JOIV : Int. J. Inform. Visualization, 8(2) - May 2024 991-1000

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Optimizing the Performance of AI Model for Non-Invasive Continuous Glucose Monitoring: Hyperparameter Tuning and Random Oversampling Approach

Karisma Trinanda Putra [a,b,*], Mahendro Prasetyo Kusumo [c], Prayitno [d], Darma Wicaksana [a], Ahmad Zaki Arrayyan [a], Sakca Garda Pratama [a], Mujib Alrahman Al-Kamel [a], Hsing-Chung Chen [e,f]

[a] *Department of Electrical Engineering, Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, Bantul, Yogyakarta, Indonesia*
[b] *Center of Artificial Intelligence and Robotics Studies, Research and Innovation Center, Universitas Muhammadiyah Yogyakarta, Bantul, Yogyakarta, Indonesia*
[c] *Master of Hospital Administration, Postgraduate School, Universitas Muhammadiyah Yogyakarta, Bantul, Yogyakarta, Indonesia*
[d] *Department of Electrical Engineering, Politeknik Negeri Semarang, Tembalang, Semarang, Indonesia*
[e] *Department of Computer Science and Information Engineering, College of Information and Electrical Engineering, Asia University, Taichung City, Taiwan*
[f] *Department of Medical Research, China Medical University Hospital, China Medical University, Taichung City, Taiwan*

*Corresponding author: *karisma@ft.umy.ac.id*

Abstract*—Diabetes Mellitus (DM) as a non-communicable disease (NCD) continues to increase every year. Continuous glucose monitoring (CGM) is essential for effective DM management. However, existing disposable glucose monitoring methods still rely on invasive techniques, cause pain, and lack continuous monitoring capabilities. On the other hand, non-invasive techniques are not feasible for CGM due to the biometric data's complexity and the classification system's inadequate performance. This study aims to develop a non-invasive technology to improve the performance of a non-invasive blood glucose classification system using Artificial Intelligence (AI), specifically Convolutional Neural Network (CNN) and an oversampling technique. The oversampling technique could improve data quantity by balancing the amount of data for each class. This study recruited twenty-three participants in the age range of 20 to 22 years comprising seven females and fifteen males. During data recording sessions, blood glucose levels were simultaneously assessed using a gold-standard glucometer and a non-invasive CGM prototype. The proposed CNN model successfully improved the classification accuracy of non-invasive blood glucose monitoring significantly. With the implementation of oversampling for augmenting the data, the accuracy of the proposed model increased to more than 88%. This study concludes that non-invasive approaches combined with AI technology have the potential to provide a convenient and pain-free alternative to traditional monitoring methods, significantly improving diabetes management and enhancing the overall quality of life for those affected by this condition. These findings could revolutionize the field of diabetes management, offering a more comfortable and accurate monitoring solution that could potentially transform the lives of millions of diabetes patients.

***Keywords*— Diabetes Mellitus; continuous glucose monitoring; non-invasive; AI, CNN.**

Manuscript received 10 Aug. 2023; revised 31 Oct. 2023; accepted 11 Nov. 2023. Date of publication 31 May 2024.

I. INTRODUCTION

The prevalence of NCD, such as DM, continues to surge annually, with Indonesia currently holding the unsettling distinction of being ranked third globally for undiagnosed DM cases. This alarming statistic is based on data from the International Diabetes Federation Diabetes Atlas of 2021, revealing that an astonishing 73.7% of the entire Indonesian population remains undiagnosed for diabetes [1]. Lack of knowledge, attitude, and behavior toward early detection of DM is the cause of the increase in cases, coupled with the difficulty of accessing health facilities for ongoing blood sugar checks [2]. Invasive blood collection techniques, e.g., venous puncture, skin puncture, and arterial puncture, are problematic for individuals who have a phobia of needles, resulting in discomfort, pain, and potential risk of infections [3]. DM is a non-communicable disease whose number continues to increase every year. DM is a chronic condition that occurs when the pancreas cannot produce insulin or the body cannot use the insulin produced properly. This NCD leads to high BGL that has the potential to cause severe complications if not appropriately managed [4]. In healthy conditions, blood sugar levels range from 60 to 140 milligrams per deciliter (mg/dL). In contrast, blood sugar levels in diabetic patients are irregular. Blood sugar levels that exceed normal limits, such as 140 to 500 mg/dL, are known as hyperglycemia, while levels below 60 mg/dL are known as hypoglycemia [5]. Continuous monitoring of blood glucose levels is crucial for patients with DM to uphold their health and enhance their quality of life.

Biometric measurements to determine the patient's condition have been widely implemented with non-invasive methods, e.g., blood pressure, heartbeat, and oxygen saturation levels. However, the development of non-invasive devices for CGM faces significant challenges. One of the primary challenges is ensuring accurate and reliable glucose readings [6]. The CGM sensors should maintain consistency in glucose measurements while overcoming issues such as sensor drift, calibration inaccuracies, and signal noise [7]. Interference from external factors, such as temperature changes or medication intake, can also impact the accuracy of CGM readings [8]. Addressing the challenge of non-independently and Identically Distributed (non-IID) data also poses a barrier to enhancing the performance of decision-making systems. The continuous stream of glucose data requires efficient algorithms and systems to provide meaningful insights for both patients and healthcare providers [9]. Addressing the challenges posed by non-IID datasets in medical data silos has emerged as a critical focus.

The nature of medical data often exhibits significant variations due to diverse sources, patient demographics, and data acquisition protocols, rendering them non-IID [10]. In recent research, the application of AI techniques to non-IID medical data has garnered substantial attention. Data oversampling approaches have been explored to mitigate the impact of non-IID characteristics [11], [12]. Oversampling strategies leverage knowledge gained from one domain to enhance performance in another, addressing the challenge of limited labeled data in non-IID settings. These advancements underscore the importance of developing robust AI models capable of effectively handling the complexities inherent in diverse medical datasets [13], thereby paving the way for improved diagnostic accuracy and personalized healthcare solutions.

This study utilizes a PPG sensor to monitor Blood Glucose Levels (BGLs) using two data sensors, red and infrared signals. It also incorporates AI technology, using CNN and oversampling techniques to classify blood glucose patterns accurately. This approach is anticipated to significantly contribute to overcoming challenges in developing sustainable, noninvasive blood glucose monitoring technology, ultimately enhancing the DM healthcare experience for patients.

The remainder of this paper is structured as follows: Section II details the study materials and the proposed method. Section III elaborates on and discusses the results. Finally, concise conclusions are summarized in Section IV.

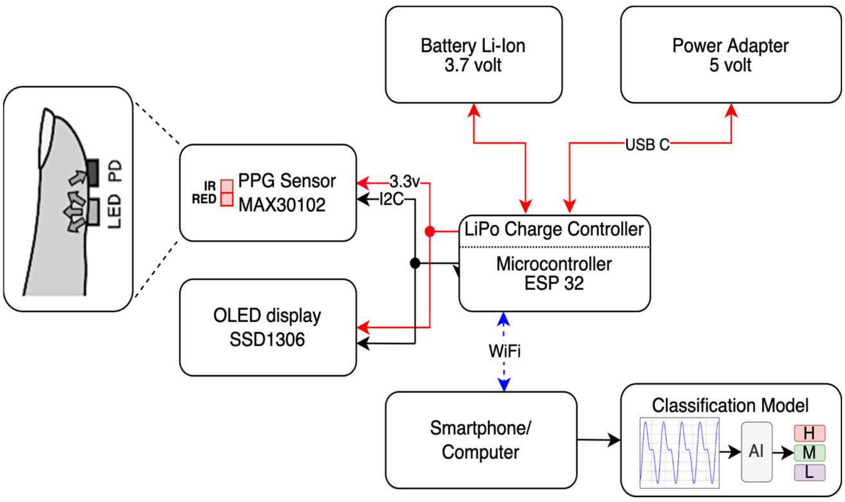

Fig. 1 A non-invasive prototype [14] was used to record PPG signals.

II. Materials and Method

CGMs have emerged as a pivotal technology in diabetes management, offering real-time insights into an individual's glucose levels [15]. DM patients often face challenges in maintaining optimal blood glucose levels, leading to the need for a proactive monitoring system. CGM systems, consisting of small sensors implanted under the skin, continuously measure glucose levels in interstitial fluid, providing a more comprehensive understanding of glucose fluctuations throughout the day. The real-time data generated by CGM enables early detection of abnormal glucose patterns [16], allowing for timely interventions and personalized treatment adjustments. This proactive approach empowers healthcare providers and patients with the tools to prevent hypoglycemic or hyperglycemic events, ultimately enhancing the quality of diabetes care. The integration of CGM with a decision-making system based on AI technology represents a significant stride towards personalized and precision medicine, promising improved health outcomes for individuals living with diabetes [17].

AI, particularly CNN models, has revolutionized medical signal processing [18] and prediction[19], demonstrating substantial promise in healthcare applications [20]. CNNs excel in extracting intricate features from complex data, making them particularly well-suited for analyzing physiological signals. In predictive analytics, AI-driven models can leverage large datasets to identify patterns and correlations within medical signals, contributing to early disease detection and prognosis. In medical signal processing, CNNs have been instrumental in tasks such as image recognition in medical imaging, enabling more accurate and rapid diagnoses. The effective management of medical data in healthcare institutions is particularly challenging, given the non-IID nature of the data recorded by various practitioners and medical institutions.

This research focuses on optimizing an AI model, specifically CNN, to enhance the performance of a non-invasive blood glucose classification system. The classification system, i.e., CGM, utilizes an experimental prototype [14] powered by ESP32 microprocessor and a PPG sensor MAX30102. As depicted in Figure 1, this prototype records PPG signals in the infrared and red spectrum and then transmits them to the smartphone or computer for AI inferencing. By leveraging the capabilities of CNN, this study aims to extract relevant features and patterns from the non-invasive blood glucose data, enabling more accurate and reliable classification of BGLs. The proposed methodology utilizes CNN's ability to learn hierarchical representations from the data, leading to improved classification results. This optimization of the deep CNN holds the potential to revolutionize non-invasive blood glucose monitoring technologies and contribute to more effective healthcare practices.

A. Designing an AI Model from the Scratch

1) Deep Learning (DL): Deep learning is a subset of machine learning that utilizes artificial neural networks to handle large datasets, making it particularly effective in supervised learning. It involves the study of artificial neural networks and related machine learning algorithms with multiple hidden layers, also known as deep structured learning, hierarchical learning, or deep machine learning. These deep nets have demonstrated significant capabilities in tackling complex tasks and patterns within data [21], [22].

2) Convolutional Neural Networks (CNN): CNN is one part of the neural network method. CNN is not much different from ordinary neural networks, which consist of neurons that have weights, biases, and activation functions [23]. CNN is the development of an Artificial Neural Network (ANN) algorithm that uses the same principle as the working system of human neurons. CNN is similar to ANN, consisting of neurons that optimize themselves through learning. Each neuron will still receive input and perform operations (i.e., a scalar product followed by a non-linear function) based on countless ANNs. From the raw image vector input to the final output of class scores, the entire network will still express one perceptive score function (weight). The last layer will contain the loss functions associated with the class, and all the usual tips and tricks developed for ANNs still apply. An essential difference between CNN and ANN is that CNN is mainly used in the field of pattern recognition in images. This allows us to encode image-specific features into the model architecture, making the network more suitable for image-focused tasks, thereby reducing the parameters required to prepare the model [24].

3) Data Oversampling: Oversampling is a method that balances a non-IID dataset by adding new samples to the minority class. There are two main approaches to oversampling: random oversampling and synthetic oversampling. In random oversampling, existing minority samples are duplicated to increase the size of the minority class. On the other hand, synthetic oversampling generates artificial samples for the minority class. These additional samples provide crucial information to the minority class, helping prevent misclassification of its instances. By employing oversampling techniques, the dataset becomes more balanced, leading to improved performance in classification tasks [25], [26], [27].

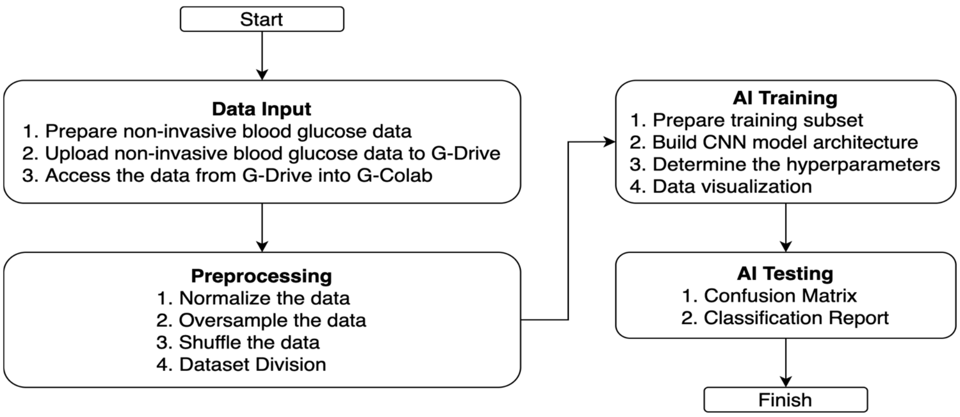

Fig. 2 The comprehensive research flow encompasses key stages such as data input, preprocessing, AI training, and AI testing.

4) Google Colaboratory: Google Colaboratory or Google Colab is an executable document that can be used to store, write, and share programs via Google Drive [28]. Like Jupyter Notebooks, Google Colab is a cloud program that runs on a browser. Google Colab allows developers to execute AI code without the installation process and other settings [29].

B. The Proposed Optimization Framework

As illustrated in Figure 2, the data files are initially uploaded to Google Drive (G-Drive), then loading the PPG data into the Google Collaboratory (G-Colab) environment. Subsequently, data preprocessing is undertaken to partition the data into training and testing subsets. An oversampling technique is then applied to equalize the number of samples in each class. The subsequent step involves constructing a Deep Learning (DL) model based on the LeNet-5 architecture. The DL model is trained to minimize error concerning the selected target, necessitating the setting and tuning of hyperparameters, including the number of neurons, epochs, learning rate, and batch size. Upon completion of the training phase, validation is conducted using a confusion matrix and classification report to assess the model's performance. This method is designed to yield a more accurate and reliable blood glucose classification system for non-invasive health monitoring.

1) PPG Data Acquisition:

For this research, twenty-three participants aged between 20 and 22 were enrolled. The group consisted of seven females and fifteen males. Before the recording sessions, all participants received thorough instructions and explanations about the measurement protocol. They were requested to attend the measurement sessions in the evening, one hour after breakfast. BGLs were assessed during the sessions using a gold-standard glucometer. Following breakfast, BGL and PPG signals were measured at the same time. Twenty seconds of measurements were collected (i.e., non-invasive data), encompassing a range of glucose levels between 65 and 245 mg/dL.

The data acquired from the CGM prototype [14] is systematically organized into files corresponding to each subject, employing the *.csv file format. The non-invasive series data, comprising 200 infrared data points and 200 red data points, extracted from the prototype, serves as the input for the AI model, while the invasive data functions as the corresponding labels. Subsequently, the data is uploaded to the Google Drive cloud storage. The data is then loaded into the Google Colab (G-Colab) environment to execute modeling, training, and performance validation of the prediction system. The data input procedure encompasses various steps, including:

- Prepare noninvasive blood glucose data: Noninvasive blood glucose data consists of PPG signals recorded using a CGM prototype and BGL data measured using a gold-standard glucometer.
- Upload non-invasive blood glucose data to cloud storage: The non-invasive blood glucose data is uploaded to G-Drive to facilitate data access and management for further analytics.
- Access the data from G-Drive into G-Colab: Data stored in G-Drive is accessed from the G-Colab environment for further preprocessing, AI modeling, and pattern validation.

2) Preprocessing the PPG Signals

During the preprocessing stage, essential steps are undertaken to adequately prepare the data for training and testing the system. These critical steps encompass:

- Normalize the data: Data is converted into a uniform and appropriate scale to facilitate model processing. In this case, the blood glucose data was normalized to have values between 0 and 1.
- Manage non-IID challenge: An oversampling technique is employed to address the non-IID nature of the blood glucose monitoring dataset. This method enhances data representation in minority classes by augmenting data points, thereby mitigating imbalances within the dataset.
- Shuffle the data: Data is randomized to reduce bias and ensure representativeness when divided into training and testing data.
- Dataset division: Non-invasive blood glucose data is divided into two parts, namely the training subset (80%) that will be used to train the model and the validation subset (20%) that will be used to test the performance of the proposed classification system.

3) Training the Proposed AI Model

During the training phase, the AI model, represented by a combination of neuron weights, undergoes iterative updates. The primary focus is on a subset of the training data, constituting 80% of the entire dataset. This training subset is pivotal in enhancing the Convolutional Neural Network (CNN) model's optimal weights by implementing the backpropagation algorithm.

- Prepare training subset: In this stage, a figure representing the training subset will be visualized as depicted in Figure 3. This visualization provides an overview of the distribution and patterns in the training data, making it easier to understand the characteristics of the input data.
- Build CNN model architecture: CNN is a type of deep learning architecture highly effective in image recognition tasks, making it suitable for use in non-invasive blood glucose classification. This research uses the LeNet-5 CNN architecture [30] which has been proven to classify spatial data points, e.g., data series and images.
- Determine the hyperparameters: At this stage, the number of neurons, the number of epochs, learning rate, and batch size will be determined. The number of epochs determines how often the entire training dataset will be used to train the model. The learning rate defines how aggressively the neuron's weights are updated, while the batch size refers to the number of data instances processed before the model weight update. This research uses 500 epochs, a learning rate of 0.001, and a batch size of 5.
- Plot training results: The training accuracy and loss are monitored so that researchers can observe how the model's performance evolves with each epoch. This visualization helps track the training progress and identify potential problems, such as overfitting or underfitting.

4) Validating the Trained AI Model

A validation subset, constituting 20% of the dataset, is utilized during the testing stage. Testing is executed using the optimal model determined through the preceding training process. The subsequent paragraphs outline the critical stages of the AI testing process.

- Confusion matrix: A confusion matrix measures the model's performance on the testing data. It is used to visualize the model prediction results and compare them with the actual value of the testing data. From the confusion matrix, the number of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) can be identified, which will be used to calculate various evaluation metrics.
- Classification report: The report encompasses diverse model performance evaluation metrics, including accuracy, precision, recall, and f1-score. Precision gauges the accuracy of optimistic predictions made by the model, recall assesses the model's ability to identify positive data correctly, and the f1-score represents the harmonic mean between precision and recall. Based on testing data, this comprehensive report thoroughly explains the model's efficacy in non-invasive blood glucose classification.

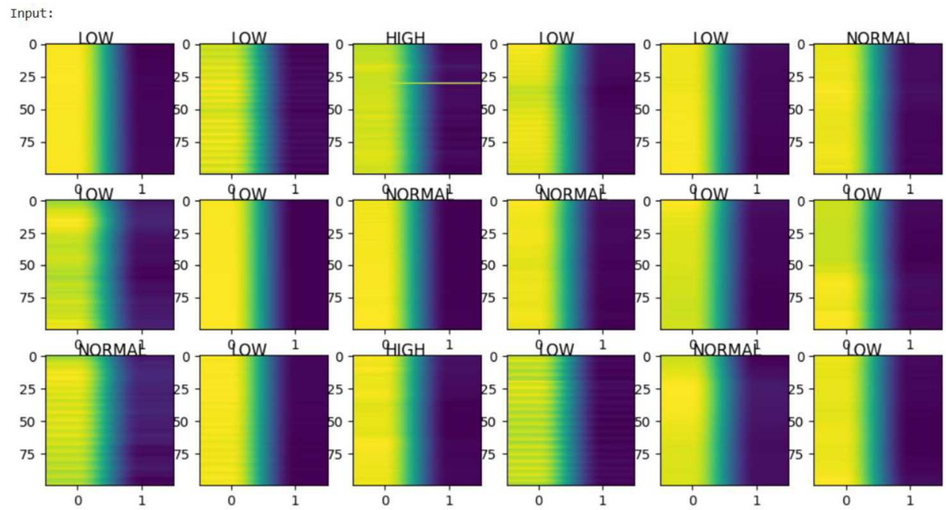

Fig. 3 Raw data visualization without oversampling

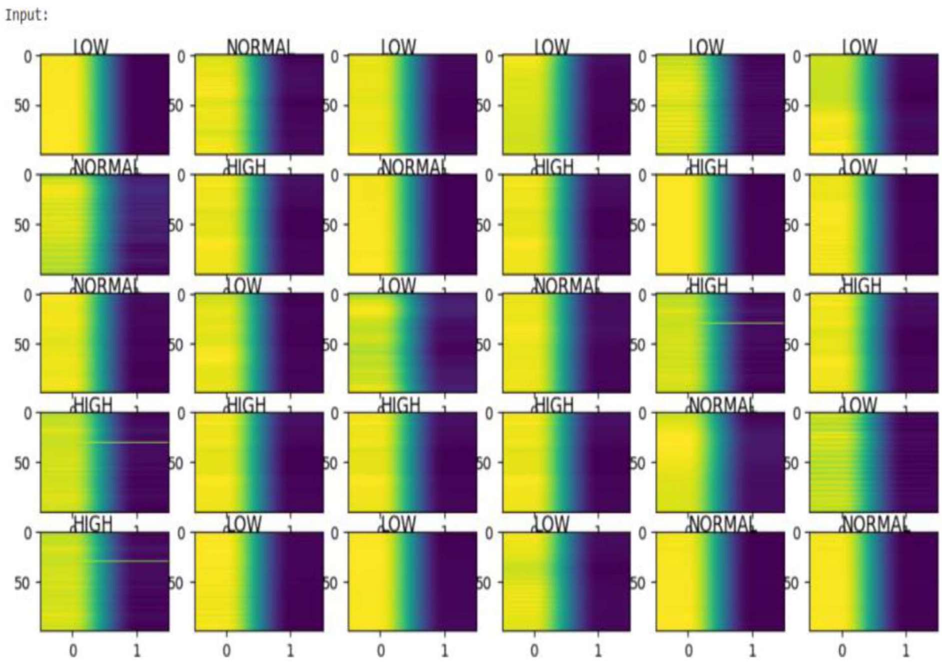

Fig. 4 Data visualization with oversampling reveal that the number of data increased by augmenting minority classes.

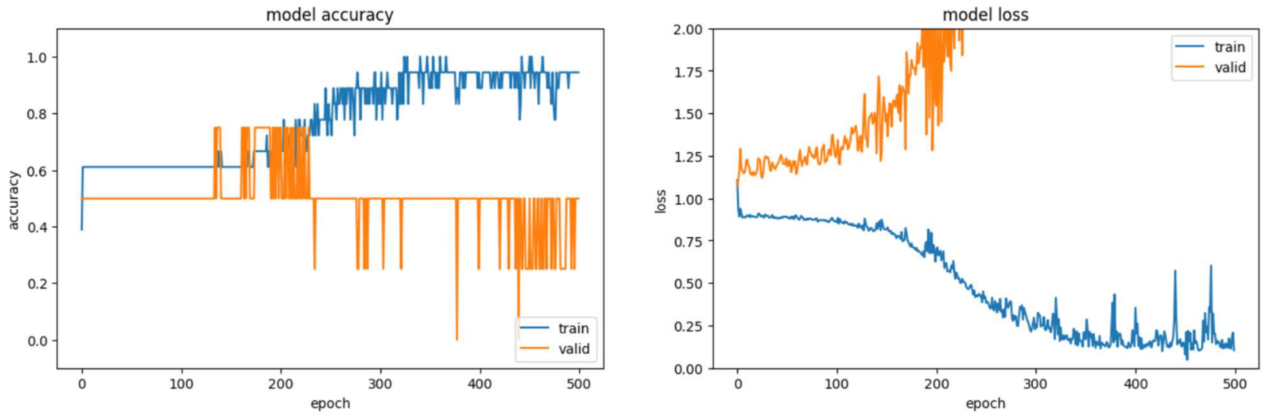

Fig. 5 Line diagram of training and testing accuracy of CNN(8,12) without oversampling

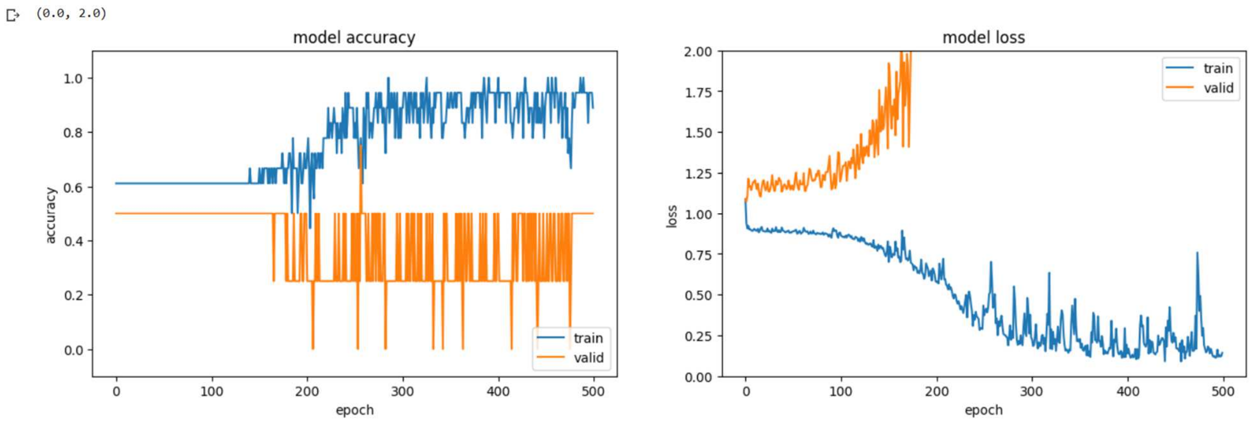

Fig. 6 Line diagram of training and testing accuracy of CNN(16,24) without oversampling

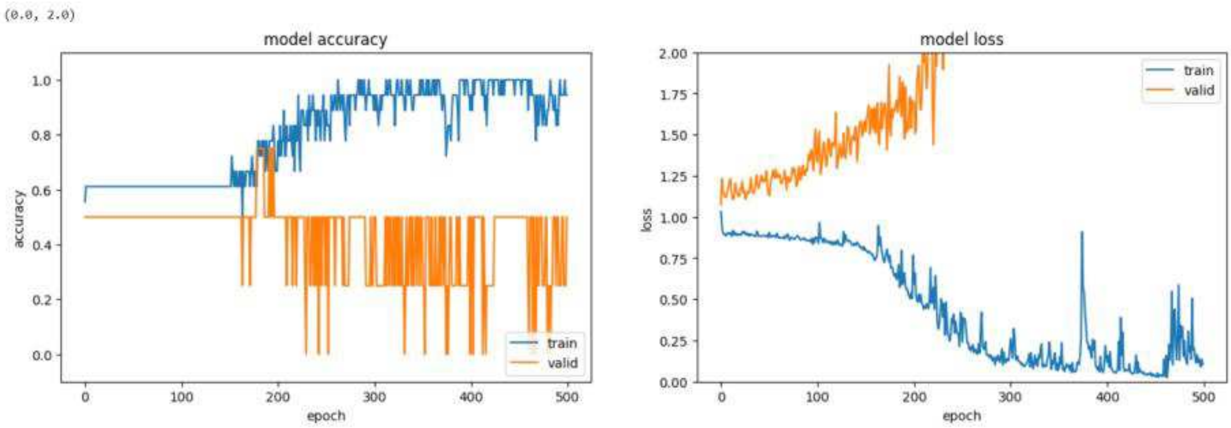

Fig. 7 Line diagram of training and testing accuracy of CNN(32,48) without oversampling

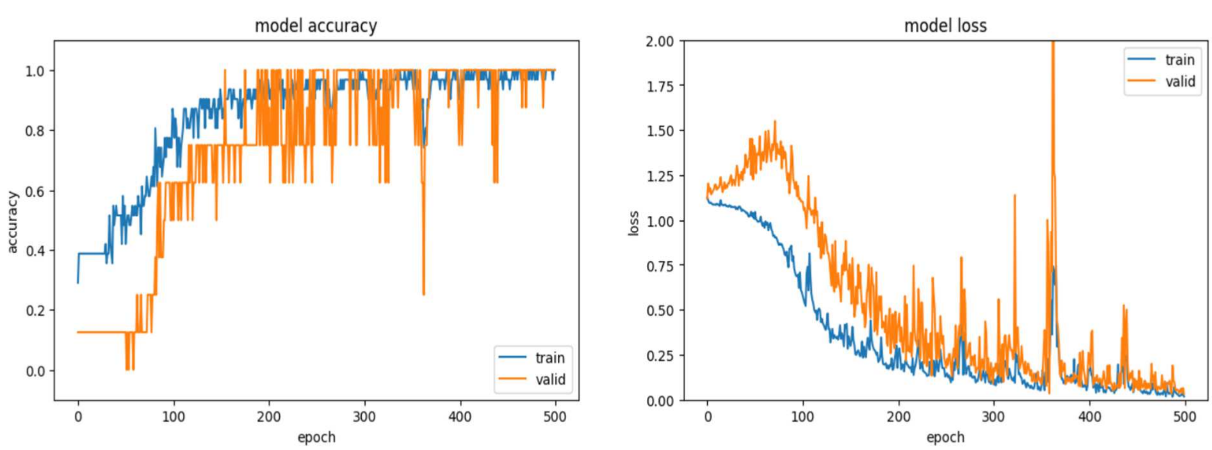

Fig. 8 Line diagram of training and testing accuracy of CNN(8,12) with oversampling.

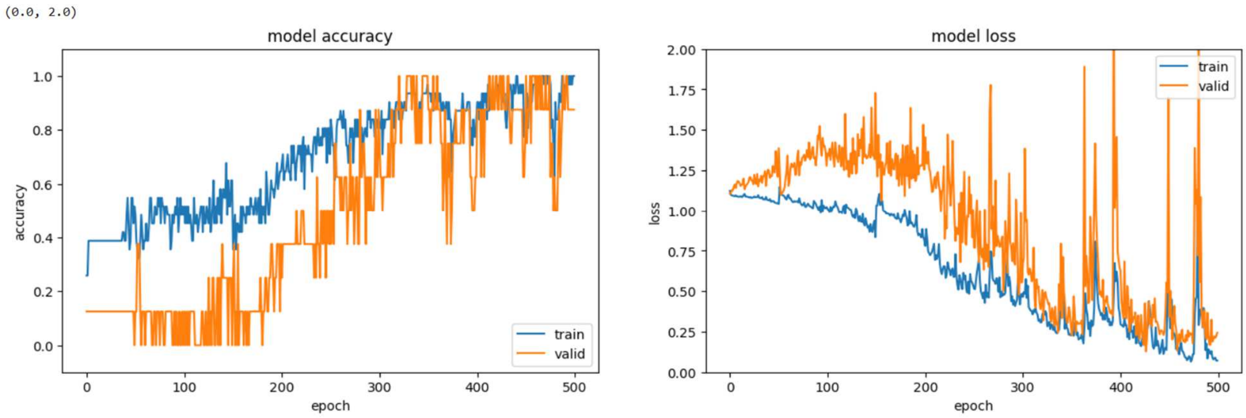

Fig. 9 Line diagram of training and testing accuracy of CNN(16,24) with oversampling

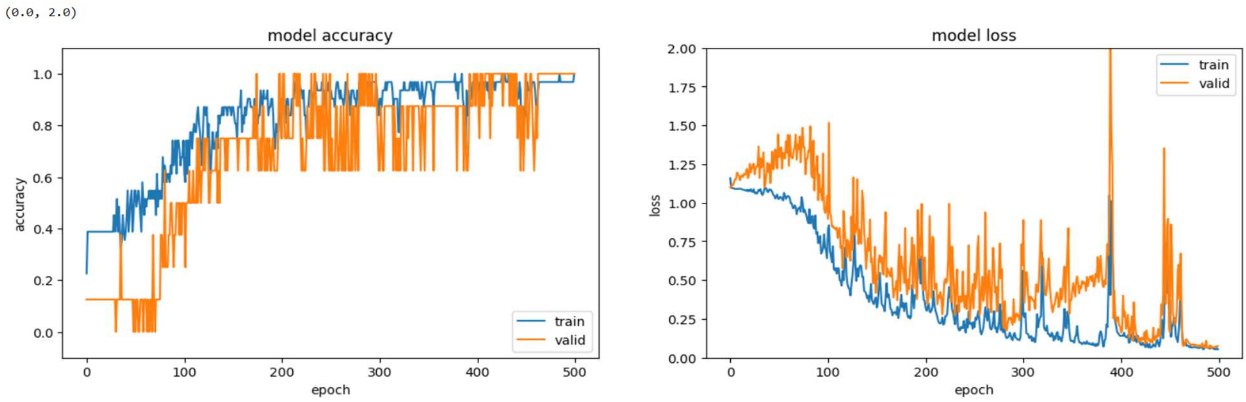

Fig. 10 Line diagram of training and testing accuracy of CNN(32,48) with oversampling.

III. Results and Discussion

This research focuses on optimizing the AI method using the CNN model to improve the performance of the noninvasive blood glucose classification system in non-IID settings. The procedure involves oversampling the data to increase the number of minority classes. It involves several stages of experimentation to achieve the desired results, which will be explained later.

In the training stage, Figure 3 visualizes figures representing the training subset, illustrating the features extracted from the red and infrared signals. Notably, these figures depict data before the application of the oversampling technique. Conversely, Figure 4 displays the training subset with random oversampling applied. Subsequently, a Convolutional Neural Network (CNN) model is established to process the data and extract meaningful features adeptly. Parameters, including the number of neurons, epochs, and batch size, are meticulously determined to refine the training process.

Following the training stage, the model undergoes testing using a dedicated testing data subset, constituting 20% of the dataset. The model's accuracy is assessed by comparing its predictions with the actual values of the testing data. The evaluation process entails constructing a confusion matrix to visually depict the model's performance and calculating diverse metrics such as precision, recall, and f1-score. These metrics collectively offer comprehensive insights into the model's classification performance.

TABLE I

The Performance of CNN(8,12) Without Oversampling

Classification Label	Precision	Recall	F1-score	Support
Low	0.50	0.50	0.50	2
Normal	1.00	1.00	1.00	1
High	0.00	0.00	0.00	1
Overall Accuracy			0.50	4

TABLE II

The Performance of CNN(16,24) Without Oversampling

Classification Label	Precision	Recall	F1-score	Support
Low	0.50	0.50	0.5	2
Normal	1.00	1.00	1.00	1
High	0.00	0.00	0.00	1
Overall Accuracy			0.50	4

TABLE III

The Performance of CNN(32,48) Without Oversampling

Classification Label	Precision	Recall	F1-score	Support
Low	1.00	0.50	0.67	2
Normal	0.50	1.00	0.67	1
High	0.00	0.00	0.00	1
Overall Accuracy			0.50	4

TABLE IV

The Performance of CNN(8,12) With Oversampling

Classification Label	Precision	Recall	F1-score	Support
Low	1.00	1.00	1.0	1
Normal	0.80	1.00	0.89	4
High	1.00	1.00	1.00	3
Overall Accuracy			0.88	8

TABLE V
THE PERFORMANCE OF CNN(16,24) WITH OVERSAMPLING

Classification Label	Precision	Recall	F1-score	Support
Low	1.00	1.00	1.00	1
Normal	0.80	1.00	0.89	4
High	1.00	1.00	1.00	3
Overall Accuracy			0.88	8

TABLE VI
THE PERFORMANCE OF CNN(32,48) WITH OVERSAMPLING

Classification Label	Precision	Recall	F1-score	Support
Low	1.00	1.00	1.00	1
Normal	1.00	1.00	1.00	4
High	1.00	1.00	1.00	3
Overall Accuracy			1.00	8

This research aims to create a robust and accurate non-invasive blood glucose classification system by optimizing the AI model and utilizing the power of the CNN model. Random oversampling is a valuable technique used to enhance the performance of PPG data analysis. In the context of PPG, the imbalance between different classes of physiological events can negatively impact the accuracy of predictive models. Random oversampling addresses this issue by generating synthetic data points for the minority class, thereby increasing its representation in the dataset. By duplicating or creating new instances of the underrepresented class, the model becomes better equipped to recognize patterns and make more precise predictions for all classes. This approach effectively mitigates the bias towards the majority class and prevents the model from being overly influenced by its prevalence. As a result, random oversampling is crucial in improving the performance and reliability of PPG data analysis, empowering researchers and healthcare practitioners to derive more accurate insights and make better-informed decisions in various applications, including disease detection, monitoring, and overall physiological assessment.

The CNN is employed for BGL prediction when developing the proposed AI model. During the training phase, the data undergoes preprocessing and is visually represented. Subsequently, the data is fed through modified convolution filter layers inspired by LeNet-5. The CNN, known for automatically extracting features from image data, effectively recognizes patterns and variations within blood glucose data. The number of neurons and filters is fine-tuned to enhance model performance. Moreover, the CNN model processes and classifies the data based on the data label, often acquired through a comparative device, i.e., a gold-standard glucometer.

In the training scenario without oversampling, 18 samples are backpropagated on the proposed CNN model with 500 epochs, a learning rate of 0.001, and a batch size of 5. The training results showed that the highest accuracy achieved was 0.6111, while the lowest loss was 0.9451. These results show that the model has successfully recognized patterns and variations in BGL data. The visualization of the training results for 500 epochs is shown as a line graph in Figure 5, Figure 6, and Figure 7. These figures reveal how the model performance increases along with the number of epochs performed. Then, in the testing stage, the optimal model generated from the training stage is used to test the model's performance on the testing data. Tables I, II, and III show the test results.

In the training scenario with oversampling, 39 samples are backpropagated on the proposed CNN model with 500 epochs and several batches of 8. The training results showed that the highest accuracy achieved was 1.000, while the lowest loss was 0.1051. This evaluation indicates that the deep learning model used in this study is very effective in classifying non-invasive blood glucose data with a very high level of accuracy. The visualization of the training results for 500 epochs is shown as a line graph in Figure 8, Figure 9, and Figure 10. The charts show how the model performance increases along with the number of epochs performed and consistently increases until it reaches the highest accuracy. The test results provide precision, recall, f1-score, and accuracy information. Tables IV, V, and VI show the test results' details.

The results of using the oversampling method in training DL models are very effective in improving model performance. Before oversampling, the model's precision, recall, and f1-score only reached 50%, which showed that the model had limitations in classifying non-invasive blood glucose data with high accuracy. However, with oversampling, the test results showed that the model's precision, recall, and f1-score increased to 100%. This evaluation indicates that the model can classify the data well and achieve high accuracy, making it reliable in predicting non-invasive blood glucose classification. The significant improvement in these evaluation metrics indicates that the use of the oversampling method successfully overcomes the problem of imbalance in the number of data samples in the minority class, as suggested by [25], [31]. Therefore, the model generated from training with oversampling is better able to recognize and classify data for each class, thus improving the accuracy and reliability of the overall non-invasive blood glucose classification system.

IV. CONCLUSION

This study introduces an innovative approach to medical data processing using the proposed DL framework. (1) This research successfully optimized a DL model by implementing a novel CNN technique, enhancing the efficacy of a non-invasive blood glucose classification system. The proposed CNN model demonstrates a high accuracy in blood glucose classification for both subset training and subset validation. The pinnacle of accuracy is attained through meticulous tuning of the CNN model. The best results are performed by the $CNN_{(32,48)}$, which utilized 32 and 48 filters for the first and the second layer, respectively. (2) This research successfully increased the number of data samples from 23 to 39 by applying the random oversampling technique, resulting in an increase of about 77.27% in the number of relevant data samples to improve the performance of the deep learning model in non-invasive blood glucose classification. The oversampling technique successfully overcomes the problem of non-IID dataset. (3) This study successfully demonstrated a significant improvement reaching 88% to 100% accuracy using the oversampling technique. This result shows that the random oversampling technique is effective in overcoming the problem of imbalance in the amount of data and improving the performance of the non-invasive blood glucose classification system. (4) Finally, this study successfully showed an increase in F1-Score for class 1 from 67% to more than 89% and for class 2 from 0% to 100% after using the random oversampling technique on the CNN model. For future research, non-invasive applications of PPG could be extended to predict hypoglycemic and hyperglycemic events and to investigate glucose levels in non-diabetic individuals, such as athletes or individuals interested in optimizing their health and performance.

Acknowledgment

This work was supported in part by Universitas Muhammadiyah Yogyakarta, Indonesia, through the Center of Artificial Intelligence and Robotic Studies, the Research and Innovation Centre, and the Cooperation and International Affairs under Grant 50/R-LRI/XII/2023; in part by Politeknik Negeri Semarang; in part by the Chelpis Quantum Tech Company Ltd., Taiwan, under Grant Asia University: I112IB120; and in part by Asia University, Taiwan; and part by the China Medical University Hospital, China Medical, University, Taiwan, under Grant ASIA-111-CMUH-16, Grant ASIA-110-CMUH-22, Grant ASIA108-CMUH-05, Grant ASIA-106-CMUH-04, and Grant ASIA-105-CMUH-04. This work was partly supported by the Department of Electrical Engineering, Politeknik Negeri Semarang.

References

[1] J. S. Flier, L. H. Underhill, and G. S. Eisenbarth, "Type I Diabetes Mellitus," New England Journal of Medicine, vol. 314, no. 21, pp. 1360–1368, May 1986, doi: 10.1056/nejm198605223142106.

[2] K. Gu, C. C. Cowie, and M. I. Harris, "Mortality in Adults With and Without Diabetes in a National Cohort of the U.S. Population, 1971–1993," Diabetes Care, vol. 21, no. 7, pp. 1138–1145, Jul. 1998, doi:10.2337/diacare.21.7.1138.

[3] C. J. Sokolowski, J. A. Giovannitti, and S. G. Boynes, "Needle Phobia: Etiology, Adverse Consequences, and Patient Management," Dental Clinics of North America, vol. 54, no. 4, pp. 731–744, Oct. 2010, doi:10.1016/j.cden.2010.06.012.

[4] "Introduction: Standards of Medical Care in Diabetes—2022," Diabetes Care, vol. 45, no. Supplement_1, pp. S1–S2, Dec. 2021, doi:10.2337/dc22-sint.

[5] M. B. Davidson, D. L. Schriger, A. L. Peters, and B. Lorber, "Relationship Between Fasting Plasma Glucose and Glycosylated Hemoglobin," JAMA, vol. 281, no. 13, p. 1203, Apr. 1999, doi:10.1001/jama.281.13.1203.

[6] G.-J. Eerdekens, S. Rex, and D. Mesotten, "Accuracy of Blood Glucose Measurement and Blood Glucose Targets," Journal of Diabetes Science and Technology, vol. 14, no. 3, pp. 553–559, Feb. 2020, doi: 10.1177/1932296820905581.

[7] L. Biagi, C. Ramkissoon, A. Facchinetti, Y. Leal, and J. Vehi, "Modeling the Error of the Medtronic Paradigm Veo Enlite Glucose Sensor," Sensors, vol. 17, no. 6, p. 1361, Jun. 2017, doi:10.3390/s17061361.

[8] O. Schubert-Olesen, J. Kröger, T. Siegmund, U. Thurm, and M. Halle, "Continuous Glucose Monitoring and Physical Activity," International Journal of Environmental Research and Public Health, vol. 19, no. 19, p. 12296, Sep. 2022, doi: 10.3390/ijerph191912296.

[9] J. Zhou et al., "Reference Values for Continuous Glucose Monitoring in Chinese Subjects," Diabetes Care, vol. 32, no. 7, pp. 1188–1193, Apr. 2009, doi: 10.2337/dc09-0076.

[10] Prayitno et al., "A Systematic Review of Federated Learning in the Healthcare Area: From the Perspective of Data Properties and Applications," Applied Sciences, vol. 11, no. 23, p. 11191, Nov. 2021, doi: 10.3390/app112311191.

[11] S. Siddiqi, F. Qureshi, S. Lindstaedt, and R. Kern, "Detecting Outliers in Non-IID Data: A Systematic Literature Review," IEEE Access, vol. 11, pp. 70333–70352, 2023, doi: 10.1109/access.2023.3294096.

[12] W. Wang, M. Zhang, Z. Wu, Q. Chen, and Y. Li, "MDFD: Study of Distributed Non-IID Scenarios and Frechet Distance-Based Evaluation," 2023 IEEE International Conference on Image Processing (ICIP), Oct. 2023, doi: 10.1109/icip49359.2023.10222893.

[13] K. T. Putra et al., "A Review on the Application of Internet of Medical Things in Wearable Personal Health Monitoring: A Cloud-Edge Artificial Intelligence Approach," IEEE Access, vol. 12, pp. 21437–21452, 2024, doi: 10.1109/access.2024.3358827.

[14] K. T. Putra, I. Surahmat, A. N. Nazilah Chamim, M. Z. Ramadhan, D. Wicaksana, and R. A. Dhea Namyra Alissa, "Continuous Glucose Monitoring: A Non-Invasive Approach for Improved Daily Healthcare," 2023 3rd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS), Aug. 2023, doi: 10.1109/ice3is59323.2023.10335328.

[15] H. Rashtian, S. S. Torbaghan, S. Rahili, M. Snyder, and N. Aghaeepour, "Heart Rate and CGM Feature Representation Diabetes Detection From Heart Rate: Learning Joint Features of Heart Rate and Continuous Glucose Monitors Yields Better Representations," IEEE Access, vol. 9, pp. 83234–83240, 2021, doi:10.1109/access.2021.3085544.

[16] A. Facchinetti, S. Favero, G. Sparacino, and C. Cobelli, "An Online Failure Detection Method of the Glucose Sensor-Insulin Pump System: Improved Overnight Safety of Type-1 Diabetic Subjects," IEEE Transactions on Biomedical Engineering, vol. 60, no. 2, pp. 406–416, Feb. 2013, doi: 10.1109/tbme.2012.2227256.

[17] B. Lobo, L. Farhy, M. Shafiei, and B. Kovatchev, "A Data-Driven Approach to Classifying Daily Continuous Glucose Monitoring (CGM) Time Series," IEEE Transactions on Biomedical Engineering, vol. 69, no. 2, pp. 654–665, Feb. 2022, doi: 10.1109/tbme.2021.3103127.

[18] C. Huang, Y. Xiao, and G. Xu, "Predicting Human Intention-Behavior Through EEG Signal Analysis Using Multi-Scale CNN," IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 18, no. 5, pp. 1722–1729, Sep. 2021, doi: 10.1109/tcbb.2020.3039834.

[19] V. Senger and R. Tetzlaff, "New Signal Processing Methods for the Development of Seizure Warning Devices in Epilepsy," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 63, no. 5, pp. 609–616, May 2016, doi: 10.1109/tcsi.2016.2553278.

[20] "Video for A CNN Based Human Bowel Sound Segment Recognition Algorithm with Reduced Computation Complexity for Wearable Healthcare System", doi: 10.1109/iscas45731.2020.9180432/video.

[21] S. Zhang, S. M. H. Bamakan, Q. Qu, and S. Li, "Learning for Personalized Medicine: A Comprehensive Review From a Deep Learning Perspective," IEEE Reviews in Biomedical Engineering, vol. 12, pp. 194–208, 2019, doi: 10.1109/rbme.2018.2864254.

[22] M. Elhadary et al., "Revolutionizing chronic lymphocytic leukemia diagnosis: A deep dive into the diverse applications of machine learning," Blood Reviews, vol. 62, p. 101134, Nov. 2023, doi:10.1016/j.blre.2023.101134.

[23] O. Abdel-Hamid, A. Mohamed, H. Jiang, and G. Penn, "Applying Convolutional Neural Networks concepts to hybrid NN-HMM model for speech recognition," 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Mar. 2012, doi:10.1109/icassp.2012.6288864.

[24] A. Ajit, K. Acharya, and A. Samanta, "A Review of Convolutional Neural Networks," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Feb. 2020, doi: 10.1109/ic-etite47903.2020.049.

[25] N. Cahyana, S. Khomsah, and A. S. Aribowo, "Improving Imbalanced Dataset Classification Using Oversampling and Gradient Boosting," 2019 5th International Conference on Science in Information Technology (ICSITech), Oct. 2019, doi:10.1109/icsitech46713.2019.8987499.

[26] S. Maldonado, J. López, and C. Vairetti, "An alternative SMOTE oversampling strategy for high-dimensional datasets," Applied Soft Computing, vol. 76, pp. 380–389, Mar. 2019, doi:10.1016/j.asoc.2018.12.024.

[27] S. J. Basha, S. R. Madala, K. Vivek, E. S. Kumar, and T. Ammannamma, "A Review on Imbalanced Data Classification Techniques," 2022 International Conference on Advanced Computing Technologies and Applications (ICACTA), Mar. 2022, doi:10.1109/icacta54488.2022.9753392.

[28] T. Potluri et al., "Secure Software Development in Google Colab," 2023 IEEE World AI IoT Congress (AIIoT), Jun. 2023, doi:10.1109/aiiot58121.2023.10174336.

[29] N. L. Z. Msomi and B. A. Thango, "Development of Dissolved Gas Analysis-based Fault identification System using Machine Learning with Google Colab," 2023 31st Southern African Universities Power

Engineering Conference (SAUPEC), Jan. 2023, doi:10.1109/saupec57889.2023.10057713.
[30] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998, doi: 10.1109/5.726791.
[31] F. Rodríguez-Torres, J. F. Martínez-Trinidad, and J. A. Carrasco-Ochoa, “An Oversampling Method for Class Imbalance Problems on Large Datasets,” Applied Sciences, vol. 12, no. 7, p. 3424, Mar. 2022, doi: 10.3390/app12073424.