higher and lower values for optimal stability and accuracy.
3. Number of Epochs and Early Stopping
   A total of 20 epochs were used in training with the addition of early stopping to stop training if the validation loss did not improve after five consecutive epochs. Setting epochs and early stopping will prevent overfitting and maintain model stability on new data.
4. Adam Optimizer
   Adam optimizer accelerates the convergence process adaptively compared to other optimizers. Adam was chosen because it provided better stability in CNN training than the SGD optimizer in early experiments.

This setting contributes to achieving stable model accuracy on validation data and helps to reduce the risk of overfitting. With the right combination of parameters, the model can maintain good generalization, as shown in the accuracy results and other evaluation metrics. Overall, the results of this study indicate that the built model has great potential for use in image classification applications. By making some improvements and further developments, the performance of this model can be further improved so that it can be more reliable and accurate in classifying images on a larger scale.

This study relies on secondary data due to practical and ethical challenges in collecting primary facial data from actual online exam participants. While secondary data provides sufficient diversity for this proof-of-concept model, it may only partially capture real-world exam conditions. Future studies should prioritize using primary data to tailor the model more closely to actual exam scenarios and improve its real-world applicability.

IV. CONCLUSION

The image classification model developed using the Convolutional Neural Network (CNN) architecture demonstrated good performance, achieving an overall accuracy of 85 percent with balanced Precision, Recall, and F1-score of both legal and illegal classes. These results highlight the model's ability to effectively differentiate between legal and illegal exam takers. Class weighting during training played a significant role in achieving balanced predictions addressing potential bias toward one class. However, some classification errors, particularly in the illegal class, reveal limitations in generalizing to real-world contexts due to reliance on secondary data. Future research can focus on increasing the diversity of training data through primary data collection or advanced augmentation techniques to improve visual representation. Additionally, exploring ensemble learning techniques or fine-tuning pre-trained architectures may enhance model robustness and reduce classification errors, ensuring excellent reliability in practical applications.

ACKNOWLEDGMENT

This study was made possible with funding support from the Directorate of Research, Technology, and Community Service of the Ministry of Education and Culture of the Republic of Indonesia. Gratitude is also extended to www.kaggle.com to provide the data set used in developing the model for this study. Lastly, sincere appreciation is given to University Technology of Yogyakarta for its full support, which greatly contributed to the completion of this research.

REFERENCE

[1] H. Santos, “COVID-19 Lockdown Effects on Student Grades of a University Engineering Course: A Psychometric Study,” *IEEE Transactions on Education*, vol. 65, no. 4, pp. 493–501, Nov. 2022, doi: 10.1109/TE.2021.3131745.

[2] R. Y. Kim, "The Impact of COVID-19 on Consumers: Preparing for Digital Sales," *IEEE Engineering Management Review*, vol. 48, no. 3, pp. 212–218, Jul. 2020, doi: 10.1109/EMR.2020.2990115.

[3] O. Tounekti, A. Ruiz-Martinez, and A. F. Skarmeta Gomez, "Users Supporting Multiple (Mobile) Electronic Payment Systems in Online Purchases: An Empirical Study of Their Payment Transaction Preferences," *IEEE Access*, vol. 8, pp. 735–766, 2020, doi: 10.1109/ACCESS.2019.2961785.

[4] N. Van Zeebroeck, T. Kretschmer, and J. Bughin, "Digital 'is' Strategy: The Role of Digital Technology Adoption in Strategy Renewal," *IEEE Trans Eng Manag*, vol. 70, no. 9, pp. 3183–3197, Sep. 2023, doi: 10.1109/TEM.2021.3079347.

[5] S. Allamsetty, M. V. S. S. Chandra, N. Madugula, and B. Nayak, "Improvement of the Quality of Question Papers for Online Examinations Toward Simultaneous Enhancement of Students' Learning," *IEEE Transactions on Learning Technologies*, vol. 17, pp. 135–142, 2024, doi: 10.1109/TLT.2023.3272361.

[6] A. W. Muzaffar, M. Tahir, M. W. Anwar, Q. Chaudry, S. R. Mir, and Y. Rasheed, "A systematic review of online exams solutions in e-learning: Techniques, tools, and global adoption," *IEEE Access*, vol. 9, pp. 32689–32712, 2021, doi: 10.1109/ACCESS.2021.3060192.

[7] M. Alguacil, N. Herranz-Zarzoso, J. C. Pernías, and G. Sabater-Grande, "Academic dishonesty and monitoring in online exams: a randomized field experiment," *J Comput High Educ*, pp. 1–17, Jul. 2023, doi: 10.1007/S12528-023-09378-X/TABLES/4.

[8] J. Jia and Y. He, "The design, implementation and pilot application of an intelligent online proctoring system for online exams," *Interactive Technology and Smart Education*, vol. 19, no. 1, pp. 112–120, Feb. 2022, doi: 10.1108/ITSE-12-2020-0246/FULL/XML.

[9] H. Ben Fredj, S. Bouguezzi, and C. Souani, "Face recognition in unconstrained environment with CNN," *Visual Computer*, vol. 37, no. 2, pp. 217–226, Feb. 2021, doi: 10.1007/S00371-020-01794-9/METRICS.

[10] A. Chaudhuri, "Deep Learning Models for Face Recognition: A Comparative Analysis," pp. 99–140, 2020, doi: 10.1007/978-3-030-32583-1_6.

[11] X. He and Y. Chen, "Transferring CNN Ensemble for Hyperspectral Image Classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 5, pp. 876–880, May 2021, doi: 10.1109/LGRS.2020.2988494.

[12] N. J. Fleischer and A. Khalil, "Limitations and recommendations for use of secondary data analysis in pediatric research," *Children's Health Care*, Oct. 2024, doi: 10.1080/02739615.2023.2279064.

[13] R. Gonzales-Martinez, J. Machacuay, P. Rotta, and C. Chinguel, "Hyperparameters Tuning of Faster R-CNN Deep Learning Transfer for Persistent Object Detection in Radar Images," *IEEE Latin America Transactions*, vol. 20, no. 4, pp. 677–685, Apr. 2022, doi: 10.1109/TLA.2022.9675474.

[14] C. Garbin, X. Zhu, and O. Marques, "Dropout vs. batch normalization: an empirical study of their impact to deep learning," *Multimed Tools Appl*, vol. 79, no. 19–20, pp. 12777–12815, May 2020, doi: 10.1007/S11042-019-08453-9/METRICS.

[15] S. Noppitak and O. Surinta, "dropCyclic: Snapshot Ensemble Convolutional Neural Network Based on a New Learning Rate Schedule for Land Use Classification," *IEEE Access*, vol. 10, pp. 60725–60737, 2022, doi: 10.1109/ACCESS.2022.3180844.

[16] L. Qian, L. Hu, L. Zhao, T. Wang, and R. Jiang, "Sequence-Dropout Block for Reducing Overfitting Problem in Image Classification," *IEEE Access*, vol. 8, pp. 62830–62840, 2020, doi: 10.1109/ACCESS.2020.2983774.

[17] R. Archana and P. S. E. Jeevaraj, "Deep learning models for digital image processing: a review," *Artif Intell Rev*, vol. 57, no. 1, pp. 1–33, Jan. 2024, doi: 10.1007/S10462-023-10631-Z/TABLES/5.

[18] M. Nagaraju, P. Chawla, and N. Kumar, "Performance improvement of Deep Learning Models using image augmentation techniques," *Multimed Tools Appl*, vol. 81, no. 7, pp. 9177–9200, Mar. 2022, doi: 10.1007/S11042-021-11869-X/TABLES/10.

[19] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J Big Data*, vol. 6, no. 1, pp. 1–54, Dec. 2019, doi: 10.1186/S40537-019-0192-5/TABLES/18.

[20] G. Naidu, T. Zuva, and E. M. Sibanda, "A Review of Evaluation Metrics in Machine Learning Algorithms," *Lecture Notes in Networks and Systems*, vol. 724 LNNS, pp. 15–25, 2023, doi: 10.1007/978-3-031-35314-7_2.

[21] G. He and Y. Jiang, "Real-time Face Recognition using SVM, MLP and CNN," *Proceedings - 2022 International*

Conference on Big Data, Information and Computer Network, BDICN 2022, pp. 762–767, 2022, doi: 10.1109/BDICN55575.2022.00149.

[22] S. Paul and S. K. Acharya, “A Comparative Study on Facial Recognition Algorithms,” *SSRN Electronic Journal*, Dec. 2020, doi: 10.2139/SSRN.3753064.

[23] P. S. S. Sreedhar and N. Nandhagopal, “Classification Similarity Network Model for Image Fusion Using Resnet50 and GoogLeNet,” *Intelligent Automation & Soft Computing*, vol. 31, no. 3, pp. 1331–1344, Oct. 2021, doi: 10.32604/IASC.2022.020918.