Heliyon 10 (2024) e36112

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

5²CelPress

An explainable Artificial Intelligence software system for predicting diabetes

Parvathaneni Naga Srinivasu^{a,b}, Shakeel Ahmed^{c,*}, Mahmoud Hassaballah^{d,e}, Naif Almusallam^f

^a Department of Teleinformatics Engineering, Federal University of Ceará, Fortaleza, 60455-970, Brazil

^b Amrita School of Computing, Amrita Vishwa Vidyapeetham, Amaravati, 522503, Andhra Pradesh, India

^c Department of Computer Science, College of Computer Sciences and Information Technology, King Faisal University, Al-Ahsa, 31982, Saudi Arabia

^d Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj, 16278, Saudi

Arabia

^e Department of Computer Science, Faculty of Computers and Information, South Valley University, Qena, Egypt

^f Department of Management Information Systems, College of Business Administration, King Faisal University, Al-Ahsa, 31982, Saudi Arabia

ARTICLE INFO

Keywords: Convolutional neural networks Bi-LSTM Blood glucose levels Spectrogram images Hyperparameters ROC curves

ABSTRACT

Implementing diabetes surveillance systems is paramount to mitigate the risk of incurring substantial medical expenses. Currently, blood glucose is measured by minimally invasive methods, which involve extracting a small blood sample and transmitting it to a blood glucose meter. This method is deemed discomforting for individuals who are undergoing it. The present study introduces an Explainable Artificial Intelligence (XAI) system, which aims to create an intelligible machine capable of explaining expected outcomes and decision models. To this end, we analyze abnormal glucose levels by utilizing Bi-directional Long Short-Term Memory (Bi-LSTM) and Convolutional Neural Network (CNN). In this regard, the glucose levels are acquired through the glucose oxidase (GOD) strips placed over the human body. Later, the signal data is converted to the spectrogram images, classified as low glucose, average glucose, and abnormal glucose levels. The labeled spectrogram images are then used to train the individualized monitoring model. The proposed XAI model to track real-time glucose levels uses the XAI-driven architecture in its feature processing. The model's effectiveness is evaluated by analyzing the performance of the proposed model and several evolutionary metrics used in the confusion matrix. The data revealed in the study demonstrate that the proposed model effectively identifies individuals with elevated glucose levels.

1. Introduction

Diabetes is a chronic condition that has been known for decades. Several instances, however, are detected in their late stages. Diabetes affects one out of every eleven adults worldwide. The World Health Organization (WHO) predicts that by 2040, the number of individuals with diabetes will rise to 642 million, or one in every 10 people [1]. In addition to the four main categories, there are other sub-categories. A person develops type 1 diabetes, also called insulin-dependent Type-2 diabetes, when the insulin production in the pancreas stops working. Insufficient insulin production by the body characterizes type-2 diabetes. People over 40 are more likely to

* Corresponding author. E-mail address: shakeel@kfu.edu.sa (S. Ahmed).

https://doi.org/10.1016/j.heliyon.2024.e36112

Received 28 January 2024; Received in revised form 9 August 2024; Accepted 9 August 2024

Available online 10 August 2024





^{2405-8440/© 2024} The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

have this problem. Gestational diabetes (GDM) occurs most often during pregnancy. prediabetes which develops whenever blood sugar levels are elevated but not severe enough to qualify as type-2 diabetes, is the fourth main category [2]. As a result, early diagnosis and appropriate treatment of diabetes may help prevent impediments and reduce the risk of serious health consequences. Several bioinformatics researchers have taken on this problem by building methods and tools for diabetes prediction. Several body composition changes may aid the early identification of Type 2 diabetes. Maintaining a good dietary cycle is critical in minimizing the risk of developing diabetes and its complications. Our presentation emphasizes earlier research investigations on the diagnostics of Type 2 diabetes and early identification of diabetes based on real-time glucose data [3].

Wearable, non-invasive, and intelligent patient monitoring technologies are in high demand. As a result of previous research, wearable glucose monitoring devices based on various technologies have been developed. It was shown that glucose concentrations could be measured using plasmonic, carbon nanotube, and fluorescence sensors [4,5]. Lengthy calibration procedures, large size, and autofluorescence sensitivity were some of these approaches' drawbacks [6,7]. There is a necessity for a model that could precisely diagnose the abnormal glucose level in the body and timely notify the primary healthcare centers to provide the appropriate medication. Due to life obligations, financial restraints, and an insufficient number of medical practitioners, particularly in rural and developing regions, communities may not seek regular health exams, resulting in late diagnosis of illnesses that can cause serious sickness if neglected or treated late [8].

The main purpose of our XAI modeling is trustworthiness since we intend to maximize a Convolution layer more reliably so that possible mishaps may be avoided. As a result, our XAI model increases the trustworthiness of a Classification algorithm by employing feasibility analysis, with some sample images and visible explanations about the decision mechanism [9]. It is a technique for estimating the impact of every input parameter. First, we assign a particular value to each input variable and feed the changed input vector into the model. Resultantly, one can calculate how much the output differs from the model's output vector whenever the input is the original input data. So, we can observe which pieces of the input greatly affect the output vector, and hence which portion of the input causes the model to determine correctly or incorrectly [10]. Some of these things may result in a major mishap if the decisions being made by these systems are incorrect. At the same time, the evaluator can approximate the total number of neurons used in deep learning decision-making models but is unsure of the exact number for each neuron in these models.

The current study primarily focuses on real-time monitoring of diabetic patients to track abnormal glucose studies. The existing studies acquire the blood glucose levels as a numerical entity and alarm the caretakers when they experience an abnormal glucose value, rather in the current study, the signal data is converted to the spectrogram images, and those spectrogram images are analyzed to determine the glucose levels as the signal to numerical conversion is prone to errors. Moreover, determining the abnormal glucose level from signal data is less effective [11]. Analyzing the spectrogram images to determine the glucose levels using a deep learning model is the novelty aspect of the current study [12,13]. Analyzing the spectrogram images for glucose tracking through the deep



Fig. 1. Spectrogram images represent the blood glucose levels.

learning model is a first of its kind. Bi-LSTM [14] would assist in precisely identifying the abnormal glucose levels. The XAI technology assists in fostering trust and enhancing transparency inside Artificial Intelligence (AI) systems, particularly in high-stakes sectors such as healthcare. This can help individuals make more informed decisions, comprehend the rationale behind suggestions or actions, and subsequently take suitable measures based on such insights.

Due to the rising percentage of geriatric people in the population, there is an expanding demand for assisted-living options that allow senior citizens to maintain their independence in their own homes for an extended period of time, thus decreasing their reliance on caretakers. The proposed methodology is largely motivated by assisted technology, which continuously monitors individuals for abnormal glucose levels. The primary objective of real-time monitoring is to make immediate adjustments to treatment protocols in order to optimize outcomes. Continuous monitoring promotes patient engagement in condition management through the provision of individualized, actionable information regarding their health in real-time [15]. The utilization of wearable technology permits healthcare professionals to monitor patients' health status remotely, eliminating the necessity for frequent in-person consultations, and thereby reducing the cost of healthcare services. The proposed methodology would allow the monitoring of multiple patients simultaneously.

The blood glucose levels are acquired through the GOD strips placed over the human body. Later, the signal data is converted to the spectrogram images, which are classified as low glucose, average glucose, and abnormal glucose levels. The labeled spectrogram images are then used to train the model for individualized monitoring. Multi-resolution wavelet analysis decomposes the mixed signal into sub-signals of diverse frequency bands. Uses wavelet transform's breakdown and rebuilding operations to split signals into discrete frequencies and eliminate high-frequency band noise from the output signal. A signal is reconstructed using low-frequency components. The short-time Fourier transform (STFT) converts the signal to the spectrogram images. Fig. 1 presents the spectrogram images used in the model's training and validation process. Where sub-figure (A) and (B) present the spectrogram images of a patient with lower glucose levels. The sub-figures (C) and (D) denote the spectrogram images are classified based on the features mapped with the glucose level. In the current study, the spectrogram images with normal glucose levels are considered as one class, and the spectrogram images with abnormal glucose levels are considered as another class.

This research compared several machine learning-based approaches for predicting diabetes and reported our findings in a comparative analysis. The following is a rundown of the many contributions that this research has made.

- To discuss and outline briefly various cutting-edge technologies in diabetes prediction.
- Acquisition and processing of the spectrogram images plotted from the blood glucose levels.
- To mechanize the model that can remember the outcomes of the previous state using Bi-LSTM for better precision.
- The machine interpretability in the current study's feature engineering process would assist in determining the feature significance in the decision process.
- Analyzing the model's performance with other contemporary techniques concerning various evaluation metrics.

The remaining sections of the paper were arranged as follows, Section 2 Section 3 discusses the process of spectrogram image generation and pre-processing, which is necessitated for precise classification of the spectrogram images. Section 4 discusses the proposed XAI framework for feature engineering. Section 5 discusses the CNN with Bi-LSTM for classifying the abnormal glucose levels from the spectrogram images. Section 6 provides the results obtained from experimentation and their statistical analysis, and Section 7 gives the conclusion and the future scope.

2. Literature review

Automated data-driven methods do not need an understanding of the structural and physiological processes since they utilize the individual's collected data. An individual's degree and duration of physical activity, as well as their experience with hypoglycemia and the function of their pancreas, influence blood glucose dynamics in people with diabetes. Previously, models with input parameter combinations accounted for these aspects were examined. In recent years, several ML and DL-based models for diabetes prediction have been designed, including Logistic Regression (LR), XGBoost (XGB), Multilayer Perceptron (MLP), Support Vector Machine (SVM), Artificial Neural Network (ANN), Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Decision Tree (DT), Random Forest (RF), AdaBoost (AB), J48, k-Nearest Neighbours (k-NN), random forest (RF), and Light Gradient Boosting Machine (LGBM) [16,17]. Various feature reduction and cross-validation strategies and the process for handling the missing data and discarding outliers have been used to boost the efficiency of ML and DL models [18,19].

Lin et al. [20] evaluated Naive Bayes, SVM, and ANN classifiers in their study utilizing a diabetic dataset. They conducted a weighted investigation in which most votes determined the likelihood of diabetes prognosis. Finally, they determined that combining models improves classification accuracy compared to single algorithm-driven models. On the diabetes dataset, Kandhasamy et al. [21] suggested a predictive analytic model employing J48 (C4.5), k-NN classifier, RF classifier, and SVM. They planned that the J48 method would perform better than others before pre-processing data, with 73.82 % accuracy. Still, KNN and Random Forest would perform a much better way of performing the pre-processing. Dey et al. [22] employed ensemble classification mechanisms like SVM, KNN, Naive Bayes, and ANN with Min-Max scaling (MMS) over a structured dataset like Pima. The ANN models with MMS have a greater accuracy of 82.35 percent than the other four methods. Sonar and JayaMalini [23] built a model to effectively identify diabetes patients through machine learning methods, including DT, ANN, NB, and SVM classification algorithms, where the DT has obtained an accuracy rate of 85 %. Beloufa and Chikh [24], report an 86.13 % accuracy associated with Support Vector Regression (SVR) and an ANN-based

algorithm for diabetes identification.

A study on the Gaussian Hidden Markov Model (GHMM) method has attained 85.69 % accuracy [25], and a Deep Extreme Learning Machine (DELM) oriented prediction technique has achieved 92.8 % accuracy [26]. Sanakal and Jayakumari [27] utilized FCM clustering and SVM for the precise classification of diabetic records over the PIMA dataset records and achieved 94.3 % accuracy. Compared to earlier efforts, Hasan Temurtas et al. [28] utilized the Levenberg–Marquardt (LM) technique for the diabetic diagnostic model. It has obtained an accuracy of 82.37 %. Another study based on CNN-LSTM for diabetes classification has achieved 90.9 % test accuracy. CNN model over the 5-fold cross-validation accuracy was 93.6 %, whereas CNN with LSTM's 95.1 %. In the study on diabetes prediction [29], the Authors gathered 289 cases and 13 characteristics from the Khulna Diabetes Center. In our analysis, we apply Logistic Regression (88 %), XGB (86.36 %), and Random Forest (86.36 %). Saha et al. [30] employed ML methods like neural networks (NN), SVM, and Random Forest to predict diabetes. Preprocessing techniques like imputation, data normalization, and feature selection techniques like principal component analysis were employed. The Neural Network was the best and most consistent model with setup settings for 100 epochs with a batch size of 10. It obtained an accuracy of 80.4 %. Wei et al. [31] deployed a DNN model for preventing type 2 diabetes and observed an accuracy of around 77.86 % percent. After performing the scaling, the ReLu activation over the Nadam optimizer was over 100 epochs with a batch size of 10. Smartphones and wearable sensor data were utilized to construct a disease pre-warning system [32].

This system classifies health issues using data analysis, ontologies, and deep learning. However, all the models mentioned above rely on pre-existing data to predict diabetes in patients. The models need to be robust to handle the real-time data, and they need a tremendous volume of data to train them to attain considerable accuracy. The proposed model is designed to monitor the real-time blood glucose level and alarm the individuals for timely medication or appropriately notify the caretakers. There is another concern related to data imbalance in diabetes prediction. Particularly big is the data of blood glucose measurements taken at normal levels which is considered as the majority class, e.g., hypoglycemia, and then the dataset of blood glucose measurements taken over the specific symptoms, which is considered as the minority class. An unbalanced dataset leads to a model with a biased outcome, which means that the precision of the minority class is substantially worse than that of the class label. This is another pivotal reason for migrating toward real-time blood glucose analysis, which is the present research problem. Most often, the RF and SVM models were used, with the DT model being utilized in conjunction with the RF model in certain circumstances. However, performance metrics revealed that the DT model frequently scored worse than the RF model [33]. The SVM model employed a variety of kernel functions because of the variety of models that were utilized. Regression models such as SVR and linear regression are often used. The LR model was used rather than the linear regression model because of the categorical nature of the dataset. The multinomial LR model was utilized since the intended feature classification output had not been binaries but three-class (low, normal, and high blood glucose level) problems to be evaluated [34].XAI technology is used in various healthcare domains to make the decision process evident in various healthcare applications [35,36]. The existing studies on XAI for healthcare have proven to outperform in efficiency and the ability to explain the decision model. A study on XAI for image analysis through a deep learning model [37]. The performances of various state-of-the-art models used in assessing Type-2 diabetes using ML and clinical approaches (CA) are presented in Table 1.

3. Spectrogram image generation and pre-processing

The spectrogram images are generated from the real-time blood glucose levels observed from the sensors connected to the human body. Blood glucose levels may be seen as a time series signal, with time over the x-axis and blood glucose concentration over the yaxis. Short-time Fourier transforms a series of Fourier transformations performed on a windowed input signal. When a signal's frequency components vary over time, the STFT offers time-localized frequency information, whereas the conventional Fourier transform delivers frequency components averaged across the whole signal period. Using multi-resolution wavelet analysis, the chaotic signal is

Table 1

State-of-art techniques used in Type-2 diabetes Prediction.

Technique	Algorithm	Data Type	Overall Accuracy	Remark
ML Technique	LR	Tabular Data	75.2 %	Used in assessing diabetes from pre-existing data using
Ahamed et al. [2]	XGB		83.3 %	feature selection.
	GBC		94.1 %	
	DT		94.4 %	
	RF		94.8 %	
	LGBM		95.2 %	
ML Technique Kulkarni et al. [38]	XGB	ECG data	96.8 %	The study is limited to the early detection of diabetes.
ML Technique	NB	Tabular Data	86.1 %	Prediction of diabetes from the pre-existing tabular data
Ahmed et al. [39]	DT		96.8 %	by feature selection.
	GBC		91.0 %	
CA Approach Saha and Saha [40]	RCT	Real-time blood sample data	95.0 %	The approach is invasive and needs frequent finger pricking.
ML Technique	NB	IoT and Embedded systems for real-	84.1 %	The approach is invasive and needs finger pricking.
Shen et al. [41]	J48	time data	99.7 %	
	LR		86.0 %	
	RF		99.6 %	

broken down into component signals in their unique frequency ranges. The wavelet transform's decomposition and reconstruction techniques are used in signal denoising to separate the noisy signal into its component bandwidths and eliminate the noisy high-frequency band. The low-frequency components are used to reconstruct a signal. Short Fourier transformations can be used to modify this signal to provide a considerably improved frequency distribution throughout time. Equation (1) demonstrates the converted signal identified by the variable $O_s(\kappa)$.

$$O_{s}(\kappa) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\kappa t} I_{s}(\lambda) \omega_{s}(\lambda - t) d\lambda$$
⁽¹⁾

From the above Equation, the variable $I_S(\lambda)$ denotes the original input signal to the STFT function, operated through the window $\omega_s(\lambda)$, where the length of the window impacts total resolving power concerning time and frequency. The STFT may be considerably enhanced by adding a window to a non-stationarity input. With this window, the frequency obtains a new temporal dimension. The intensity magnitude is determined by Equations (2) and (3), which assumes a homogenous signal filter band.

$$\Theta(s) = \omega \times \sum_{p = -\infty}^{\infty} m(p) \times (\alpha - p)$$
⁽²⁾

$$f(\alpha,\omega) = \Theta(s) \times \sum_{p=-\infty}^{\infty} e^{-c\omega p}$$
 (3)

From the above Equation, the variables α and ω designate the time and frequency components in the signal processing. The variable *m* represents the related window function with a zero-centered interval *p*.

The change in glucose levels may manifest as changes in the frequency content of the signal, which would be visible as variations in the spectrogram. For example, higher glucose levels might correspond to increased activity in certain frequency bands, leading to more intense areas in the spectrogram. Windowing functions may emphasize certain frequency elements in the spectrogram. To emphasize low-frequency components, the window functions with a peak in the frequency domain. One such example is the Gaussian window. The mathematical definition of the Gaussian window $\omega_g(x)$ is shown in Equation (4).

$$\omega_g(\mathbf{x}) = e^{-\frac{\mathbf{x}^2}{2\sigma^2}} \tag{4}$$

 σ is a parameter that determines the width of the main lobe. A bigger σ number creates a broader main lobe, emphasizing lower frequencies. The filtering method is employed to extract low-frequency components from spectrogram images. The mathematical expression for the output Y(f) of a low-pass filter applied to the spectrogram image X(f) is as shown in Equation (5).

$$Y(f) = X(f) \times H(f) \tag{5}$$

From the above equation, H(f) represents the low-pass filter's frequency response. This filter selectively passes sounds below a predetermined cutoff frequency, highlighting low-frequency elements in the spectrogram.

The dataset employed in this study is processed for the class imbalance issue, where numerous spectrogram images of abnormal glucose instances are present, which slows down the model training process and impacts the model's accuracy. The modified entropy loss function is used in the current study to address this issue. By lowering the weight of easily classifiable samples, this function may improve the model's focus on the difficult-to-classify samples in the training phase. The modified cross-entropy loss function formula is shown in Equation (6).

$$L_{f} = -tx(1-x')^{-1}\log x' - (1-x)\log (1-x')$$
(6)

From the above equation, the variables *t* target balancing the positive and negative data samples concerning loss value. The variable η parameter directs the model's attention to indistinguishable data samples, improving the model's overall classification accuracy [9]. Furthermore, the quality of the spectrogram images is being enhanced for better processing and recognition of the image features. The spectrogram images are processed using the Otsu-based Adaptive Weighted Bilateral Kernel (AWBK) for enhanced contrast. AWBK approach uses the Otsu thresholding approach. The non-linear kernels maintaining the target's boundary information (edge) may reduce noise by noisy aggregating sites in the picture concerning pixels directly around it. Assuming the deformed pixel is reconstructed using the contra-harmonic average of pixel intensities and that the weight will be significant. The weight's significance decreases as we step away from the pixel's centroid and becomes negligible far away [36]. In which the average pixel intensity of the surrounding pixels has been maintained, and the noise has been disregarded, the noise would be weakly correlated. When a low pass kernel is used to compute the average of the pixel intensities during the image enhancement stage, it will smoothen the picture, i.e., blur the area of the image, but the bilateral edge kernel. The texture information is kept just for processing convenience. To utilize the bilateral kernel, the pixel's degree of belongingness D_x and maximum likelihood m_s , A Gaussian distribution function is used over the centroid. The weighted sum of image pixels p of kernel size k is determined using Equation (7).

$$B_f = \frac{1}{m_s} \sum_{K} D_x(\|a - b\|) m_s(\|i_a - i_b\|)$$
(7)

From the above equation, the variables a and b denote the corresponding neighboring pixel, and similarly, the variables i_a and i_b

denotes the intensities of both the pixels a and b, respectively. The variable D_x is used to assess the corresponding pixel's belongingness. The variable m_s denotes the contra-harmonic mean calculated as shown in Equation (8).

$$m_{s} = \frac{\alpha_{1}^{2} + \alpha_{2}^{2} + \alpha_{3}^{2} + \dots + \alpha_{n}^{2}}{\alpha_{1} + \alpha_{2} + \alpha_{3} + \dots + \alpha_{n}}$$
(8)

The thresholding strategy, such as Otsu, reduces the interclass correlation variance by evaluating the weighted probabilistic average across several classes that may be assessed using the following process for a spectrogram image. ω_p and ω_q the probability weights of categories *m* and *n* are indeed categorized using an approximation of the threshold *t* as shown in Equation (9).

$$V(t) = \omega_p \nu_p(t) + \omega_q \nu_q(t) \tag{9}$$

On pre-processing the spectrogram images, the quality of images has tremendously improved, and they would assist in better classifying the spectrograms by abnormal glucose levels. The spectrogram images after the pre-processing are shown in Fig. 2, where the sub-figure (a) represents the original spectrogram image and (b) represents the pre-processed image. The histogram of the original image and the pre-processed image is shown in Fig. 3 for understanding the impact of the pre-processing technique, where the sub-figure (a) is corresponding to the original spectrogram image and (b) is corresponding to the pre-processed spectrogram image. Where the sequences are distributed across the image rather than accumulating towards a confined range of frequencies as in the original spectrogram image.

The spectrogram images are generated from the glucose levels at the instance of time, but the long-term dependencies are not considered. Long-term dependencies are desired for better precision of the classification outcome. There is a possibility of Temporal Resolution Loss (TRL) [42] and Dynamic Range Compression (DRC) [43], which leads to a considerable loss in the spectrogram image generation process. The current study was limited to addressing both the challenges that are discussed above.

4. XAI-driven feature engineering

The weights of neurons cannot be immediately interpreted as information, and the size and sensitivity of activations are not suitable indicators of a neuron importance for a specific task. As we know, the deep learning models are still not fully understood. A full examination of the underlying architecture, procedures, and predictable statics that machine-interpretable and explainable models can obtain. Diagnostic data is analyzed using XAI techniques to better display the outcomes of a machine learning system [44]. Tasks like feature extraction and feature extraction with feature weights fall within the purview of feature engineering. The XAI model is then used to optimize the loss function and initialize the feature weights. The elements considered more important in the assessment phase are given greater weightage than the rest. Transparency is anticipated to ensure that the model is visible. The corresponding figure for XAI-driven feature engineering is shown in Fig. 4.

Each iteration updates the weights in each neural network layer. The performance of the classification algorithm depends on initial weights, biases, and activation functions. Transparency of weight approximation functions helps reveal the Operational Procedure Model (OPM). In weight initialization, more important traits are given higher weights and are later optimized [37]. Equation (10) shows the weight initialization process for better model transparency.

$$\omega = \sum_{x=1}^{p} \sum_{y=1}^{q} \left| \omega_{xy} \times \omega_{yo} \right| \tag{10}$$

From Equation (7), the variable ω_{xy} designates the association among the weights of the input node *x* and the corresponding hidden node *y*. Similarly, the weight ω_{yo} represents the associated weights among the hidden node *y* and the corresponding output node represented by *o*. Over iterations, weights are tweaked for greater performance. Considering model-specific characteristics and loss functions optimizes weight. The loss caused over the training phase is denoted by the variable provided in Equation (11) using the variable L_{train} .

The variable ω_y , represents the feature's weight, which is being updated to use the validation and training component $\{\omega_y\}_{y=1}^{j_t}$. Equation (12) determines the loss function linked with the feature weights from data used for training, which should be the smallest.



(a) Original Spectrogram Image

(b) Pre-processed Spectrogram Image

Fig. 2. Image representing the original and pre-processed spectrogram image.



Fig. 3. Histograms of original and pre-processed images.



Fig. 4. Block diagram of the proposed XAI-driven CNN with Bi-LSTM model.

$$L_{train} = \frac{1}{f_t} \sum_{\nu=1}^{f_t} f_x(\theta) \times \omega_y \tag{11}$$

$$\omega' = \frac{1}{f_t} \sum_{y=1}^{f_t} f'_y(\theta \times \omega)$$
(12)

Throughout the training phase, the weight is optimized, the number is finetuned, and the values are advised to be higher than zero for an effective diagnostic model [2]. The accurate assessment depends largely on feature selection, and the feature selection method is desired to be transparent for any trustworthy system. The feature set that is thought useless is deleted from deliberation in the assessment process. In contrast, the feature set presumed to be significant is made more obvious [45]. Equations (13) and (14) are then used to update the fitness for the most important feature.

$$f = f_x + \eta((t_u - t_l)\vartheta + t_l)\gamma \ge 0$$
(13)

$$f' = f_x - \eta((t_u - t_l)\vartheta + t_l)\gamma < 0 \tag{14}$$

From the above Equations, the variables t_u and t_l denotes the upper and the lower threshold values respectively. The other coefficients, like η , ϑ , γ , are the balancing factors, whose values are calculated using Equations (15) and (16).

$$\eta = 2e^{-\left(\frac{4m}{m_t}\right)^2} \tag{15}$$

 $\vartheta = \delta_x$ (16)

Where in the above equations, the variable *m* denotes the current iteration and the m_t denotes the sum of all epochs. The variable δ_x

denotes the chaotic gradient of the current epoch. The value of *g* is likely determined during the training or optimization phase of the model. It may be adjusted based on empirical results, experimentation, or domain knowledge. Suppose we have a loss function $L(\gamma)$ that we want to minimize with respect to the variable *g*. This loss function could represent the error of our model on a training dataset or some other measure of performance. Set $g = g_0$ with a learning rate α that controls the step size in each iteration. Determine the gradient of the loss function relative to *g* is designated by $\frac{\partial L}{\partial r}$. The *g* values are updated as shown in Equation (17).

$$g' = g - \alpha \frac{\partial L}{\partial g} \tag{17}$$

In the above equation, the g' is the new updated value of g that is determined based on the old value. Equation (18) determines the fitness value for assessing the feature set.

$$f(\mathbf{x}) = \omega \times acc + (1 - \omega) \times \left(1 - \frac{I_{gain}}{tot_f}\right)$$
(18)

The *acc* variable denotes the accuracy associated with the current epoch, I_{gain} is the associated information gain, and the variable *tot_f* denotes the total number of associated features.

5. CNN with Bi-LSTM for spectrogram classification

The integration of CNN and Bi-LSTM models in this approach effectively exploits the respective advantages of each architecture, enabling the simultaneous extraction of spatial characteristics and temporal dependencies within the dataset. CNN has shown its efficacy in capturing spatial characteristics present in images, including but not limited to edges, textures, and forms. On the other hand, Bi-LSTM networks excel at capturing temporal patterns and dependencies that exist within sequences. Also, the Bi-LSTM models avoid the overfitting issue.

Convolutional neural network is a type of Multilayer perceptron (MLP). The following are some of how they are analogous to neural networks: The neurons are updated with the weights and biases that need to be acquired through training. There is some input reaching each neuron. After that, a dot product implementation is carried out, which may or may not be followed by a non-linear function. Initial applications of CNNs were found in image processing, where the network would take in unprocessed image pixels on one end, convert them via a series of hidden layers, and then produce class scores on the other. The major layers of CNN may be broken down into three categories. The convolutional, pooling, and fully connected layers long with rectified linear activation (ReLU) functions are the layers included in the network [38]. Finally, the softmax layer is used in probabilistic measures to determine the belongingness of the class. In this case, CNN uses one-dimensional time-series data organized in order of successive time instants. A CNN's primary notion is to extract local characteristics using upper-layer data and transmit it to subsequent layers for more significant features. The layered network architecture of the CNN with the Bi-LSTM model is illustrated in Fig. 5. The size of the input spectrogram image in the current study is 224×224 . The original spectrogram images are resized before they are fed as input to the CNN model.

The convolutional layer has kernels for determining associated feature maps. Those kernels roll out over the input using stride to



Fig. 5. The layered network diagram of the CNN with the Bi-LSTM model.

create integer output volume dimensions. The input volume shrinks when the convolution is utilized to stride [39]. Padding with zeros is needed to pad the input vector with zeros and keep its low-level dimensions. The equation for the convolutional layer over the input vector v over the 2D kernel k of size $m \times n$ is shown in Equation (19).

$$f(p,q) = \sum \sum s(p+m,q+n) \times k(m,n)$$
⁽¹⁹⁾

ReLU layers boost non-linearity in the feature maps [41]. ReLU calculates activation with zero thresholds, as shown in Equation (20).

$$f(r) = \max\left(0, l\right) \tag{20}$$

The pooling layer down samples a given input dimension to limit the number of parameters. The most frequent approach is max pooling, which creates the highest values over the input region. The FC layer serves as a classifier, making decisions based on the characteristics gathered from convolution and pooling. To find the probability distribution over labels, use the softmax function. The likelihood of classifying that use the softmax function. The equation for softmax is given in Equation (21).

$$p(m^{(x)} = y | n^{(x)}; \theta) = \frac{e^{(\theta_y^x) n^{(x)}}}{\sum\limits_{k=1}^{z} e^{(\theta_y^x) n^{(x)}}}$$
(21)

For example, when a convolution accompanies a normalizing procedure, the batch normalization settings were examined to speed up training. All of the batch normalization settings are constant. The convolution process may be used to include these constant parameters. In the training process, the batch normalization b_n is assessed through the standard deviation σ and mean m over the input component i, with an assumption that for each mini-batch m_b , the variance should be 1 or the average should be 0.

$$b_n = \sigma \times \frac{i - \Upsilon}{\sqrt{(m_b)^2 + \kappa}} + m \tag{22}$$

$$m_b^2 = \frac{1}{p} \sum_{x=1}^{p} (i - Y)^2$$
(23)

$$\Upsilon = \frac{1}{p} \sum_{x=1}^{p} i \tag{24}$$

From Equations (22)–(24), the variable κ denotes the hyperparameters used in finetuning the training time of the model. Resultantly, the learning phase is stabilized, and the sum of epochs in the training phase required to build neural networks is significantly reduced.

5.1. Integration on Bidirectional-LSTM

The Bi-LSTM component consists of an LSTM component that could function in either of the directions, allowing it to include previous and prospective context summaries. Bi-LSTM would learn long-term dependencies by normalizing the redundant background information [46]. The findings have demonstrated outstanding efficiency in sequential modeling issues and are widely utilized in text categorization. The Bi-LSTM framework, when compared to LSTM architecture, includes dual concurrent layers that move in different directions by forwarding and backward passes to capture interdependence in other contexts. It takes the features created by the CNN process and identifies them from the final hidden layer. Bi-LSTM provides feasibility in accessing previous and following context data, and the information produced by Bi-LSTM may be thought of as two distinct futuristic representations. The CNN features are supplied into a Bi-LSTM model, which generates a sequence characterization. This final feature map is put into an attention layer, selecting the characteristics strongly associated with the final classification. The forward LSTM network learns from prior data in the forward move, whereas the backward LSTM learns from forthcoming values in the opposite direction. In the final layer, the learned information from the previous concealed states is merged [46]. Equations (20) and (21) describe the processes carried out in the Bi-LSTM unit identified by *BiLSTM* over the forward LSTM component $\overline{f(x_t)}$ and the backward component $\overline{f(x_t)}$.

$$(\mathbf{x}_{t}) = \sigma(\omega_{1}\mathbf{p}_{t} + \omega_{2}\mathbf{p}_{t-1}) \times \tanh(\mathbf{k}_{t})$$
(25)

$$\overleftarrow{\mathbf{f}(\mathbf{x}_t)} = \sigma(\omega_3 \mathbf{p}_t + \omega_4 \mathbf{p}'_{t-1}) \times \tanh(\mathbf{k}'_t)$$
(26)

$$BiLSTM = \omega_5 \overline{f(\mathbf{x}_t)} + \omega_6 \overline{f(\mathbf{x}_t)}$$
(27)

From Equations (25)–(27), the variable p_t is the input at the time stamp t, over the weights $\omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ associated with the gates. An L2 regularize, and an activation function tanh are used to complete normalization, which may assist in preventing over-fitting. When the network is run, the following layer combines vectors of size $\alpha \times \beta$, where the variable β denotes the neurons in each LSTM unit.

5.2. Batch normalization

Batch Normalization (BN) is often used to enhance the efficiency of the training process, which influences the accuracy of the prediction. When tuning the learning rate to be high, most neural networks may explode or vanish. In such a context, batch normalization is to solve these issues. Normalizing activation throughout the network helps to prevent modest changes to parameters that might result in these parameters amplifying or disappearing so abruptly. Since implementing batch normalization and ReLU, classification accuracy has dramatically increased. Batch normalization led to faster training, shorter training and testing times, and decreased sensitivity to initialization in our trials. Before the activation function, batch normalization is implemented, which is represented as shown in Equation (28).

$$n = \omega p \times \beta$$

(28)

The weight is denoted by the variable ω , and the variable β denotes the associated bias. The normalization is performed over the input component *p*, where Normalization satisfies the convolution property, in which various parts of the same feature map are normalized in the same manner at different instances.

5.3. Dataset description

This information comes from the HR waves created by the GOD sensors connected to the human body that monitor glucose levels. The spectrogram images are generated using a short-time Fourier transform. The glucose levels are manually tagged as -1, 0, and 1 to properly identify the spectrogram images. After the images have been tagged, the CNN with the Bi-LSTM model is trained on the data. To make the model simple to implement and evaluate, the spectrogram images are considered as two classes, i.e., the images with normal and abnormal glucose levels. The total dataset consists of 6735 spectrogram images, among which 1347 images are used to identify people with aberrant glucose levels, of which 2806 are related to normal glucose levels (GL normal), 2658 samples are high glucose levels (HG), and 1271 are linked to low glucose levels (GL lows) (abnormal GL). The entire dataset comprises 80-20 training and testing sets, respectively. The same dataset images are shown in Fig. 6. The images are randomly picked from the dataset and shown in a 4 × 4 grid.



Fig. 6. Images with class labels used in model building.

5.4. Hyperparameters

The hyperparameter choices for every model used in the tests are described in this section. Determining the best hyper-parameters is difficult and changes depending on the dataset's features [47,48]. For that purpose, we ran multiple experiments to determine which hyper-parameters would provide the highest performance outcomes for the classifiers. Hyperparameters of the CNN and Bi-LSTM layers are presented in Table 2. It is not significant that data points were zero-padded before processing them as input for the deep neural networks, ensuring that all input vectors have the same size. The feature vectors are supplied to a fully-connected layer with 64 nodes in the CNN as a baseline model. A ReLu function and the norm approach train and regularize the dense hidden layers. The associated weights are then driven to the output layer using a sigmoid activation function, which calculates the final classification probabilities. During the training phase of the model, the Adam algorithm is optimized with a predefined initial learning rate of 0.00001, and loss minimization of the cross-entropy mechanism is used [49]. The training and testing instances are obtained from the same image dataset with a proportionate split of the samples. The validation set used to update and fine-tune the model and update the hyperparameters is an integral part of the testing samples.

A machine learning model's overall performance may be significantly enhanced using a loss function. During the training and validation phases, the loss is calculated, and its significance is determined by how well the model performs in each set. It is the sum of errors made in either the training or validation sets for each episode [50]. An accuracy measure is used to quantify the algorithm's performance in a useful way. The accuracy is usually estimated after the parameters that describe the model's correctness to the real true labels have been determined. Fig. 7 illustrates the hyperparameters, where the sub-figure (a) represents the model loss graphs and sub-figure (b) represents the accuracy graphs, and Table 2 provides an in-depth discussion of the environmental conditions used in this investigation. Both of these are provided in the context of the current research. The graphs show that the model has shown decent behavior without excessively fitting to the training data.

The learning rate of a deep learning model is extremely important in assessing the model's performance. In a very low learning rate, the model's training process is relatively slower in adjusting the associated weights in the network. Similarly, a higher learning rate might result in deviation from the desired outcome. The optimal learning rate limit is attained at the model's first point of divergence. When selecting the ideal learning rate, the loss should ideally continue to decrease at this point. The ideal learning rate of the Bi-LSTM with CNN model is highlighted with a red dot over the learning rate curve shown in Fig. 8.

5.5. Implementation environment

The proposed model is implemented on a stand-alone computer working in a Windows environment. The details of the implementation platform and the packages used in its evaluation process are presented in Table 3.

5.6. Model explainability using SHAP

The model explainability is exceptionally significant in healthcare studies like tumor identification, healthcare informatics, and real-time monitoring to ensure the AI model's transparency and trustworthiness in the classification process. In the current study, Shapley values are used to explain the significance of the feature in making the prediction. Shapley values measure the feature's prediction contribution. They explain how input characteristics affect model prediction. The black-box models, like deep learning mechanisms, need to make the feature contribution explainable. The Shapley values associated with the spectrogram images are shown in Fig. 9.

Spectrogram images visually depict glucose dynamics across time, highlighting patterns that may not be apparent from raw data alone. SHAP values provide a way to recognize each input feature's contribution to the prediction model's output, allowing for a more granular understanding of model predictions. In the context of spectrogram analysis, SHAP values can help identify which features or frequency components in the images are most influential in predicting glucose levels. This increased interpretability may help discover biomarkers, comprehend physiological processes, and eventually improve diabetes treatment options. Furthermore, the openness

CNN + Bi-LSTM		
0.00001		
4		
128		
2		
ReLu (CNN) and Tanh (Bi-LSTM)		
30		
0.3		
64		
Adam		
2 imes 2		
40		
3		

Details	of the	hyperparamete	ers.

Table 2



Fig. 7. Hyperparameters associated with CNN with Bi-LSTM.



Fig. 8. Learning rate graph of CNN with Bi-LSTM.

Table 3
Details of Experimental setup.

Environment Parameters	Specifications	
Machine OS	Windows 11 64-Bit	
Processor	Intel core TM i7-8550U CPU @ 1.99 GHz	
Memory	8 GB DDR3 RAM	
Implementation Language	Python	
Implementation Platform	PyTorch, DL	
Packages	Numpy, pandas, Scikit-learn, sklearn, Flask	

afforded by XAI approaches, such as SHAP values, is critical in establishing confidence in AI-powered healthcare systems.

In dependency graphs of SHAP values, pink nodes represent features that positively affect predictions, while blue nodes represent features that have a negative impact. This color scheme helps visualize how different features contribute to model predictions. This gradual shift in colors visually illustrates the varying degrees of influence that features have on model predictions, allowing for a clearer understanding of the complex relationships within the data. The SHAP method effectively demonstrates its value in combining model-agnostic and model-specific justifications. By considering both the cumulative imputed values and the characteristics in the input vector. All possible feature combinations, including and excluding the feature set, must be checked for correctness before calculating the Shapley value. The number of potential combinations has also increased significantly.

As indicated by Equation (29), the explainable model's results are affected by the set of characteristics to which it is coupled. Independent variables *m* and *n* in the range up to *x*, are utilized to estimate the interdependence among the features. The Shapley value that is approximated is represented through the $\theta(m)$ notation.

(30)



Fig. 9. The SHAP values that are associated with the input data.

$$\theta(m) = \left\{ \frac{\delta m}{\delta n_1} + \frac{\delta m}{\delta n_2} + \dots + \frac{\delta m}{\delta n_x} \right\}$$
(29)

The notation θ stands for the change in weights associated with each feature vector. Equation (30) illustrates the connection between the feature vector and weight changes.

$$\delta m = [\omega_1 \times n_1, \omega_2 \times n_2, \dots, \omega_x \times n_x]$$

The XAI models are considerably important in determining the feature significance in making the outcome. Shapley values provide insights into how individual characteristics affect the overall forecast by quantifying the marginal effect of each input variable, enabling transparency and interpretability.

6. Results and discussion

The CNN with Bi-LSTM model performance is measured over the underlying facts linked with the spectrogram images. Other studies have used CNN with Bi-LSTM in the healthcare domain studies as EE-Based Emotion Recognition [51], schizophrenia detection using MSST-spectral images [52], epileptic seizure detection [53]. Similarly, there are various other domains in which the CNN with Bi-LSTM has outperformed in classification. The suggested CNN with the Bi-LSTM model is being assessed by measuring the model's true positive, true negative, false positive, and false-negative predictions. Based on the assessments, measures such as sensitivity, specificity, accuracy, and recall are calculated. Proper prediction of normal glucose level is termed true positive, while correct detection of aberrant glucose level is considered true negative. Similarly, misinterpreting excessive glucose levels as normal glucose

13

levels results in a false positive, whereas misinterpreting normal glucose levels as abnormal glucose levels results in a false negative. Fig. 10 depicts the confusion matrix linked with the predictions.

The performance evaluation measures like sensitivity, specificity, accuracy, and F1-score are assessed from the generated confusion matrix [54,55]. The performances of various other state-of-the-art models concerning the above-discussed metrics are presented in Table 4. The comparison includes various approaches like Small-World Feed Forward Artificial Neural Network (SW-FFANN), ResNet18, ResNet50, SVM, Adaptive neuro-fuzzy inference system (ANFIS), Logistic Regression(LR), and Gaussian support vector machine (GSVM). The comparison is made over divergent datasets like PIMA, photoplethysmographic (PPG), Real-time sensor data, and electrocardiographic (ECG) signal data. The N/A in the table indicates the data is not available. The CNN with Bi-LSTM models has exhibited a reasonable accuracy compared to the other models used in the same field of study.

It can be depicted from the tabulated values shown in Table 4 that the efficiency of the CNN with Bi-LSTM performs well analyzed with the other robust models in the classification of diabetic individuals. Moreover, the Bi-LSTM with CNN model works with real-time spectrogram images rather than tabular datasets like PIMA. The ROC curve is a graph-based representation of the trade-offs between false and true positive rates. The X-axis usually presents the false-positive numbers, and the Y-axis presents the true-positive values. A true-positive rate closer to 1 is desired. The ROC curve of the CNN with the Bi-LSTM model is presented in Fig. 11.

It can be depicted from the ROC curves that the CNN with Bi-LSTM model has proven to have a fair and better performance in classifying the spectrogram images associated with abnormal glucose levels. Cross-validation (CV), is one of the many performance assessment measures that may be used to classify issues. CV operates by folding the data into numerous folds and guaranteeing that each fold is used as a test dataset. Its single value parameter specifies the number of folds, or groups that should be created from each test data sample before it can be validated. As a result, we call this procedure fold cross-validation. The results of applying the model's evaluation to several folds are shown in Table 5.

Results show that the models' performance has continuously increased across the data samples in the assessment procedure. The diagnostic Odds Ratio (DOR) is a measure used to evaluate the effectiveness of a certain prediction associated with illness. DOR is calculated by dividing the likelihood that the test will return a positive result if the patient has an abnormal glucose level by the probability that the test would produce a positive result if the patient has a normal glucose level [67].

$$DOR = \frac{True \ (positive \times negative)}{False \ (positive \times negative)}$$
(31)

In the current study, the DOR obtained is 949.6 on a standalone direct execution of the model. A test with high specificity and sensitivity and a low occurrence of false positives and false negatives has a higher DOR. DOR rises when test specificity improves, even while the test's sensitivity remains constant.





14

Table 4

Table representing the performances of various State-of-art models.

Approach	Dataset	Sensitivity	Specificity	Accuracy	F1-Score
SW-FFANN [56]	PIMA	0.85	0.96	0.91	N/A
ResNet18 [55]	PIMA	0.88	0.66	0.80	0.85
ResNet50 [55]	PIMA	0.94	0.57	0.80	0.85
SVM [55]	PIMA	0.95	0.83	0.90	0.93
SVM + RF + MLP + ANFIS [57]	Real-time sensor data	N/A	N/A	0.90	N/A
Deep transfer learning [58]	OhioT1DM dataset	0.59	0.98	0.95	0.61
LR [59]	PPG Signal data	0.73	0.64	0.69	N/A
GSVM [60]	PPG Signal data	0.79	0.83	0.81	N/A
Bayesian classifier [61]	PPG Signal data	1.00	0.87	0.93	N/A
SVM [62]	PPG Signal data	0.98	0.96	0.97	N/A
AdaBoost [63]	ECG Signal data	0.92	0.88	0.90	N/A
IGRNet [64]	ECG Signal data	0.80	0.77	0.77	N/A
GoogLeNet [64]	ECG Signal data	0.69	0.83	0.75	N/A
AlexNet [64]	ECG Signal data	0.77	0.82	0.74	N/A
HoG with K-NN [64]	ECG Images	0.70	0.79	0.76	N/A
AlexNet [19]	Spectrogram images	0.93	0.96	0.95	0.92
ResNet [19]	Spectrogram images	0.93	0.96	0.95	0.92
KNN [12]	Spectrogram images	0.89	0.89	0.86	0.86
SVM [12]	Spectrogram images	0.91	0.91	0.89	0.89
DenseNet-121 [65]	Spectrogram images	0.72	0.73	0.98	N/A
ResNet with SVM [66]	Spectrogram images	0.98	0.98	0.98	0.98
Proposed Approach	Spectrogram images	0.97	0.96	0.96	0.93



Fig. 11. ROC curve associated with CNN with Bi-LSTM.

Table 5					
Accuracy	of	the	Bi-LSTM	model	concerning
multi-fold	cro	ss-va	lidation.		

k value	Accuracy
2	0.964
3	0.951
4	0.956
5	0.959
6	0.969
7	0.972
8	0.979

6.1. Future perspective model

The future perspective model is a web-based application, where the caretakers are given a provision to access the real-time blood glucose levels and the heartbeat rate of the patients that are part of the ambient assisted community. The application would provide them the ability to access real-time healthcare data, update patient-related information, notify the patient individually or as a group, perform multi-patient monitoring simultaneously, and provide them with access to the messages that are sent by the patients. The chatbot facility would allow the future perspective model to have a friendly conversation with the patients for better convenience. The

application is a Flask web framework that relies on the Jinja template engine to interface the business logic in the backend with the web interface for the user interface. NoSQL MongoDB [68] is used to save information on users and the data being tracked on them. The model is provided with user authentication to avoid unauthorized access to sensitive data. Fig. 12 shows the authentication interface to the application, and Fig. 13 shows the patient search page, where the user profile is searched based on the reference number of the patient. Fig. 14 shows the patient's dashboard, where the real-time monitoring details of glucose level and heartbeat rate are shown along with other basic information.

The future perspective model would rely on the spectrogram images generated from the sensor signal attached to the patient's body. The model's performance largely depends on the quality of the spectrogram image. The other potential challenge associated with the body sensor-based diabetes assessment from the spectrogram images includes the time frame between the spectrogram images and the length of the spectrogram images. The images are stripped to a confined length, impacting the model's performance. The current study uses spectrogram images of length 5sec to make predictions. However, more statistical analysis is needed on the length of the spectrogram images to ensure the accuracy of the prediction. The size of the spectrogram image is exceptionally significant in determining the prediction accuracy, as the larger spectrogram images can hold more signal data, which would result in better analysis of the underlying patterns in the data more precisely. The size of the spectrogram images is confined and fixed throughout the experimental analysis, which could be considered as one of the potential limitations of the current study. The hyperparameter tuning could yield a better accuracy [69], which was not done in the current study is also a potential limitation of the study.

7. Conclusion

The current study primarily categorizes the spectrogram images with abnormal glucose levels. By recognizing them, the model intends to notify the caretakers about the situation and provide timely medication and treatment. The CNN with Bi-LSTM model has performed well in precisely classifying the spectrogram images, making it unique. Considering the data samples as two classes, i.e., normal and abnormal glucose levels, the model has obtained a reasonable accuracy of 96.44 %. The findings demonstrate Bi-LSTM's ability to execute sequential data models and retrieve context data through forwarding and backward correlations over the dependencies from feature sequences. Compared to baseline approaches, the Bi-LSTM-driven model is more productive and accurate in classification quality. The XAI techniques would better assist in making the decision model more interpretable to the stakeholder. The model has outperformed the previous models in the current study concerning accuracy and F1-measure findings. The future perspective model would assist the caretakers in monitoring the patients remotely for a better monitoring experience.

The data in the current study is considered to be two classes, which could be considered as three classes, i.e., normal, low, and high blood glucose levels. So that the caretakers are given more appropriate information about the patient's condition, which would result in better medical outcomes. The acquisition of spectrogram images is a tedious task, and they are obtained at fixed time intervals, which could be considered one of the major technical limitations of the current study. The images of shorter time intervals could result in better outcomes. In the future, designing a model with minimal computational latency is necessary as the model is deployed over a real-time scenario to assist the patients in a much better way. Ablation studies are needed for problems such as diabetes prediction, which is considered to be the limitation of the results section of the current study. The ablation study could be considered in future research for better analysis of the performances of the model. Patient-specific sub-class training and classification would result in better results. The hyperparameters in the current study are decided based on the existing studies, and fine-tuning of the hyperparameters could result in better accurate results. The optimization of the hyperparameters is one of the future research directions of the current study. The multi-class problem with three classes, normal, low, and high blood glucose levels, is considered binary, i.e., normal and abnormal blood glucose levels. Multi-class analysis of the blood glucose levels could be fruitful and can be considered as a future research direction.

Funding

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King

	Diaheed
Username: Password:	Username Password
	Login Clear
	rve gov userinanire passovimu
Future P	erspective model for real time diabetes monitoring

Fig. 12. Agent authentication page of future perspective model.



Fig. 13. The patient search page of the future perspective model.

ඛ		Diaheed
(0)	Name: Lucas Jackson Patient Ref. No.: 09J71145	NO NO 100 100 200 200 200 200 200
\boxtimes	Age: 65	Diabetes Monitoring
æ Ø	Location: Bulnes Illness: Diabetes Type-2	₩
Q		marcuntelesservellesserve
		Heartbeat Monitoring
		Refresh
		Future Perspective model for real time diabetes monitoring

Fig. 14. Patient dashboard of the future perspective model.

Faisal University, Saudi Arabia [Grant No. 5633].

Ethical approval

The authors declare that they have no ethical approval required for this article.

Data availability statement

The authors declare that the research data will be available for readers on request.

CRediT authorship contribution statement

Parvathaneni Naga Srinivasu: Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Shakeel Ahmed:** Supervision, Software, Investigation, Formal analysis, Data curation. **Mahmoud Hassaballah:** Writing – original draft, Validation, Resources, Project administration, Investigation, Conceptualization. **Naif Almusallam:** Validation, Supervision, Resources, Funding acquisition, Formal analysis, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- I. Tasin, T.U. Nabil, S. Islam, R. Khan, Diabetes prediction using machine learning and explainable AI techniques, Healthcare Technology Letters 10 (2023) 1–10.
- [2] B.S. Ahamed, M.S. Arya, A.O. Nancy V, Prediction of type-2 diabetes mellitus disease using machine learning classifiers and techniques, Front. Comput. Sci. 4 (2022) 835242.
- [3] L. Singh, M. Khanna, D. Singh, Feature subset selection through nature inspired computing for efficient glaucoma classification from fun- dus images, Multimed. Tool. Appl. (2024) 1–72, https://doi.org/10.1007/s11042-024-18624-y.
- [4] S. Ahmadi, N. Rabiee, M. Bagherzadeh, F. Elmi, Y. Fatahi, F. Farjadian, N. Baheiraei, B. Nasseri, M. Rabiee, N.T. Dastjerd, A. Valibeik, M. Karimi, M.R. Hamblin, Stimulus-responsive sequential release systems for drug and gene delivery, Nano Today 34 (2020) 100914, https://doi.org/10.1016/j.nantod.2020.100914.
- [5] H. Jiang, C. Xia, J. Lin, H.A. Garalleh, A. Alalawi, A. Pugazhendhi, Carbon nanomaterials: a growing tool for the diagnosis and treatment of diabetes mellitus, Environ. Res. 221 (2023) 115250, https://doi.org/10.1016/j.envres.2023.115250.
- [6] N.A. Eracle, R. Lavinia, V. Monica, A non-invasive glucose analysis model with a carbon nanotube sensor, ARS Medica Tomitana 25 (2019) 189–192.
- [7] T.T. Islam, M.S. Ahmed, M. Hassanuzzaman, S.A. Bin Amir, T. Rahman, Blood glucose level regression for smartphone ppg signals using machine learning, Appl. Sci. 11 (2021) 618.
- [8] L. Singh, M. Khanna, H. Garg, D. Singh, Emperor penguin opti- mization algorithm- and bacterial foraging optimization algorithm-based novel feature selection approach for glaucoma classification from fundus images, Soft Comput. 28 (2023) 1–37, https://doi.org/10.1007/s00500-023-08449-6.
- [9] P. Naga Srinivasu, M.F. Ijaz, M. Woźniak, XAI-driven model for crop recommender system for use in precision agriculture, Comput. Intell. 40 (1) (2024) e12629, https://doi.org/10.1111/coin.12629.
- [10] H.S. Kim, I. Joe, An xai method for convolutional neural networks in self-driving cars, PLoS One 17 (2022) e0267282.
- [11] E. Susana, K. Ramli, H. Murfi, N.H. Apriantoro, Non-invasive clas- sification of blood glucose level for early detection diabetes based on photo- plethysmography signal, Information 13 (2022) 59, https://doi.org/10.3390/info13020059.
- [12] E. Susana, K. Ramli, P.D. Purnamasari, N.H. Apriantoro, Non-invasive classification of blood glucose level based on photoplethysmography using time-frequency analysis, Information 14 (2023) 145, https://doi.org/10.3390/info14030145.
- [13] O. Yildirim, M. Talo, B. Ay, U.B. Baloglu, G. Aydin, U.R. Acharya, Automated detection of diabetic subject using pre-trained 2d-cnn models with frequency spectrum images extracted from heart rate signals, Comput. Biol. Med. 113 (2019) 103387.
- [14] V. Vankadaru, P.N. Srinivasu, S.H.H. Prasad, P. Rohit, P.R. Babu, M.D.C. Raju, Text identification from handwritten data using Bi-LSTM and CNN with FastAI, in: 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA), Uttarakhand, India, 2023, pp. 215–220, https:// doi.org/10.1109/ICIDCA56705.2023.10099715.
- [15] D.C. Klonoff, Improved outcomes from diabetes monitoring: the benefits of better adherence, therapy adjustments, patient education, and telemedicine support, J. Diabetes Sci. Technol. 6 (2012) 486–490.
- [16] C. Taramasco, C. Rimassa, F. Martinez, Improvement in quality of life with use of ambient-assisted living: clinical trial with older persons in the Chilean population, Sensors 23 (2023) 268, https://doi.org/10.3390/s23010268.
- [17] Usharani Bhimavarapu, Nalini Chintalapudi, Gopi Battineni, Automatic detection and classification of diabetic retinopathy using the improved pooling function in the convolution neural network, Diagnostics 13 (2023), https://doi.org/10.3390/diagnostics13152606.
- [18] U. Bhimavarapu, G. Battineni, Deep learning for the detection and classification of diabetic retinopathy with an improved activation function, Healthcare 11 (2023) 97, https://doi.org/10.3390/healthcare11010097.
- [19] M. Khanna, L. Singh, S. Thawkar, M. Goyal, Deep learning based computer-aided automatic prediction and grading system for dia- betic retinopathy, Multimed. Tool. Appl. 82 (2023) 1–48, https://doi.org/10.1007/s11042-023-14970-5.
- [20] L. Li, Diagnosis of diabetes using a weight-adjusted voting approach, in: 2014 IEEE International Conference on Bioinformatics and Bioengineering, Boca Raton, FL, USA, 2014, pp. 320–324, https://doi.org/10.1109/BIBE.2014.27.
- [21] J.P. Kandhasamy, S. Balamurali, Performance analysis of classifier models to predict diabetes mellitus, Procedia Computer Science 47 (2015) 45–51.
- [22] S.K. Dey, A. Hossain, M.M. Rahman, Implementation of a web ap- plication to predict diabetes disease: an approach using machine learning algorithm, in: 2018 21st International Conference of Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2018, pp. 1–5, https://doi.org/10.1109/ ICCITECHN.2018.8631968, 2018.
- [23] P. Sonar, K. JayaMalini, Diabetes prediction using different machine learning approaches, in: 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2019, pp. 367–371, https://doi.org/10.1109/ICCMC.2019.8819841.
- [24] F. Beloufa, M.A. Chikh, Design of fuzzy classifier for diabetes disease using modified artificial bee colony algorithm, Comput. Methods Progr. Biomed. 112 (2013) 92–103.
- [25] S. Perveen, M. Shahbaz, T. Saba, K. Keshavjee, A. Rehman, A. Guergachi, Handling irregularly sampled longitudinal data and prognostic modeling of diabetes using machine learning technique, IEEE Access 8 (2020) 21875–21885.
- [26] A. Rehman, A. Athar, M.A. Khan, S. Abbas, A. Fatima, A. Saeed, et al., Modelling, simulation, and optimization of diabetes type ii prediction using deep extreme learning machine, J. Ambient Intell. Smart Environ. 12 (2020) 125–138.
- [27] R. Sanakal, T. Jayakumari, Prognosis of diabetes using data mining approach-fuzzy c means clustering and support vector machine, Int. J. Comput. Trends Technol. 11 (2014) 94–98.
- [28] H. Temurtas, N. Yumusak, F. Temurtas, A comparative study on diabetes disease diagnosis using neural networks, Expert Syst. Appl. 36 (2009) 8610–8615.
- [29] M.M. Hassan, Z.J. Peya, S. Mollick, M.A.M. Billah, M.M.H. Shakil, A.U. Dulla, Diabetes prediction in healthcare at early stage us- ing machine learning approach, in: 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 1–5, https://doi. org/10.1109/ICCCNT51525.2021.9579869, 2021.
- [30] P.K. Saha, N.S. Patwary, I. Ahmed, A widespread study of diabetes prediction using several machine learning techniques, in: 2019 22nd International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2019, pp. 1–5, https://doi.org/10.1109/ICCIT48885.2019.9038559.
- [31] S. Wei, X. Zhao, C. Miao, A comprehensive exploration to the machine learning techniques for diabetes identification, in: 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), Singapore, 2018, pp. 291–295, https://doi.org/10.1109/WF-IoT.2018.8355130.
- [32] Zehra Amatul, Tuty Asmawaty, Abdul Kadir, M.A.M. Aznan, A comparative study on the pre-processing and mining of pima Indian diabetes dataset, in: 3rd International Conference on Software Engineering & Computer Systems (ICSECS - 2013), 20-22 August 2013, Universiti Malaysia Pahang, 2013, pp. 1–10.
- [33] R. Bala, A. Sharma, N. Gooel, Comparative analysis of diabetic retinopathy classification approaches using machine learning and deep learning techniques, Arch. Comput. Methods Eng. 31 (2023) 1–37, https://doi.org/10.1007/s11831-023-10002-5.
- [34] B. Jeong, H. Cho, J. Kim, S.K. Kwon, S. Hong, C. Lee, T. Kim, M.S. Park, S. Hong, T.Y. Heo, Comparison between statistical models and machine learning methods on classification for highly imbalanced multiclass kidney data, Diagnostics 10 (2020), https://doi.org/10.3390/diagnostics10060415.
- [35] P.N. Srinivasu, N. Sandhya, R.H. Jhaveri, R. Raut, From blackbox to explainable ai in healthcare: existing tools and case studies, Mobile Inf. Syst. 20 (2022) 8167821, https://doi.org/10.1155/2022/8167821, 2022.
- [36] G. Yang, Q. Ye, J. Xia, Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: a mini-review, two showcases and beyond, Inf. Fusion 77 (2022) 29–52, https://doi.org/10.1016/j.inffus.2021.07.016, 77.
- [37] B.H. Van Der Velden, H.J. Kuijf, K.G. Gilhuijs, M.A. Viergever, Explainable artificial intelligence (xai) in deep learning-based medical image analysis, Med. Image Anal. 79 (2022) 102470, https://doi.org/10.1016/j.media.2022.102470.
- [38] A.R. Kulkarni, A.A. Patel, K.V. Pipal, S.G. Jaiswal, M.T. Jaisinghani, V. Thulkar, L. Gajbhiye, P. Gondane, A.B. Patel, M. Mamtani, et al., Machine-learning algorithm to non-invasively detect diabetes and pre-diabetes from electrocardiogram, BMJ Innovations 9 (2022) 32–42, 2023.
- [39] N. Ahmed, R. Ahammed, M.M. Islam, M.A. Uddin, A. Akhter, M.A.A. Talukder, B.K. Paul, Machine learning based diabetes prediction and development of smart web application, International Journal of Cognitive Computing in Engineering 2 (2021) 229–241.

- [40] S. Saha, S. Saha, A systematic review and meta-analysis of randomized controlled trials comparing the safety of dapagliflozin in type 1 diabetes patients, Iran J Health Sci 8 (2) (2020) 45–58. https://publish.kne-publishing.com/index.php/JHS/article/view/4028.
- [41] X. Shen, L. Wei, S. Tang, Dermoscopic image classification method using an ensemble of fine-tuned convolutional neural networks, Sensors 22 (2022) 4147, https://doi.org/10.3390/s22114147.
- [42] H. Liu, X. Liu, Q. Kong, W. Wang, M.D. Plumbley, Learning Temporal Resolution in Spectrogram for Audio Classification, 2024 arXiv:2210.01719.
- [43] H. Shi, Q. Sheng, Y. Wang, B. Yue, L. Chen, Dynamic range compression self-adaption method for SAR image based on deep learning, Remote Sens 14 (2022) 2338, https://doi.org/10.3390/rs14102338.
- [44] L.K. Singh, M. Khanna, H. Garg, R. Singh, Efficient feature selec- tion based novel clinical decision support system for glaucoma prediction from retinal fundus images, Med. Eng. Phys. 123 (2024) 104077, https://doi.org/10.1016/j.medengphy.2023.104077.
- [45] L.K. Singh, M. Khanna, R. Singh, Feature subset selection through nature inspired computing for efficient glaucoma classification from fundus images, Multimed Tools Appl (2024), https://doi.org/10.1007/s11042-024-18624-y.
- [46] E. Khodabandehloo, D. Riboni, A. Alimohammadi, Healthxai: col- laborative and explainable ai for supporting early diagnosis of cognitive decline, Future Generat. Comput. Syst. 116 (2021) 168–189.
- [47] M.I. Hertzog, U. Brisolara Corrêa, R.M. Araujo, SpreadOut: a kernel weight initializer for convolutional neural networks, in: 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 2019, pp. 1–7, https://doi.org/10.1109/IJCNN.2019.8852161.
- [48] M.Z. Islam, M.M. Islam, A. Asraf, A combined deep CNN-LSTM net- work for the detection of novel coronavirus (covid-19) using x-ray images, Inform. Med. Unlocked 20 (2020) 100412, https://doi.org/10.1016/j.imu.2020.100412.
- [49] J. Zhang, Y. Li, J. Tian, T. Li, LSTM-CNN hybrid model for text classification, in: 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 2018, pp. 1675–1680, https://doi.org/10.1109/IAEAC.2018.8577620.
- [50] P. Hobden, S. Srivastava, E. Nurellari, et al., Fpga-based cnn for real-time uav tracking and detection, Frontiers in Space Technologies 3 (2022) 878010.
- [51] Z. Huang, Y. Ma, R. Wang, W. Li, Y. Dai, A model for EEG-based emotion recognition: CNN-Bi-LSTM with attention mechanism, Electronics 12 (2023) 3188, https://doi.org/10.3390/electronics12143188.
- [52] K. Jindal, R. Upadhyay, P.K. Padhy, L. Longo, 6 bi-lstm-deep cnn for schizophrenia detection using msst-spectral images of eeg signals, in: V. Bajaj, G. Sinha (Eds.), Artificial Intelligence-Based Brain-Computer Interface, Academic Press, 2022, pp. 145–162, https://doi.org/10.1016/B978-0-323-91197-9.00011-4.
 [53] K.B. Aravind Britto, Saravanan Srinivasan, Sandeep Kumar Mathivanan, Muthukumaran Venkatesan, M.B. Benjula Anbu Malar, Mallik Sauray, Hong Oin.
- [53] K.R. Aravind Britto, Saravanan Srinivasan, Sandeep Kumar Mathivanan, Muthukumaran Venkatesan, M.B. Benjula Anbu Malar, Mallik Saurav, Hong Qin, A multi-dimensional hybrid CNN-BiLSTM framework for epileptic seizure detection using electroencephalogram signal scrutiny, Systems and Soft Computing 5 (2023) 200062, https://doi.org/10.1016/j.sasc.2023.200062.
- [54] M. Dinesh, N. Bacanin, S. Askar, M. Abouhawwash, Diagnostic abil- ity of deep learning in detection of pancreatic tumour, Sci. Rep. 13 (2023) 9725.
- [55] M.F. Aslan, K. Sabanci, A novel proposal for deep learning-based diabetes prediction: converting clinical data to image data, Diagnostics 13 (2023) 796, https:// doi.org/10.3390/diagnostics13040796.
- [56] O. Erkaymaz, M. Ozer, Impact of small-world network topology on the conventional artificial neural network for the diagnosis of diabetes, Chaos, Solit. Fractals 83 (2016) 178–185.
- [57] A. Ahmed, S. Aziz, U. Qidwai, A. Abd-Alrazaq, J. Sheikh, Per- formance of artificial intelligence models in estimating blood glucose level among diabetic patients using non-invasive wearable device data, Com- puter Methods and Programs in Biomedicine Update 3 (2023) 100094.
- [58] Y. Deng, L. Lu, L. Aponte Becerra, A. Angelidi, V. Novak, G. Karniadakis, C. Mantzoros, Deep transfer learning and data augmentation improve glucose levels prediction in type 2 diabetes patients, npj Digital Medicine 4 (2021) 109.
- [59] S. Usman, N. Harun, R.A. Dziyauddin, N.A. Bani, Estimation of hba1c level among diabetic patients using second derivative of photo- plethysmography, in: 2017 IEEE 15th Student Conference on Research and Development (SCOReD), vol. 40, 2017, pp. 89–92.
- [60] G. Zhang, Z. Mei, Y. Zhang, X. Ma, B. Lo, D. Chen, Y. Zhang, A noninvasive blood glucose monitoring system based on smartphone ppg signal processing and machine learning, IEEE Trans. Ind. Inf. 16 (2020) 7209–7218.
- [61] A. Keikhosravi, H. Aghajani, E. Zahedi, Discrimination of bilateral finger photoplethysmogram responses to reactive hyperemia in diabetic s healthy subjects using a differential vascular model framework, Physiol. Meas. 34 (2013) 513.
- [62] N. Nirala, R. Periyasamy, B.K. Singh, A. Kumar, Detection of type-2 diabetes using characteristics of toe photoplethysmogram by applying support vector machine, Biocybern. Biomed. Eng. 39 (2019) 38–51.
- [63] U. Rajendra Acharya, O. Faust, N. Adib Kadri, J.S. Suri, W. Yu, Automated identification of normal and diabetes heart rate signals using nonlinear measures, Comput. Biol. Med. 43 (2013) 1523–1529.
- [64] L. Wang, Y. Mu, J. Zhao, X. Wang, H. Che, IGRNet: a deep learning model for non-invasive, real-time diagnosis of prediabetes through electrocardiograms, Sensors 20 (2020) 2556, https://doi.org/10.3390/s20092556.
- [65] J. Alvarado, J.M. Velasco, F. Chavez, F.F. de Vega, J.I. Hidalgo, Combining wavelet transform with convolutional neural networks for hy- poglycemia events prediction from cgm data, Chemometr. Intell. Lab. Syst. 243 (2023) 105017, https://doi.org/10.1016/j.chemolab.2023.105017.
- [66] B. Das, A deep learning model for identification of diabetes type 2 based on nucleotide signals, Neural Comput. Appl. 34 (2022), https://doi.org/10.1007/ s00521-022-07121-8.
- [67] M.R. Rajput, S.S. Khedgikar, Diabetes prediction and analysis us- ing medical attributes: a machine learning approach, J. Xi'an Univ. Archit. Technol. 14 (2022) 98–103.
- [68] P.S. Sen, N. Mukherjee, An ontology-based approach to designing a NoSQL database for semi-structured and unstructured health data, Cluster Comput. 27 (2024) 959–976, https://doi.org/10.1007/s10586-023-03995-y.
- [69] J. Hancock, T.M. Khoshgoftaar, Impact of hyperparameter tuning in classifying highly imbalanced big data, in: 2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science (IRI), Las Vegas, NV, USA, 2021, pp. 348–354, https://doi.org/10.1109/IRI51335.2021.00054.