JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY
https://e-journal.uum.edu.my/index.php/jict

How to cite this article:
Rao, S., Kulkarni, S., & Bhat, R. (2024). Revolutionising diabetic retinopathy diagnosis with modified regularisation long short-term memory framework. *Journal of Information and Communication Technology, 23*(4), 745-776. https://doi.org/10.32890/jict2024.23.4.6

Revolutionising Diabetic Retinopathy Diagnosis with Modified Regularisation Long Short-Term Memory Framework

[1]Sudesh Rao, [2]Sanjeev Kulkarni & [3]Radhakrishna Bhat
[1]Department of AIML, NMAM Institute of Technology,
Nitte (Deemed to be University), India
[1&2]Institute of Engineering and Technology,
Srinivas University, India
[3]Department of Computer Science and Engineering,
Manipal Institute of Technology,
Manipal Academy of Higher Education, India

[1]sudesh.rao@nitte.edu.in
[2]sanjeev.d.kulkarni@gmail.com
*[3]bhat.radhakrishna@manipal.edu
*Corresponding author

Received: 31/12/2023 Revised: 20/10/2024 Accepted: 20/10/2024 Published: 28/10/2024

ABSTRACT

The diagnosis of Diabetic Retinopathy (DR) demands a paradigm shift towards more accurate and efficient solutions to overcome vision impairment. Therefore, the current study introduces a new Modified Regularisation Long Short-term Memory (MR-LSTM) framework approach for DR diagnosis. The proposed framework leverages the

power of deep learning and provides a dynamic and robust solution for the early detection of DR, which in turn preserves a patient's vision. The proposed framework uses a DR Debrecen Dataset from the UCI database with 21 distinct features relevant to retinal health, and employs a series of data preprocessing steps, including data cleaning, normalisation, and transformation, to ensure data quality and compatibility. The MR-LSTM framework excels at capturing temporal dependencies in sequential retinal images, offering a unique advantage in understanding the progression of DR. The MR-LSTM framework is implemented using Python libraries, and the results are compared with those of other popular models. It is observed that the MR-LSTM framework outperforms other models and achieves an accuracy of 97.12 percent and an F1 Score of 98.49. Furthermore, the Receiver Operating Characteristic (ROC) curve reveals an area under the curve of 0.97, highlighting the exceptional ability to discriminate between positive and negative cases of the proposed framework. By revolutionising DR diagnosis with the proposed MR-LSTM framework, it can achieve accurate, timely, and accessible solutions in the fight against vision-threatening conditions.

Keywords: Deep Learning Model, Diabetic Retinopathy, Early Detection, Healthcare Challenges, Long Short-Term Memory.

INTRODUCTION

Diabetic Retinopathy (DR), a disorder caused by microvascular changes in the retina caused by diabetes, is the leading preventable cause of blindness in people of working age worldwide. However, a complication of DR called diabetic macular edema (DME) may appear at any stage of DR and presents as a hardening of the retina or an accumulation of fluid. It is observed that 30 percent of people with diabetes suffer from DR, and 10 percent have vision-threatening DR, according to the International Council of Ophthalmology. It's the sixth most frequent cause of blindness in India and one of the severe neurovascular consequences of diabetes. Recent epidemiological data, as reported by the American Academy of Ophthalmology, paints a concerning picture of the global diabetes burden. With an estimated 387 million individuals affected by diabetes mellitus, this number is projected to surge to a staggering 592 million by the year 2035

(Abbood et al., 2022). Among these, a substantial 93 million people worldwide grapple with DR.

The prevalence of DR is notably high, affecting the majority of the individuals with type 1 diabetes and remaining with type 2. Alarmingly, a significant proportion of these cases, roughly 25 percent to 30 percent, are at risk of developing vision-threatening DME. Moreover, the clinical management of DR is a matter of utmost concern, with approximately 5 percent to 8 percent of patients necessitating laser treatment. For some, the situation is even more critical, as up to 5 percent of patients will ultimately require vitrectomy surgery to address the advanced stages of this condition. These statistics underscore the pressing need for innovative and effective diagnostic and therapeutic approaches to combat the growing global DR crisis, highlighting the potential impact of technological advancements in this field (Alahmadi, 2022). The loss of eyesight from DR can be avoided if it is diagnosed and treated early. Therefore, diabetic individuals are advised to have frequent consultations and biannual or yearly follow-ups for retina screening.

Expert physicians and the foundational healthcare infrastructure needed to treat the eye are crucial to reducing avoidable visual impairment (Liang et al., 2022; Huang et al., 2022; Nunes et al., 2021). Compared to the national ratio of one eye care expert per 1,07,000 people, this ratio varies widely across the Indian subcontinent, with certain regions having as few as one eye care expert every 6,08,000 people. Management of DR necessitates the creation of a computer-aided diagnostic tool because of the high number of persons who need constant follow-up and the scarcity of ophthalmologists. To better serve the clinical practice needs, biomedical engineers and computer scientists may use current developments in computing power, communication networks, and machine learning approaches. To create, validate, compare, and enhance DR lesion detection algorithms used in clinical applications (He et al., 2021; Zhou et al., 2021; Yang et al., 2022), the scientific community can benefit from having access to the raw pictures with ground realities. The ability to compare the effectiveness of various lesion segmentation techniques requires accurate pixel-level annotation of DR-related abnormalities such as hard exudates, soft exudates, haemorrhages and microaneurysms.

The accurate data on the severity of DR and DME lies in the assessment of image analysis and retrieval methods for early diagnosis of the

disease (Qiao et al., 2020; Li et al., 2020). Diabetes has now become a global epidemic, affecting millions of lives and straining healthcare systems. Since DR is mostly asymptomatic in its early stages, it can cause severe vision loss by gradually damaging the retinal blood vessels (Chaudhary & Pachori, 2022; Farag et al., 2022) if it is not treated promptly and precisely. Traditional procedures for diagnosing this illness are labour-intensive, lengthy, and prone to human errors, and they rely on manual examinations. Therefore, early identification and action are essential for efficient treatment (Feng et al., 2023). Recent research on artificial intelligence (AI) assisted diagnosis of DR has revealed paramount importance for earlier, more precise, and more effective identification. These deep learning (DL) and AI-assisted technologies have demonstrated superiority in revolutionising screening processes by analysing retinal images and identifying signs of DR with high precision (Kukkar et al., 2023; Niu et al., 2022).

Integrating telemedicine and remote screening has expanded access to essential eye care services and connects patients in remote or underserved areas. However, collecting extensive datasets to train AI models effectively and ensure the reliability and safety of remote diagnostic tools is a real challenge. Additionally, AI can aid only in screening. A comprehensive diagnosis and treatment plan often require the involvement of an expert ophthalmologist. Therefore, a multidisciplinary approach that combines technological advancements with clinical expertise is crucial to effectively address the issues with the complex nature of DR (Ghouali et al., 2022; Aurangzeb et al., 2023; Zhu et al., 2020). In this pursuit, the MR-LSTM framework acts as a convergence technology to mitigate the side effects of DR. This innovative model, rooted in DL, introduces a dynamic and robust approach to DR diagnosis. In this study, the transformative journey through the intricacies of DR diagnosis reveals the significant ramifications of combining modern DL methods with the unique regularisation schemes in MR-LSTM.

RELATED WORKS

Since DR is found to be the primary cause of blindness, its early identification is essential. Manually determining the degree of DR has always been complicated and lengthy for ophthalmologists. Therefore, creating a fully automated approach for DR detection in

retinal fundus pictures was crucial. Extensive studies have shown that pre-trained DL networks are significantly more effective than their untrained counterparts for image classification. For DR detection and grading tasks, optimal feature weights and parameters for a Transfer Learning (TL) approach were proposed by Wong et al. (2023). In this, features were gathered from the average pooling layers to enhance training generalisation, maximise categorisation, and feed into an ECOC ensemble configuration. They utilised the Adaptive Differential Evolution (ADE) technique to improve DR detection and grading, which involves selecting features and tweaking ensemble parameters simultaneously. Experimental findings showed that the suggested strategy outperformed baseline DL models and was competitive with previous studies.

By accurately diagnosing diabetic patients and early identification of DR, blindness could be prevented. IoMT-capable computer-aided design (CAD) systems were used by Palaniswamy and Vellingiri (2023) to recognise and categorise the severity of DR fundus images in diabetic patients who were in danger of going blind. The DR Classification (DRC) system was proposed using hybrid DL model-optimisation technique to classify DR images. This system reduced edge noise in the preprocessing stage as a preliminary step. The most critical region of the pictures was then isolated and subjected to further analysis using the recommended cluster-based K-mean segmentation technique. Extract and classify the features into one of four severity categories from the ROI, RESNet, and a Convolutional Neural Network (CNN) model, which was used with a hybrid genetic and ant colony optimisation (HGACO) approach. Diverse CNN and HGACO settings were used to assess the system's performance. While CNN made it simpler to create image features, the HGACO-based method made it easier by automating the hyperparameter-building process.

Manual diagnosis of DR required an ophthalmologist to examine photographs of the retina in a patient's eye, which was costly and time-consuming. It is even more challenging, especially in the early stages of the disease, when imaging signs are less noticeable. It has been demonstrated that medical image evaluation utilising machine learning helps evaluate retinal fundus images, and DL techniques have been utilised to aid in the early detection of DR. DL was used in Atwany et al. (2022) for the self-supervised, supervised and vision transformer configurations to categorise and detect retinal fundus

pictures. It was investigated, and a summary of DRC, including referable, non-referable, and proliferative, was provided. The paper also analysed datasets from DR retinal fundus for segmentation, categorisation, and detection tasks. Finally, the authors record the research gaps and issues that need additional study and exploration in DRC.

Protecting patient privacy and fostering collaboration are the most important considerations during DL network training. Therefore, it was necessary to develop a method of automated DR detection that would safeguard both patient information and confidentiality. In this regard, Federated Learning (FL) was used by Mohan et al. (2023) for the task of disaster severity ranking (DRFL). In fact, FL introduced a novel research approach to train DL models in a group setting without sharing sensitive clinical data. Researchers used the categorical cross-entropy loss median with the Federated averaging (FedAvg) approach to create DRFL in this work. For those clients who were overfitted or underfitted, median cross-entropy was preferred to find FedAvg. It was also advised to use a particular centralised server to gather multi-scale features from fundus images to spot any present small lesions. The proposed model outperformed existing techniques with accuracy, specificity, precision, and F1 scores of 98.6 percent, 99.3 percent, 97.5 percent, and 97.3 percent, respectively.

The current study aims to use DL techniques to categorise fundus photos into five DRC. After merging three separate DR datasets, the combined raw picture count leads to 5,819. The image preparation methods were used to clean up the images and reduce the noise before the model training. An even-keeled dataset was achieved by applying three types of augmentation: geometrical, photometric, and elastic deformation. Three blocks of convolutional layers and max pooling layers were combined to create a shallow CNN using categorical cross-entropy loss function, Adam optimiser, learning rate of 0.0001 and batch size of 64 as the starting point. This foundational model was also used to select an optimal data augmentation strategy for subsequent analysis (Raiaan et al., 2023). The experiments were conducted to enhance performance by adjusting various parts and hyperparameters of the original model.

An ensemble deep network with many streams was used by Mustafa et al. (2022) to classify the severity of DR. This study used deep

neural networks and principal component analysis (PCA) techniques to identify changes in the original visual features across classes and between classes. To obtain great accuracy in categorisation and durable effectiveness, a group of machine learning classification techniques was used to gather the deep characteristics. DenseNet-121 and ResNet-50 were the main feature extractors to construct a multi-stream network with pre-trained deep-learning frameworks. The variation space was split into intra-class and inter-class subspaces after the dimension of the feature data was reduced using PCA. Lastly, an ensemble machine learning classifier was utilised to improve categorisation precision further by merging the AdaBoost and random forest approaches. The suggested method was compared with other CNN-based methods on many different datasets. As observed in the experiments, the proposed method is highly effective, with an accuracy of up to 95.58 percent, making it a top contender for automated DR identification.

Despite all the benefits, the existing models have notable limitations that hinder their broad applicability. These models often need help with generalisation, making them less effective when applied to diverse patient populations or images from various sources. Also, the imbalanced datasets create bias, with most models performing well on the majority class but poorly on the minority class. Additionally, resource-intensive computational requirements can limit their feasibility in resource-constrained healthcare settings, particularly in low-income regions, where access to advanced hardware and infrastructure is limited. These limitations underscore the need for more robust and adaptable approaches in DR diagnosis.

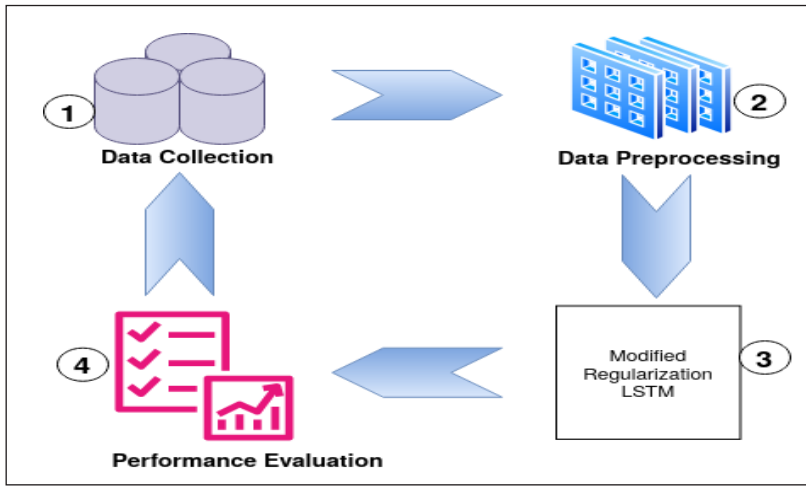
THE PROPOSED MR-LSTM FOR DR DIAGNOSIS

Figure 1 illustrates the comprehensive workflow for DR diagnosis using the MR-LSTM framework. It begins with data collection, where retinal images are gathered, followed by a critical data preprocessing phase that ensures data quality through cleaning, normalisation, and transformation. The preprocessed data is then fed into the MR-LSTM model, which incorporating modified regularisation techniques to enhance model robustness and prevent overfitting. The central LSTM unit in the framework captures temporal dependencies in sequential retinal images, crucial for detecting progressive changes associated with DR. Finally, the model's performance is evaluated using various

metrics, ensuring the accuracy and reliability of the DR diagnosis system.

Figure 1

The Proposed MR-LSTM Framework-based System



Data Collection

The DR Debrecen Dataset, available at the UCI Machine Learning Repository, is a valuable resource in medical image analysis. The primary purpose of this dataset is to create and test algorithms that utilise machine learning to categorise DR, a severe eye condition caused by diabetes (Balint, 2014). This dataset is primarily centred on the diagnosis of DR, explicitly discerning the presence or absence of this critical medical condition. It contains 1,158 instances of retinal images categorised into four stages of DR (No DR, Earlier Stage, Mid-Stage, Severe Stage), split into training and validation datasets using an 80/20 ratio. The training set, comprising 80 percent of the data (926 instances), was utilised to train machine learning models to recognise and classify various stages of DR. The remaining 20 percent of the dataset (232 instances) formed the validation set to assess the trained models' performance and ensure their generalizability to new, unseen data. This partitioning strategy ensures robust model development and evaluation, which is essential for accurate diagnosis and management of DR.

Table 1 shows the dataset features and their description. The dataset categorises data into four distinct groups, each signifying a different stage of DR: No DR denotes retinal images without any evidence of DR. The images appear healthy, devoid of characteristic signs of the condition. DR will be mild at the earlier stage, with subtle retinal changes that haven't reached mid-stage or severe diagnosis criteria. Minor abnormalities may be present, but they don't significantly threaten vision. The mid-stage category indicates DR beyond the early stage but not yet at the severe condition. More noticeable retinal changes affect significant damage, yet vision may still be manageable with timely treatment. The severe stage is the most advanced stage of DR. It involves substantial retinal damage, including widespread haemorrhages and extreme pathological changes. Therefore, vision is at risk, necessitating prompt intervention to prevent further deterioration and potential blindness.

Table 1

The Dataset Features and Their Descriptions

Feature	Description
Quality Assessment	Binary attribute indicating the quality of the image. 1 represents high-quality images, while 0 denotes lower-quality images.
Pre-screening	Binary attribute indicating whether pre-screening took place. 1 means pre-screening occurred, while 0 indicates otherwise.
MA Detection	Binary attribute for microaneurysm detection. 1 signifies the presence of microaneurysms, while 0 indicates their absence.
MA Localisation	Binary attribute for the localisation of microaneurysms. 1 represents localisation, while 0 represents the lack of localisation.
Exudates	Binary attribute indicating the presence of exudates. 1 suggests the presence of exudates, while 0 suggests their absence.
Exudates Localisation	Binary attribute for exudates localisation. 1 indicates localisation, and 0 indicates a lack of localisation.
Haemorrhages	Binary attribute for the presence of haemorrhages. 1 indicates haemorrhages, and 0 denotes their absence.

(continued)

Feature	Description
Haemorrhages Localisation	Binary attribute for haemorrhage localisation. 1 indicates localisation, while 0 indicates a lack of localisation.
Cup Disc Ratio (CDR)	The continuous attribute represents the optic cup's vertical diameter ratio to the optic disc's vertical diameter.
AM/FM Classification	Binary attribute for AM/FM classification. 1 denotes a positive result, and 0 represents a negative result.
Class Label	No DR, Earlier, Mid and Severe stages.

Data Preprocessing

Data preprocessing, including the DR Debrecen Dataset, is critical when working with datasets. The dataset must undergo various preprocessing steps in DRC to ensure its suitability for DL analysis. Data preprocessing for the DR dataset involves several crucial steps to enhance data quality and ensure compatibility with machine learning algorithms. Initially, data cleaning is performed to identify and address any missing or incomplete data points. This process starts by detecting which attributes or image instances lack values due to factors like data collection errors or equipment malfunctions during image acquisition. For instance, attributes with missing data might be imputed based on clinical relevance using mean or median imputation methods, while instances with missing images may be excluded from the analysis to maintain dataset integrity. This meticulous approach ensures the dataset is complete and reliable for subsequent analysis.

Normalisation and data transformation are then applied to standardise the dataset. In the normalisation step, each feature is centred around zero and scaled according to the standard deviation by subtracting the mean value of that feature from all data points. This is crucial for algorithms sensitive to feature scaling, such as logistic regression and support vector machines. Following normalisation, the Min-Max scaler is employed to transform the data into a specified range, typically between 0 and 1. This transformation is particularly beneficial for neural networks and k-nearest neighbour algorithms, which perform better when features have a consistent scale. Additionally, converting numerical attributes to the float32 data type optimises memory usage and computational efficiency, which is essential for handling large

datasets and ensuring smooth processing during model training. This comprehensive preprocessing pipeline enhances data quality, making the dataset robust for accurate DRC and model development.

Data Cleaning

The first step in data cleaning is identifying the missing or incomplete data points within the dataset. Missing data can result from various factors, such as data collection errors or equipment failures during image acquisition. Therefore, detecting which attributes or image instances have missing values is essential. After identifying the missing data, a decision must be made regarding how to handle these gaps. The dataset may have images or attributes with incomplete information. This could involve understanding the clinical significance of the missing information for the attribute-based missing data. Finding missing image instances may be more challenging and require careful consideration.

Normalisation Using Mean Method

Normalisation is a common data transformation step in DL, mainly when working with numerical features. The mean method applied means that each feature or attribute is scaled by subtracting the mean value of that feature from all data points. This process centres the data around zero and scales it based on the standard deviation. This step ensures that all features have a comparable scale, which is crucial for algorithms like logistic regression or support vector machines sensitive to feature scaling.

Data Transformation using Min-max Scaler

Another data transformation method is Min-max scaling. This approach scales the data to fall within a specific range, often between 0 and 1. This transformation is beneficial to ensure that the values are consistent. For the DR dataset, this may involve scaling attributes that measure the characteristics of retinal images to a standardised range. Min-max scaling can be significant when using algorithms like neural networks and k-nearest neighbours, which can benefit from feature scaling to improve convergence and model performance, as in Equation 1

$$x_{scaled} = \frac{\big(x - min(x)\big)}{\big(max(x) - min(x)\big)} \quad (1)$$

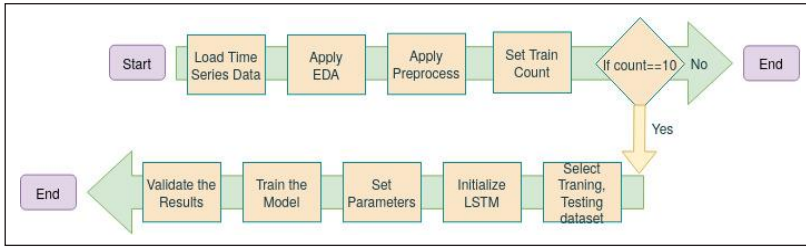
In Equation 1, x_{scaled} is a number in the range [0, 1] that represents the feature *x* after it has been rescaled. The feature's initial value, i.e., *x*, was never modified. The function $min(x)$ returns the lowest value of the feature in the dataset. The greatest value of the feature in the dataset is denoted by the function $max(x)$.

Data Type Change (Converted into Float32)

Converting data into the float32 data type typically involves representing each numerical value using 32 bits of memory. This change is often made to save memory and improve computational efficiency, particularly in resource-constrained settings or when working with large datasets. In the context of the DR dataset, this conversion is typically applied to numerical attributes that represent features of retinal images. It ensures that numerical values are represented with a level of precision suitable for most machine-learning algorithms while conserving memory resources. Figure 2 shows the flowchart of the suggested system.

Figure 2

The Proposed Flow Diagram



Modified Regularisation of Long Short-Term Memory (MR-LSTM)

The MR-LSTM framework is a revolutionary approach to DR diagnosis. It leverages DL and regularisation techniques tailored

specifically to the unique characteristics of DR data. What sets MR-LSTM apart is its ability to analyse sequential retinal images (Jaskari et al., 2022). It recognises the temporal aspect of disease progression and employs techniques like dropout and L2 regularisation to balance the model's predictive power and generalisation capacity. MR-LSTM is trained on a comprehensive dataset comprising 21 distinct features, including critical aspects of retinal health. This approach is instrumental in revolutionising DR diagnosis, as it combines the strengths of DL, data preprocessing, and regularisation, ultimately contributing to more accurate and reliable results in the early detection of DR.

Roles of LSTM

LSTM is a type of recurrent neural network (RNN) designed to address the challenge of modelling sequences and capturing long-term dependencies within sequential data. In the context of DR diagnosis, where changes in retinal health indicators over time are crucial for early detection, LSTM plays a pivotal role. LSTM units are structured to process sequences of retinal images, and each LSTM unit maintains an internal state that allows it to remember and store information about previous steps. This enables the model to capture the sequential nature of retinal health progression. Also, LSTM units are equipped with gates (e.g., input gate, forget gate, output gate) that regulate the flow of information. These gates allow LSTM to selectively remember or forget information over long sequences, which is essential for capturing subtle changes in retinal health indicators such as microaneurysms, haemorrhages, and other abnormalities associated with DR. The unique feature of LSTM is its ability to integrate past information into current predictions. By retaining and updating the internal state based on historical data points, LSTM can provide context for current retinal image analysis. This integration helps in making informed predictions about the progression of DR stages, thereby supporting early intervention and treatment planning.

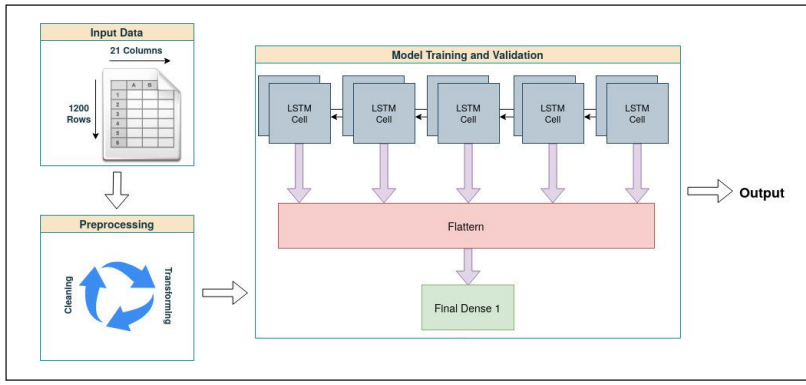
Modified Regularisation Methods

Regularisation techniques are crucial for preventing overfitting and improving the generalisation of DL models. In the context of MR-LSTM for DR diagnosis, specialised regularisation techniques such as dropout and L2 regularisation are tailored to optimise model

performance on retinal image data. The randomly selected LSTM units are ignored (i.e., their outputs are set to zero) during training in dropout. This process introduces noise and redundancy in the network, forcing it to learn more robust features and reducing the likelihood of overfitting the training data. In MR-LSTM, dropout is applied to LSTM layers to ensure the model does not rely too heavily on specific sequences or patterns in the retinal images. Additionally, L2 regularisation imposes a penalty on the squared magnitude of weights in LSTM layers. Adding this regularisation term to the loss function during training, L2 regularisation encourages the model to learn simpler patterns and reduces the variance in weight values across the network. This constraint promotes model stability and helps prevent the amplification of noise or irrelevant features in retinal image data.

Figure 3

The Architecture of the Proposed System



Fine-tuning the dropout rates, regularisation strengths (for L2 regularisation), and other hyperparameters specific to the MR-LSTM framework. The model can achieve optimal diagnostic accuracy and reliability by adapting these parameters to the characteristics of retinal health indicators and the complexity of DR progression. The combination of enhanced temporal modelling through LSTM architecture and specialised regularisation techniques in the MR-LSTM framework represents a significant advancement in DR diagnosis. By capturing temporal dependencies and integrating past information, MR-LSTM enables accurate tracking of retinal health indicators over time, which is essential for early detection

and intervention in DR. Moreover, employing dropout, L2 regularisation, and customised optimisation, the framework enhances model robustness and generalisation capabilities, ensuring reliable performance across diverse datasets and clinical scenarios. This integrated approach underscores the potential of DL in revolutionising healthcare practices, offering new avenues for precision medicine and improved patient outcomes in DR management. Figure 3 shows the detailed LSTM architecture of the proposed method.

The normalisation techniques used by MR-LSTM, a unique approach for automating the diagnosis of DR, increase the precision and robustness of the LSTM-based model. LSTM, a recurrent neural network, excels at processing consecutive information. In medical image analysis, this is very helpful for disorders like DR. LSTM neural networks are superior to conventional neural networks for simulating recurring patterns and temporal transitions (Ali et al., 2023). In the context of DR, LSTMs can be used to analyse sequences of retinal images taken over time from the same patient. These sequences can reveal how the disease is progressing and help in the early detection of abnormalities. LSTMs can also be employed to process image data in a way that considers the spatial and temporal relationships between features, which is crucial for identifying subtle changes in the retina. The normalisation techniques used by MR-LSTM represent a unique approach in DR diagnosis due to their tailored application to sequential retinal image data.

Unlike conventional normalisation methods that standardise data across static features, MR-LSTM integrates normalisation within its LSTM architecture to adaptively scale and transform temporal sequences of retinal images. This adaptive normalisation process ensures that the model can effectively handle variations in retinal image characteristics over time, such as changes in brightness, contrast, and overall image quality, which are critical for accurate DR detection. By dynamically adjusting to the evolving nature of retinal health indicators, MR-LSTM enhances model performance and diagnostic precision, advancing the capability to detect early signs of DR and supporting timely clinical interventions. LSTM cell input and activation Equations 2 - 7:

Input Gate (i_t):

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (2)$$

Forget Gate (f_t):

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \quad (3)$$

Cell State Update (g_t):

$$g_t = tanh\left(W_g x_t + U_g h_{t-1} + b_g\right) \quad (4)$$

Output Gate (o_t):

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (5)$$

Cell State (c_t):

$$c_t = f_t * c_{t-1} + i_t * g_t \quad (6)$$

Hidden State (h_t):

$$h_t = o_t * tanh(c_t) \quad (7)$$

where x_t = input at time t, h_{t-1} = hidden state at time (t-1). An input gate, forget gate, cell state update, and output gate activation are denoted by i_t, f_t, g_t and o_t respectively. Cell state at time t is denoted by c_t. The secret condition at time t is denoted by h_t. To indicate a sigmoid activation function, use the symbol. In this context, $tanh$ stands for the hyperbolic tangent activation function. The weight matrices for the input, forget, cell state update, and output gates are represented as W_i, W_f, W_g, W_o. Input, forget, cell state update, and output gates all have unique recurrent weight matrices denoted by U_i, U_f, U_g, U_o. b_i, b_f The bias values for the input, forget, cell state update, and output gates are b_g and b_o.

Regularisation Techniques for Enhanced Performance

Overfitting in machine learning algorithms is avoided via regularisation techniques. When a model gets too complicated and matches the training data too closely, it is said to be overfitted. This happens when the model cannot generalise successfully to new, untried data. Regularisation methods include batch normalisation, dropout, and L1 and L2 regularisation. In the context of DL and DR diagnosis, the choice of regularisation technique is of paramount importance. Applying regularisation methods to an LSTM-based model can lead to a more robust and accurate diagnostic tool. This is especially relevant given the complexity of the data and the potential for the model to overfit.

Automating the diagnosis of DR through the utilisation of LSTM neural networks represents a novel and progressive approach. What sets this concept apart is the integration of specialised regularisation techniques designed to elevate the model's overall performance and robustness. This innovative framework commences with collecting retinal data from diabetic patients, often obtained through medical imaging equipment. Subsequently, the data undergoes a series of preprocessing stages, encompassing data cleaning, the management of missing values, and the preparation of the dataset into a format conducive to DL. Ensuring meticulous preprocessing is essential for the model's success, as the quality and consistency of the data have a profound impact on the outcome. In the context of supervised learning, labelled data is indispensable for model training. Here, the pivotal role of ophthalmologists and medical experts comes to the fore as they are tasked with labelling retinal images according to the severity of DR. These labels serve as the bedrock of truth for the model, facilitating its learning process in making precise diagnoses. At the core of this framework lies the LSTM architecture, which is adept at processing data sequences, making it tailor-made for analysing sequential retinal images. Intricate patterns and temporal relationships within the information are carefully captured by LSTM layers, providing a key to understanding the evolution of DR.

However, implementing specific regularisation techniques that are thoughtfully tailored to the unique nature of DR data distinguishes this concept from conventional LSTM models. These methods include L2 regularisation, which adds a penalty term to the loss function in line with the size of the model variables, and dropout, which adds a penalty term to the loss function to supplement the loss function and prevent overfitting by randomly deactivating a subset of neurons during training. Integrating these regularisation methods works synergistically to ensure that the LSTM model not only generalises effectively to unseen data but also maintains a balanced level of complexity. Model training unfolds as the model is exposed to a subset of the labelled data, and its performance is subsequently assessed using a validation dataset that remains unseen throughout the training process. The training journey involves iterative parameter adjustments to minimise the disparity between the model's predictions and the ground truth labels. This allows the model to refine its predictive accuracy.

Utilising a validation dataset at this stage critically evaluates the model's aptitude for generalising its knowledge to entirely new data. Fine-tuning of hyperparameters is an integral facet of this framework, involving the meticulous calibration of parameters such as learning rates and the extent of regularisation. Hyperparameter tuning represents an iterative endeavour to identify the optimal settings that culminate in the model's peak performance. By incorporating LSTM neural networks fortified with specialised regularisation techniques, this framework holds the potential to significantly augment diagnostic accuracy, increase efficiency, and facilitate early intervention in the progression of DR. Nonetheless; it is essential to acknowledge the challenges associated with data quality, ethical considerations, clinical validation, model interpretability, and resource constraints that must be thoughtfully addressed to realise this innovative concept's full potential. Regularisation terms can be added to the bias and weight matrices to prevent overfitting. For example, L2 regularisation can be applied as follows:

L2 Regularisation term for weight matrices from Equations 8 – 11:

$$L2(W_i) = 0.01 * \left||W_i|\right|^2 \quad (8)$$

$$L2(W_f) = 0.01 * \left||W_f|\right|^2 \quad (9)$$

$$L2(W_g) = 0.01 * \left||W_g|\right|^2 \quad (10)$$

$$L2(W_o) = 0.01 * \left||W_o|\right|^2 \quad (11)$$

L2 Regularisation term for recurrent weight matrices from 12 – 15:

$$L2(U_i) = 0.01 * \left||U_i|\right|^2 \quad (12)$$

$$L2(U_f) = 0.01 * \left||U_f|\right|^2 \quad (13)$$

$$L2(U_g) = 0.01 * \left||U_g|\right|^2 \quad (14)$$

$$L2(U_o) = 0.01 * \left||U_o|\right|^2 \quad (15)$$

Regularisation terms are added to the weight and recurrent weight matrices to prevent overfitting and enhance the model's generalisation capabilities. These equations can be implemented in a DL framework with appropriate data and training procedures.W_i, W_f, W_g, W_o are weight matrices and U_i, U_f, U_g, U_o are the recurrent weight matrices of L2 regularisation. L4 is the fourth layer in the MR-LSTM

framework. It comprises 125 units, each undergoing the LSTM process without additional regularisation. This layer and other LSTM layers (L1, L2, and L3) contribute to the fully connected layer (LFC) that combines their outputs before final model training.

Algorithm 1: Modified Regularisation LSTM

```
1.Initialize Sequence Lan = 4, class:4        // Define the parameters
2.Initialize Regularisation Bias = [4 = 0.0, L2 = 0.01]
3.Initialize Regularisation Weight = [4 = 0.01, L2 = 0.0]
    For i in range (0, seq, Lan)                  // Layer 1
      If i == 0
          //Initial process
      Else if i == 1
        Unit = 150, prop-out = 0.2
        for j in range (0, 150)
        L1 = LSTM process with regularisation (Regularisation Bias)
        end for
      Else if i == 2
        Unit = 150, prop-out = 0.2
       for j in range (0, 150)
        L2 = LSTM process
        end for
      Else if i == 3
        Unit = 150, prop-out = 0.2
        for j in range (0, 150)
        L3 = LSTM process with regularisation (Regularisation Bias)
        end for
  elseif i == 4
        for j in range (0, 125)
        L4 = LSTM process
        end
4.LFC = L1 + L2 + L3 +L4                      // Fully Connect Layer
5.Model = Train (optimisation, Train data, Label)
```

Novelty of the Proposed Framework

MR-LSTM introduces regularisation techniques tailored to the unique characteristics of DR datasets. By applying dropout and L2 regularisation, the model maintains a delicate balance between predictive power and robust generalisation. This targeted regularisation approach contributes significantly to the model's reliability and effectiveness. Unlike traditional techniques that treat retinal images as isolated snapshots, this work recognises the sequential nature of

the data. MR-LSTM is designed to analyse retinal data in the context of their temporal evolution, allowing it to capture subtle changes and patterns indicative of DR's progression. This sequential data analysis is a groundbreaking approach in DR diagnosis. MR-LSTM integrates the power of DL with the finesse of specialised regularisation. This unique blend ensures that the model can effectively identify intricate patterns in the retinal images, a critical aspect in the early detection of DR.

RESULTS AND DISCUSSIONS

The proposed framework's key contributions are as follows: i) MR-LSTM leverages the power of DL, particularly LSTM networks, to analyse sequential retinal images. ii) The framework uses regularisation techniques such as dropout to enhance model reliability and performance. iii) The dataset features cover various aspects, including quality assessment, pre-screening information, the critical class label, etc., to denote DR's presence or absence.

It was found that MR-LSTM architecture helps to forecast the severity of DR. In implementing the proposed methodology, the system configuration was pivotal in facilitating the smooth execution of complex computational tasks and data processing. The operating system, Windows 10, provided a stable and user-friendly environment for software development and experimentation. With 16GB of RAM and a powerful Core i5 processor, the system demonstrated robust computational capabilities, ensuring efficient multitasking and handling of resource-intensive applications. 1TB hard disk storage capacity allowed for storing large datasets and model parameters, which are essential for DL tasks. This system configuration served as a reliable foundation, enabling the current research to explore and develop advanced algorithms and models responsively and efficiently. Overall, system model efficiency is assessed through the accuracy metric in Equation 16.

$$A = T_{pos} + \frac{T_{neg}}{T_{pos}+T_{neg} + F_{pos}+ F_{neg}} \quad (16)$$

where A=accuracy, T_{pos}= true positive, T_{neg}=true negative, F_{pos}= false positive and F_{neg}=false negative. Precision (P) also characterises both the similarity and correctness of multiple estimates. It is given in Equation 17. The correlation between accuracy and precision

underscores how perspectives can often vary. Recall (R) is the percentage of all relevant findings correctly classified through the procedures. The true positives in Equation 18 are determined by dividing genuine positivity by falsely negative values. The calculation of the F1-Score combines both recall and accuracy. To compute the F1-Score, use Equation 19, which divides recall by accuracy.

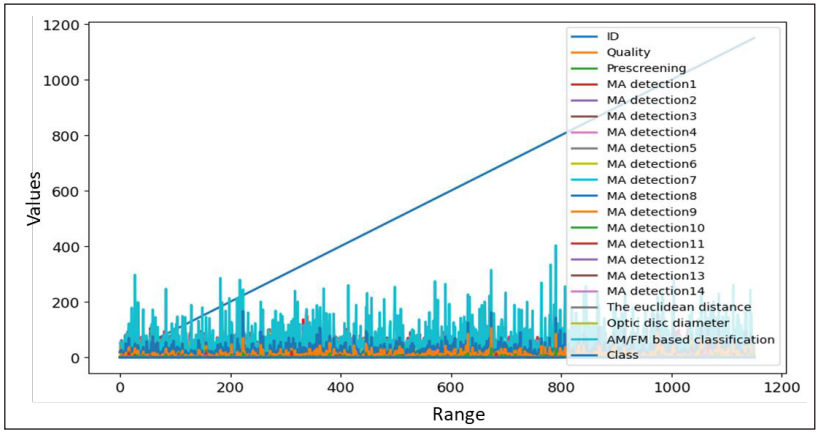
$$P = \frac{T_{pos}}{T_{pos}+F_{pos}} \tag{17}$$

$$R = \frac{T_{pos}}{T_{pos}+F_{neg}} \tag{18}$$

$$F1\ score = \frac{2\times P\times R}{P+R} \tag{19}$$

Figure 4

The Dataset Exploration



The UCI dataset employed for detecting DR, shown in Figure 4, comprises 21 distinctive features, each of which has a critical role in the diagnostic process. These meticulously selected features offer diverse insights into retinal health and pathology. Quality Assessment, Pre-screening, MA Detection1, MA Detection2, MA Detection3, MA Detection4, MA Detection5, MA Detection6, MA Detection7, MA Detection8, MA Detection9, MA Detection10, MA Detection11, MA Detection12, MA Detection13, MA Detection14, Euclidean Distance, Cup Disc Ratio (CDR), AM/FM Classification, and Class Label Features were used. The amalgamation of these 21 features provides a holistic and comprehensive set of data essential for accurately

detecting DR. This dataset serves as a vital resource for developing predictive models to facilitate early diagnosis and intervention in DR cases, ultimately contributing to improved patient outcomes and preserving visual health.

A correlation matrix is a valuable statistical tool used to analyse and visualise the relationships between multiple variables or features within the dataset. In the context of the features of the above-mentioned dataset, a correlation matrix would show the pair-wise correlations between them, quantifying how they relate. In the correlation matrix, as shown in Figure 5, numerical values indicate the strength and direction of the correlations. Coefficients near to 1 indicate a positive relationship between the two parameters. When the value of the correlation coefficient is negative (near -1), it indicates a negative impact. Close to zero indicates little to no linear relationship between the parameters. This information is crucial for feature selection, model building, and understanding the underlying relationships between the features in DR detection.

Figure 5

The Correlation Matrix

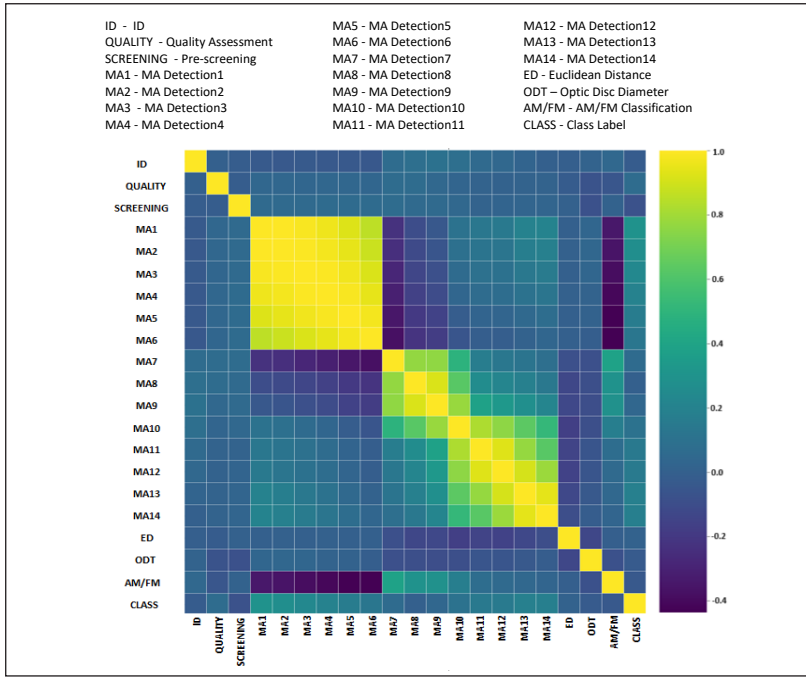
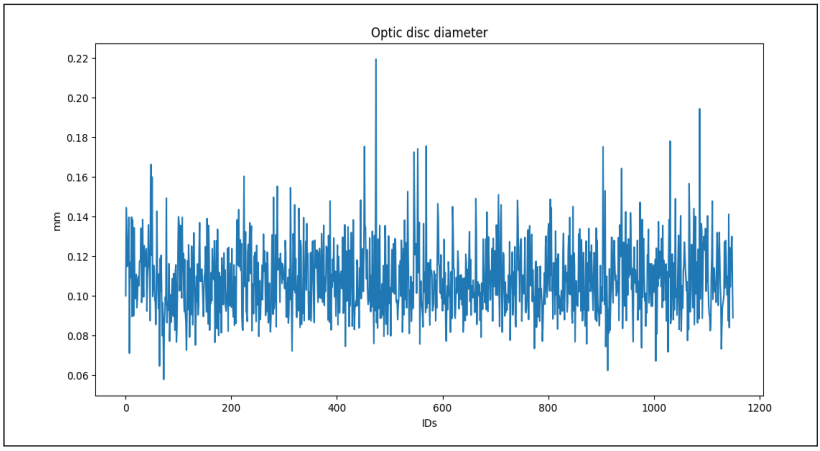


Figure 6 displays the Optic Disc of various peoples from the datasets. The optic nerve head, also called the optic disc, is a critical structure in the human eye, where the optic nerve enters the retina. It is a vital conduit for transmitting visual information from the retina to the brain. In this context, showcasing the optic disc likely indicates a comparative analysis or a visual representation of the diversity or variability in optic disc characteristics observed in different datasets as shown in Figure 6. Graphical representations are invaluable for training machine learning models or clinicians to make informed decisions about diagnosis and treatment strategies based on the observed variations in optic disc characteristics across different datasets.

Figure 6

The Optic Disc Diameter of Various Data



After completing the data preprocessing steps, including data cleaning to handle missing data, normalisation using the mean method, and data transformation with Min-max scaling, the dataset is now in a suitable format for model training. The next step is to apply the MR-LSTM model for training. MR-LSTM is a specialised neural network architecture for sequence data, making it particularly well-suited for tasks like DR detection. Dividing datasets into training and test sets using an 80:20 ratio is a well-established machine-learning model development practice. This deliberate ratio allocation is chosen to strike a delicate equilibrium, optimising data utilisation for training while reserving a distinct segment to assess the model's performance.

This data split procedure unfolds in the following manner: The training set, constituting 80 percent of the dataset, assumes a pivotal role in the training process. Within this space, the MR-LSTM model is diligently nurtured and fine-tuned. As the model traverses the data, it adapts its parameters to decipher intricate patterns and relationships inherent in the dataset. This substantial training segment is instrumental in preparing the model with the knowledge and capabilities required for accurate predictions.

The remaining 20 percent of the dataset is devoted to the test set, which serves as a sanctum of independent evaluation. Distinct from the training data, the test set has remained unseen by the model during its training phase. It provides a litmus test for the model's real-world prowess, illuminating its ability to generalise to fresh, previously unencountered instances. The test set is indispensable in unveiling the model's effectiveness in practical applications by offering an objective arena for model assessment. The 80:20 ratio thoughtfully employed in this process epitomises the balance between thorough training and robust evaluation. It guards against the model being overly simplistic (underfit) or excessively complex (overfit), ensuring it can deliver precise predictions when confronted with novel, unseen data. This meticulous approach underpins the development of a steadfast and dependable diagnostic tool for the intricate domain of DR. During model training; the prepared dataset will teach the MR-LSTM model to recognise patterns and relationships within the data indicative of DR. The model will be fine-tuned to optimise its performance and ensure it can generalise well to unseen data. Figures 7, 8 and 9 thoroughly compare several machine learning and DL models, measuring their performance across various crucial metrics for detecting DR and making medical diagnoses. Accuracy, precision, recall (sensitivity), specificity, and the F1 Score are some of these indicators, and each one provides different insights into a model's efficiency.

Support Vector Machine (SVM) demonstrates an accuracy of 70.71 percent, signifying its ability to correctly classify a substantial portion of dataset instances (Mishra et al., 2020). While it exhibits a precision of 78.57 percent and a recall of 79.71 percent, indicating robust performance in accurate positive predictions and capturing true positives, its specificity of 50.00 percent reveals challenges in correctly identifying true negatives. The F1 Score of 79.14 percent highlights a

well-balanced trade-off between precision and recall. Random Forest (RF) with an accuracy of 73.27 percent, showcases slight improvement over SVM. It achieves a precision of 81.08 percent and a recall of 82.19 percent, indicating better positive predictions and true positive capture. However, like SVM, its specificity is 50.00 percent. The F1 Score of 81.64 percent emphasises a balanced performance.

XG Boost enhances accuracy to 79.63 percent with a precision of 86.21 percent and a recall of 88.24 percent, signifying strong positive predictions and true positive capture. At 47.83 percent, specificity suggests room for improvement in true negative identification. The F1 Score of 87.22 percent reflects a well-balanced performance. Decision Tree (DT) leads with an accuracy of 85.45 percent. It exhibits a high precision of 90.43 percent and an impressive recall of 92.39, emphasising its ability to make accurate positive predictions and capture true positives. Like previous models, specificity remains at 50.00 percent. The F1 Score of 91.41 percent showcases a well-balanced performance. LSTM excels with a high accuracy of 91.51 percent. It demonstrates a precision of 95.79 percent and a recall of 94.79 percent, indicating minimal false positives and strong true positive capture. Its specificity of 60.00 percent is relatively higher. The F1 Score of 95.29 percent reflects a robust balance between precision and recall. Bidirectional Long Short-term Memory (BI-LSTM) mirrors LSTM's accuracy of 91.51 percent. It boasts a precision of 95.88 percent and a recall of 94.90 percent, similar to LSTM. However, its specificity is at 50.00 percent.

The F1 Score of 95.39 percent demonstrates a strong balance between precision and recall. MR-LSTM leads with an impressive accuracy of 97.12 percent. It excels in precision with 98.99 percent and achieves a recall of 98.00 percent, signifying an impressive ability to make accurate positive predictions and capture true positives. Its specificity significantly improves to 75.00 percent. The F1 Score of 98.49 percent marks an exceptional balance between precision and recall. In summary, MR-LSTM is the top-performing model, exhibiting remarkable accuracy, precision, and recall. Accurately predicting positive and negative cases makes it an outstanding diagnostic tool for DR. These metrics hold great significance in healthcare applications, where correctly identifying cases of the condition while minimising false positives and negatives is paramount.

Figure 7

The Accuracy and Precision Comparisons

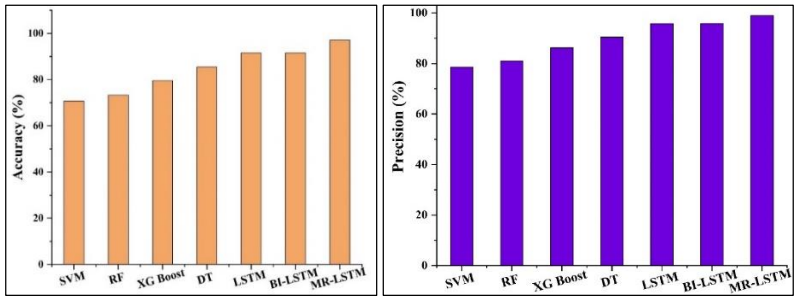


Figure 8

The Sensitivity and Specificity Comparisons

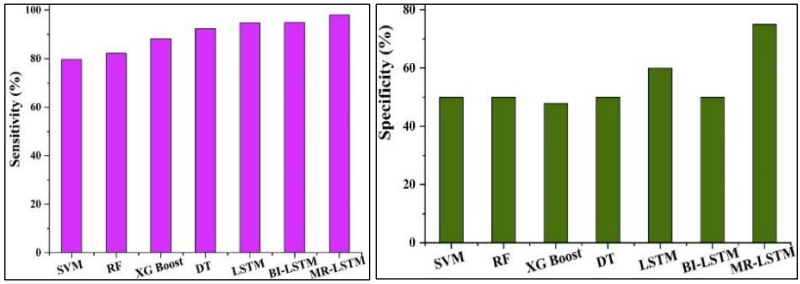
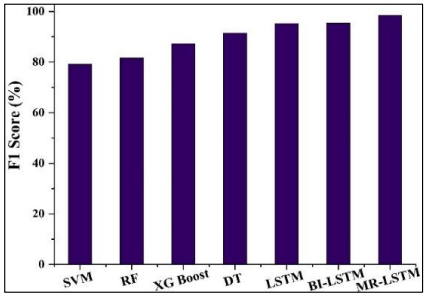


Figure 9

The F1 Score Comparison

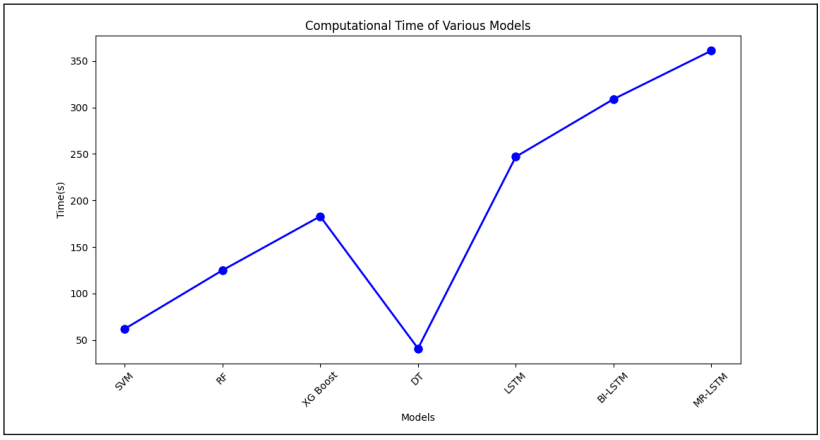


A thorough breakdown of the computation durations (measured in seconds) needed to train various machine learning and DL models can

be found in Figure 10. These time figures give essential information about each model's effectiveness in computing power and processing speed. SVM exhibits a relatively short training time of 62 seconds, making it a time-efficient choice for classification tasks. RF, known for its ensemble learning capabilities, takes 125 seconds for training, indicating slightly more computational demands than SVM. With its gradient boosting techniques, the XG Boost model requires 183 seconds for training, offering a balance between training time and predictive performance. Decision Tree, a simple yet effective model, has a training time of 41 seconds, showcasing its computational efficiency. LSTM, a DL model capable of processing sequential data, takes 247 seconds for training, indicating moderate computational demand. BI-LSTM, an extension of LSTM, requires 309 seconds for training, which is longer but still reasonable for a DL model. MR-LSTM, a specialised model, exhibits a training time of 361 seconds, signifying a slightly longer but manageable duration for model training.

Figure 10

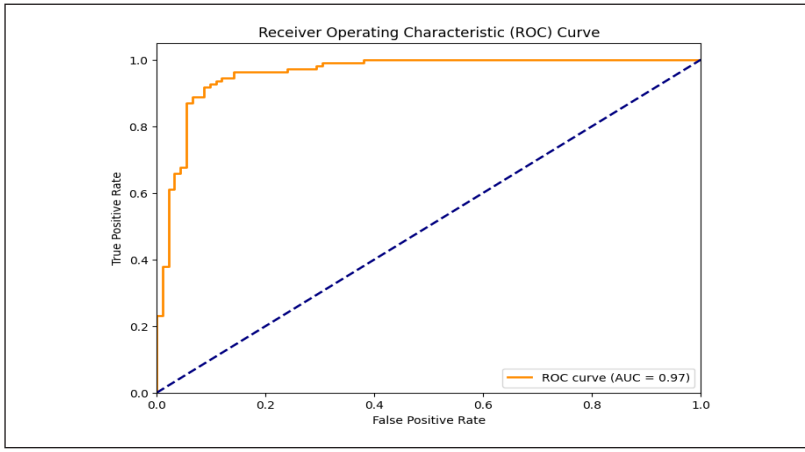
The Computational Time Comparison



The Receiver Operating Characteristic (ROC) curve shown in Figure 11 is a crucial visual assessment tool for assessing the performance of binary classification models, such as MR-LSTM, in DR. Figure 11 shows a clear distinction between positive and negative cases across a spectrum of classification thresholds. The ROC curve is plotted as a false positive rate (1 - Specificity) vs true positive rate (Sensitivity).

Figure 11

The ROC Curve of the Proposed Model



Sensitivity measures the model's capability to correctly identify true positive cases, while (1 – Specificity) quantifies the model's tendency to incorrectly label true negative cases as false positives. The ROC curve offers a dynamic view of the model's performance by examining how these rates change at different decision thresholds. The notable feature of an ROC curve is its AUC, a scalar value ranging from 0 to 1 that summarises the model's discriminative power. An AUC of 0.97, as observed in the case of MR-LSTM, signifies a highly effective model. This value is close to the maximum achievable AUC of 1, indicating that MR-LSTM distinguishes DR cases from non-cases. Practically speaking, an AUC of 0.97 indicates that MR-LSTM demonstrates an incredibly high likelihood of correctly rating a randomly chosen positive example higher than a randomly chosen negative case when it is used to categorise new, previously unknown occurrences. This underscores the robustness and reliability of MR-LSTM in DR detection, positioning it as a potent and trustworthy tool for this critical medical task. Such outstanding AUC values reinforce the model's potential to substantially impact healthcare applications by enhancing diagnostic accuracy and patient outcomes.

CONCLUSION AND FUTURE WORK

The proposed MR-LSTM framework represents a significant advancement in DR diagnosis. This work has successfully harnessed

the power of DL, temporal analysis, and specialised regularisation to create an accurate model that can generalise to new and diverse patient populations. Results show that the MR-LSTM framework is highly effective, with superior accuracy, precision, recall, and F1 score. The MR-LSTM framework performs substantially better than its competitors when evaluating the accuracy of different DL models for DR identification. The model with the best performance is MR-LSTM, which obtains an exceptional accuracy rate of 97.12 percent. LSTM and BI-LSTM share an accuracy ranking of 91.51 percent. These models exhibit remarkable performance, making them strong contenders in DR diagnosis. However, LSTM-based methods consume more computational resources for large datasets. Therefore, future work in the context of DR diagnosis may involve the development of a hybrid ensemble and optimisation models. The hybrid ensemble model will involve integrating different DL algorithms. To increase the precision and resilience of the diagnostic system, the optimisation model may be used to adjust hyperparameters, pick the best feature selection techniques, or apply specific optimisation.

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Abbood, S. H., Hamed, H. N. A., Rahim, M. S. M., Rehman, A., Saba, T., & Bahaj, S. A. (2022). Hybrid retinal image enhancement algorithm for diabetic retinopathy diagnostic using deep learning model. *IEEE Access*, *10*, 73079–73086. https://doi.org/10.1109/ACCESS.2022.3189374

Alahmadi, M. D. (2022). Texture attention network for diabetic retinopathy classification. *IEEE Access*, *10*, 55522–55532. https://doi.org/10.1109/ACCESS.2022.3177651

Ali, G., Dastgir, A., Iqbal, M. W., Anwar, M., & Faheem, M. (2023). A hybrid convolutional neural network model for automatic diabetic retinopathy classification from fundus images. *IEEE Journal of Translational Engineering in Health and Medicine*, *11*, 341–350. https://doi.org/10.1109/JTEHM.2023.3282104

Atwany, M. Z., Sahyoun, A. H., & Yaqub, M. (2022). Deep learning techniques for diabetic retinopathy classification: A survey. *IEEE Access*, *10*, 28642–28655. https://doi.org/10.1109/ACCESS.2022.3157632

Aurangzeb, K., Alharthi, R. S., Haider, S. I., & Alhussein, M. (2023). Systematic development of AI-enabled diagnostic systems for glaucoma and diabetic retinopathy. *IEEE Access*, *11*, 105069–105081. https://doi.org/10.1109/ACCESS.2023.3317348

Balint, A. H. (2014). *DR Debrecen* [dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C5XP4P

Chaudhary, P. K., & Pachori, R. B. (2022). Automatic diagnosis of different grades of diabetic retinopathy and diabetic macular edema using 2-D-FBSE-FAWT. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1–9. https://doi.org/10.1109/TIM.2022.3140437

Farag, M. M., Fouad, M., & Abdel-Hamid, A. T. (2022). Automatic severity classification of diabetic retinopathy based on DenseNet and convolutional block attention module. *IEEE Access*, *10*, 38299–38308. https://doi.org/10.1109/ACCESS.2022.3165193

Feng, M., Wang, J., Wen, K., & Sun, J. (2023). Grading of diabetic retinopathy images based on graph neural network. *IEEE Access*, *11*, 98391–98401. https://doi.org/10.1109/ACCESS.2023.3312709

Ghouali, S., Onyema, EM., Guellil, MS., Wajid, M. A., Clare, O., Cherifi, W., & Feham, M. (2022). Artificial intelligence-based teleopthalmology application for diagnosis of diabetics retinopathy. *IEEE Open Journal of Engineering in Medicine and Biology*, *3*, 124–133. https://doi.org/10.1109/OJEMB.2022.3192780

He, A., Li, T., Li, N., Wang, K., & Fu, H. (2021). CABNet: Category attention block for imbalanced diabetic retinopathy grading. *IEEE Transactions on Medical Imaging*, *40*(1), 143–153. https://doi.org/10.1109/TMI.2020.3023463

Huang, S., Li, J., Xiao, Y., Shen, N., & Xu, T. (2022). RTNet: Relation transformer network for diabetic retinopathy multi-lesion segmentation. *IEEE Transactions on Medical Imaging*, *41*(6), 1596–1607. https://doi.org/10.1109/TMI.2022.3143833

Jaskari, J., Sahlsten, J., Damoulas, T., Knoblauch, J., Särkkä, S., Kärkkäinen, L., Hietala, K., & Kaski, K. K. (2022). Uncertainty-aware deep learning methods for robust diabetic retinopathy Classification. *IEEE Access*, *10*, 76669–76681. https://doi.org/10.1109/ACCESS.2022.3192024

Kukkar, A., Gupta, D., Beram, S. M., Soni, M., Singh, N. K., Sharma, A., Neware, R., Shabaz, M., & Rizwan, A. (2023). Optimising Deep learning model larameters using socially implemented IoMT systems for diabetic retinopathy classification problem. *IEEE Transactions on Computational Social Systems*, *10*(4), 1654–1665. https://doi.org/10.1109/TCSS.2022.3213369

Li, X., Hu, X., Yu, L., Zhu, L., Fu, C. W., & Heng, P. A. (2020). CANet: Cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading. *IEEE Transactions on Medical Imaging*, *39*(5), 1483–1493. https://doi.org/10.1109/TMI.2019.2951844

Liang, N., Yuan, L., Wen, X., Xu, H., & Wang, J. (2022). End-to-end retina image synthesis based on CGAN using class feature loss and improved retinal detail loss. *IEEE Access*, *10*, 83125–83137. https://doi.org/10.1109/ACCESS.2022.3196377

Mishra, S., Hanchate, S., & Saquib, Z. (2020, October). Diabetic retinopathy detection using deep learning. In *2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)* (pp. 515-520). IEEE.

Mohan, N. J., Murugan, R., Goel, T., & Roy, P. (2023). DRFL: Federated learning in diabetic retinopathy grading using fundus images. *IEEE Transactions on Parallel and Distributed Systems*, *34*(6), 1789–1801. https://doi.org/10.1109/TPDS.2023.3264473

Mustafa, H., Ali, S. F., Bilal, M., & Hanif, M. S. (2022). Multi-stream deep neural network for diabetic retinopathy severity classification under a boosting framework. *IEEE Access*, *10*, 113172–113183. https://doi.org/10.1109/ACCESS.2022.3217216

Niu, Y., Gu, L., Zhao, Y., & Lu, F. (2022). Explainable diabetic retinopathy detection and retinal Image generation. *IEEE Journal of Biomedical and Health Informatics*, *26*(1), 44–55. https://doi.org/10.1109/JBHI.2021.3110593

Nunes, F., Madureira, P., Rêgo, S., Braga, C., Moutinho, R., Oliveira, T., & Soares, F. (2021). A mobile tele-ophthalmology system for planned and opportunistic screening of diabetic retinopathy in primary care. *IEEE Access*, *9*, 83740–83750. https://doi.org/10.1109/ACCESS.2021.3085404

Palaniswamy, T., & Vellingiri, M. (2023). Internet of things and deep learning enabled diabetic retinopathy diagnosis using retinal fundus images. *IEEE Access*, *11*, 27590–27601. https://doi.org/10.1109/ACCESS.2023.3257988

Qiao, L., Zhu, Y., & Zhou, H. (2020). Diabetic retinopathy detection using prognosis of icroaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms. *IEEE Access*, *8*, 104292–104302. https://doi.org/10.1109/ACCESS.2020.2993937

Raiaan, M. A. K., Fatema, K., Khan, I. U., Azam, S., Rashid, Md. R. U., Mukta, Md. S. H., Jonkman, M., & De Boer, F. (2023). A lightweight robust deep learning model gained high accuracy in classifying a wide range of diabetic retinopathy images. *IEEE Access*, *11*, 42361–42388. https://doi.org/10.1109/ACCESS.2023.3272228

Wong, W. K., Juwono, F. H., & Apriono, C. (2023). Diabetic retinopathy detection and grading: A transfer learning approach using simultaneous parameter optimisation and feature-weighted ECOC ensemble. *IEEE Access*, *11*, 83004–83016. https://doi.org/10.1109/ACCESS.2023.3301618

Yang, Y., Shang, F., Wu, B., Yang, D., Wang, L., Xu, Y., Zhang, W., & Zhang, T. (2022). Robust collaborative learning of patch-level and image-level annotations for diabetic retinopathy grading from fundus image. *IEEE Transactions on Cybernetics*, *52*(11), 11407–11417. https://doi.org/10.1109/TCYB.2021.3062638

Zhou, Y., Wang, B., Huang, L., Cui, S., & Shao, L. (2021). A benchmark for studying diabetic retinopathy: Segmentation, grading, and transferability. *IEEE Transactions on Medical Imaging*, *40*(3), 818–828. https://doi.org/10.1109/TMI.2020.3037771

Zhu, S., Liu, H., Du, R., Annick, D. S., Chen, S., & Qian, W. (2020). Tortuosity of retinal main and branching arterioles, venules in patients with type 2 diabetes and diabetic retinopathy in China. *IEEE Access*, *8*, 6201–6208. https://doi.org/10.1109/ACCESS.2019.2963748