A Deep Learning Approach for Diabetic Retinopathy and Cataract Detection

Zahraa H. Ali[1*], Adala M. Chaid[2], Zaid Ameen Abduljabbard[3], Vincent Omollo Nyangaresie[4] and Ali Hasan Ali[5]

[1]Department of Geology, College of Science, University of Basrah, Basra, Iraq.
[2]Department of Computer Information Systems, College of Computer Science & Information Technology, University of Basrah, Iraq
[3]Department of Computer Science, College of Education for Pure Sciences, University of Basrah, Iraq
[4]Department of Computer Science and Software Engineering, Jaramogi Oginga Odinga University of Science & Technology, Kenya
[5]Department of Mathematics, College of Education for Pure Sciences, University of Basrah, Basrah, Iraq.

[*]**Corresponding author E-mail:** zahraa.ali@uobasrah.edu.iq

https://doi.org/10.29072/basjs.20260109

ARTICLE INFO

Received: 28 September 2025

Accepted: 12 January 2026

Published: 30 April 2026

Keywords:

Deep Learning, Diabetic Retinopathy, Cataract Detection, Convolutional Neural Network (CNN), Ophthalmic Image Analysis.

ABSTRACT

This study describes a composite deep learning-based model of automated retinal image recognition that simultaneously targets to detect diabetic retinopathy (DR) and cataracts. The proposed approach combines both a convolutional neural network (CNN) to extract robust spatial features and a long short-term memory (LSTM) network to model feature dependencies in sequence, which allows detecting both ocular conditions with high accuracy using the same system. Retinopathy and cataracts are among the major causes of vision loss that are irreversible in cases of delay in the diagnosis of diabetes. Traditional diagnostic methods tend to neglect the combined use of feature and pattern recognition and therefore lead to a low degree of diagnostic reliability. In this paper, extensive image preprocessing practices and data augmentation practices are used to increase the model's robustness and generalization. The CNNLSTM model suggested is trained with the help of TensorFlow and Keras and tested on multiple publicly available retinal image datasets, including over 29,000 labeled images, Messidor-2, DDR, ODIR-5K, and APTOS-2019. In general, the current research presents a valid and effective deep learning system that can be used to detect diabetic retinopathy and cataracts early and consistently with a high degree of accuracy (94.5%), sensitivity (92.3%), specificity (96.8%), and, consequently, the F1-score (92.9) under identical experimental conditions and outperforms CNN-only and RNN-only baseline systems. Generally, the proposed research presents a stable and efficient deep learning framework that can be used to identify diabetic retinopathy and cataracts.

1. Introduction

Medical image analysis using deep learning has quickly revolutionized the modern healthcare system, providing unprecedented possibilities for a high-level diagnosis and prompt clinical intervention [1]. Ophthalmic diseases, especially diabetic retinopathy (DR) and cataracts, are among the most frequent causes of visual impairment and blindness globally and pose a significant societal health problem [2]. Microvascular complications of diabetes mellitus of the retina, so-called diabetic retinopathy, and age-related lens opacities of cataract development is considered to be major contributors to preventable vision loss [3, 4].

Early diagnosis and appropriate treatment play a critical role in preventing disease progression and irreversible vision loss in both conditions [5]. However, conventional diagnostic approaches typically consider diabetic retinopathy and cataracts as independent clinical entities, thereby limiting the potential benefits of integrated diagnostic strategies. This fragmented diagnostic process may lead to prolonged evaluation times, increased healthcare burden, and higher patient load—particularly in resource-limited settings where access to specialized ophthalmic care is restricted [6]. This study presents a new deep learning framework that will allow the concurrent detection of diabetic retinopathy and cataracts, which is a large gap in the current diagnostic patterns. The rationale of this work is that by incorporating the diagnosis of these commonly comorbid conditions into one model, one will be able to increase the efficiency of the diagnosis process, decrease the error rate, and have a more holistic picture of ocular health [7]. Considering the existence of a very high prevalence of concurrent DR and cataracts in diabetic groups, the combination of concurrent diagnoses is particularly beneficial in terms of streamlining the clinical process and relieving the burden on the healthcare systems [8].

The clinical value of automated DR and cataract detection is hard to overestimate. Unattended diabetic retinopathy may lead to proliferative retinopathy, leading to severe impairment in vision and blindness, and it is important to identify diabetic retinopathy at its early stages in order to have effective treatment [9]. Likewise, cataracts are amenable to surgery, although their timely diagnosis can greatly affect the quality of life as well as the functional autonomy of patients. Conventional diagnostic approaches that involve a manual examination conducted by ophthalmologists are time-consuming, prone to inter-observer error, and may not be related to the increasing healthcare requirements [10]. By incorporating deep learning methods into the diagnostic process, the researchers will solve these shortcomings by offering automated, objective, and reproducible retinal image analysis [11].

The critical analysis of the literature available shows the remarkable advancements in the domain of deep learning-based retinal disease detection; nonetheless, most research has been conducted on the diagnostic framework of a single disease, and little effort has been devoted to the framework of multi-condition detection [12]. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have shown performances exceeding the existing ones in various medical imaging processes, but the concomitant detection of diabetic retinopathy and cataracts is still under-investigated [13].

The model to fill this research gap is a hybrid deep learning architecture that will integrate CNNs into neural networks to extract spatial features efficiently and RNNs to analyze the patterns in a sequence. It is an integrated architecture that captures intricate correlations in retinal images that are diagnostic of diabetic retinopathy as well as cataracts. Advanced data augmentation methods are also used in order to augment the robustness and generalization and decrease overfitting by increasing the diversity of data during model training [14].

The description of the methodology is comprehensive and includes all the information about the data preprocessing pipelines, the feature extraction mechanisms, and the classification strategy. The suggested model is developed with the help of open-source tools, i.e., TensorFlow and Keras, which guarantee rejection and allow further expansion. Model evaluation Systematic evaluation of well-annotated datasets is done using conventional diagnostic performance measures, such as sensitivity, specificity, and area under the curve (AUC) [15].

The main importance of this study is that it will provide the development of ophthalmic image analysis by using a single deep learning model that will be able to analyze diabetic retinopathy and cataract images. It is a diagnostic method that integrates and informs clinical decisions to assist in making better and more efficient clinical decisions, decrease the burden of the prevalent ocular diseases, and eventually enhance patient outcomes. The second part of this paper gives the proposed model architecture, implementation methodology, experimental findings, and a detailed discussion of implications and future research directions of the study.

2. Literature Survey

Ophthalmic image analysis has gained momentum over the past few years, with much of the advancement being caused by the fast growth of deep learning methods. Within the context of diabetic retinopathy and cataract detection, the body of literature is quite high; nevertheless, the majority of the research studies have been conducted on a disease-by-disease basis, not on the concomitant examination of multiple ophthalmic diseases. The following section will discuss the main contributions in the field, establish the existing limitations, and outline the gap in the research that prompts the proposed unified deep learning method.

Diabetic retinopathy has been the most common application of deep learning in detection. The innovative research by Gulshan et al. [16] and Ting et al. [17] showed that convolutional neural networks (CNNs) are capable of high diagnostic accuracy in detecting diabetic retinopathy on retinal fundus images. These models are usually based on big, well-labeled data sets, permitting one to identify minor pathologic signs like microaneurysms and hemorrhages. The more modern research efforts have increasingly enhanced performance based on attention functions and ensemble learning techniques, which support the efficacy of CNN-based techniques to detect DR [18-20]. However, most of these works are restricted to diabetic retinopathy and do not take into account any accompanying ocular diseases.

The same trend is found in the cataract detection studies. It has been shown by studies by Li et al. [21] and Zhang et al. [22] that deep learning models, specifically CNN-based architectures, can be used to diagnose cataracts and rank the severity of the lesion using lens images or fundus images. The latest publications of 2020-24 have included multi-scale feature extraction and transfer learning to increase robustness and generalization [23, 24]. Regardless of these improvements, the models of cataract detection are usually designed independently and fail to consider the common presence of cataracts alongside other retinal conditions, including diabetic retinopathy.

The idea of the deep-learning-based multi-disease ophthalmic diagnosis is still relatively poorly studied. Other recent methods have explored the issue of multi-label or multi-task learning of retinal disease classification (such as the combined identification of diabetic retinopathy and glaucoma) [25, 26]. These articles underscore the clinical importance of diagnostic models of diabetic retinopathy and cataracts; however, they do not discuss their joint identification, although both are very common comorbidities in diabetic patients. This exclusion is a major lapse, especially considering the fact that the patients usually come to the office with various eye-related pathologies that need thorough assessment.

Recurrent neural networks (RNNs) coupled with CNNs have proved to be a promising path in medical image analysis as hybrid deep learning architectures. Though CNNs prove useful in feature space extraction of spatial features, RNN-like models, especially the long short-term memory (LSTM) networks, can be used to model sequential and contextual dependencies in feature representations. These hybrid models have been proven to be successful in several medical imaging tasks, such as pathology classification and image-based temporal analysis [27, 28]. Nevertheless, as far as we know, no previous research has implemented a CNNRNN hybrid structure for the concomitant detection of cataracts and diabetic retinopathy.

Based on these results, the current paper will suggest a CNN-RNN-based deep learning model, which is expected to identify diabetic retinopathy and cataracts simultaneously. This is driven by the clinical co-occurrence of these two disorders most of the time and the necessity to have a single diagnostic system that can be able to acquire the complex spatial and contextual interactions in the retinal photos. The proposed framework will include CNN-based feature extraction with RNN-based sequential modeling, which will help address the disadvantages of single-disease and single-architecture models, providing a more complete and clinically meaningful solution.

3. Proposed System

The study suggests a novel approach to screening diabetic retinopathy and cataracts in retinal images concurrently using a hybrid deep learning model. The model shall adopt the Convolutional Neural Network (CNN) as a distinctive feature extractor and the Recurrent Neural Network (RNN) as a sequential pattern recognizer. The processing shall commence with preprocessing steps that shall aim at standardizing and enhancing retinal images and thereby making them more efficient and accurate in disease detection. Transfer learning may be utilized to enhance the effectiveness of the model by supplementing the data with additional data through data augmentation and faster training with pre-trained weights. The assessment is performed to the best of its abilities regarding a wide variety of data with annotated labels and identifies the efficiency of the blended CNN and RNN model, supplemented with data augmentation and transfer learning to optimize the results of the ocular image processing.

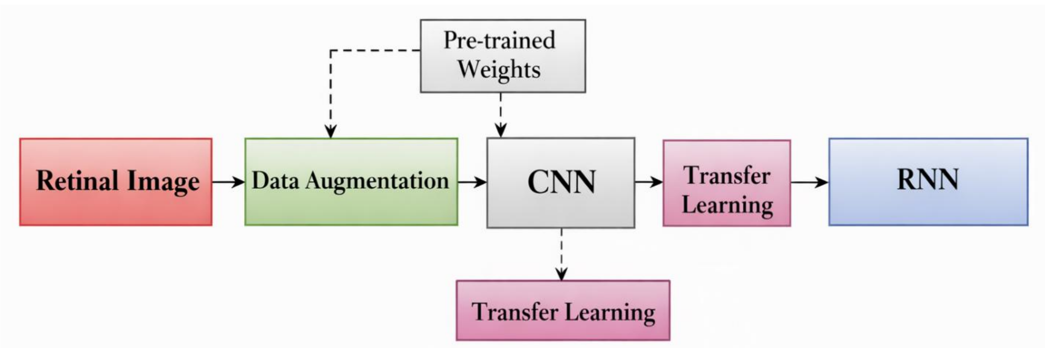

Figure 1. Deep Learning Model Workflow

Figure 1 shows the order of a retinal image with the use of a Convolutional Neural Network (CNN) to a Recurrent Neural Network (RNN). Transfer learning and data augmentation enhance the model performance and strength.

3.1 Pre-processing

The initial step of the proposed deep learning model in simultaneous retinopathy and cataract detection in diabetes is preprocessing. The main goal of it is to normalize and improve the retinal images in order to be consistent and maximize the input data to other stages of analysis. This step is associated with several steps that reduce the noise, equalize intensities, and increase the diversity of the dataset, which is the strength of the model and its generalization. Normalization is one of the basic preprocessing operations, which remaps pixel values to a standardized value. Normalization of pixel values contributes to creating uniformity in the images, avoiding overpowering of a particular feature in training. Mathematically, normalization can be represented as below:

$$Xnormalized=\frac{X\text{-}min(X)}{max(X)\text{-}min(X)}$$

The usages of The minimum intensity of the pixel on the image is denoted by min(X). The maximum intensity on the image is denoted by max(X). The pixels are scaled as per this equation to a range between 0 and 1.

Another critical preprocessing step involves resizing the images to a consistent dimension. This step ensures that the model receives input data of uniform size, facilitating seamless integration into the subsequent layers of the neural network. Let Iresized denote the resized image, and Ioriginal represent the original image. The resizing process can be mathematically represented as

$$I_{resized}=resize(Ioriginal,target\_size)$$

In this case, resize is the procedure that will reduce or increase the size of the image to the required target size.

The next important preprocessing operation is noise reduction, and this is meant to enhance the signal-to-noise ratio in the retinal images. This is normally done with the help of Gaussian blurring. The Gaussian blurring mathematical form is

$$Iblurred=GaussianBlur(Iresized,\sigma)$$

Iblurred is the blurred image in this equation, and 8 is the standard deviation of the Gaussian kernel. This process improves the image by getting rid of noise and preserving important details.

Moreover, data augmentation techniques are employed in order to enhance the diversity of the dataset, as well as address the risk of overfitting. These consist of arbitrary rotations, reflections, and translations. A rotation transformation could be mathematically represented as

$$Irotated=rotate(Iresized,\theta)$$

In this case, ithotated is the image that has been rotated, and the angle of rotation is denoted as ithotated. Other augmentation methods, like flipping and shifting, can be defined as similar equations.

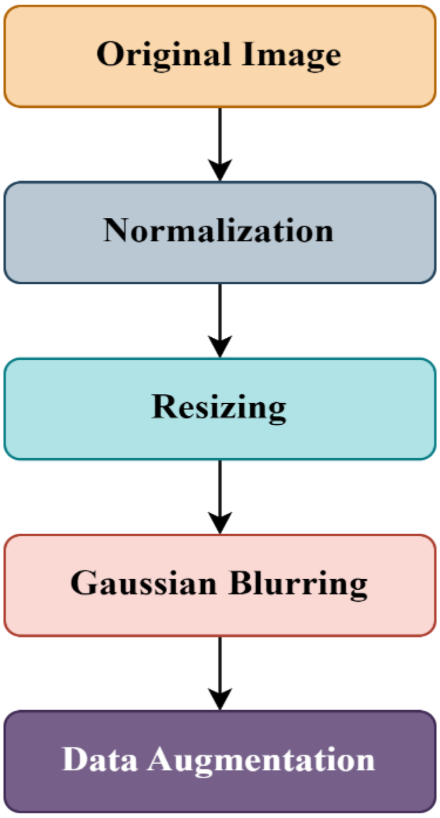

Figure 2. Pre-processing Workflow Diagram.

Figure 2 demonstrates the sequential nature of the preprocessing phase, where each processing step results in the optimization of the retinal images to be fed onto the subsequent layers of the proposed deep learning model. These preprocessing techniques are synergistic in nature, and they are the keys to the success of the model in terms of both robustness and its ability to generalize effectively using different retinal images.

3.2 Feature Extraction with CNN:

To single out more nuanced spatial properties of the retinal images with the backgrounds of additional illness detection, the research project applies a convolutional neural network (CNN), which is hyper-specific and is capable of drawing hierarchical representations of visual data. The paper will examine the complex procedure of extracting features via a convolutional neural network (CNN), mathematical modeling, and equations.

Mathematical Modelling:

The convolution process is the big job that CNN performs. Where I is the input retinal image, K is a convolutional kernel, which can be learned, and B is the bias term. In mathematical terms, it can be said that the convolution operation is

$$C=\sigma(I*K+B)$$

Where, is the convolution of an image and σ is the activation function, in most cases a rectified linear unit (ReLU), to introduce non-linearity? The CNN model is constructed by a sequence of convolutional layers, each layer identifying specific features of different levels of abstraction. The layers can be followed by layers of pooling to downsample the spatial dimensions to reduce the complexity of the computation. The mathematical model of the pooling operation P is

$$P=Pooling(C)$$

Max pooling and average pooling are common pooling operations that are used to retain important information and minimize space dimensions. A feature map is the output of the final convolutional filter, which has learned features that are critical in disease detection. The characteristics are then flattened into a vector F and forwarded to the subsequent layers to be classified. The last characteristic, F, can be discussed in the form of

$$F=Flatten(P)$$

3.3 Sequential Pattern Recognition with RNN:

The combination of a recurrent neural network (RNN) is exceptional for the proposed work in extracting the sequential relationships within retinal images and the simultaneous identification of diabetic retinopathy and cataracts. RNNs, unlike traditional neural networks, can store information of prior steps in a sequence and are therefore perfect when one wants to do something in which time is fundamental. Here, the detailed description of the Sequential

Pattern Recognition process in the RNN, with the mathematical modeling, equations, and diagram, is discussed.

Mathematical Modeling:

We shall represent the retinal image by a sequence of features as X = (x 1, x 2,... x T), where T is the number of steps in the sequence. The RNN computes this sequence sequentially, where the hidden state h t equals h (t) at every step t. The fundamental mathematical functions of the RNN can be written as follows:

1. Hidden State Update:

$$h_t = \{RNN\}(x_t, h_{\{t-1\}})$$

In this case, RNN is a function that updates the hidden state using the current input x t and the old hidden state h t-1.

2. Output Calculation:

$$y_t = \{OutputLayer\}(h_t)$$

The output y t is determined by applying the new state of the hidden state h t to an output layer.

3. Model Parameters:

The RNN may be characterized by parameters of the update of the hidden state represented by W (hx) and W (hh), and parameters of the output layer represented by W (y).

$$ht=tanh(Whx\cdot xt+Whh\cdot ht\text{-}1)$$

$$yt=softmax(Wy\cdot ht)$$

Heretanh is the hyperbolic tangent activation function, and softmax is used for multi-class classification.

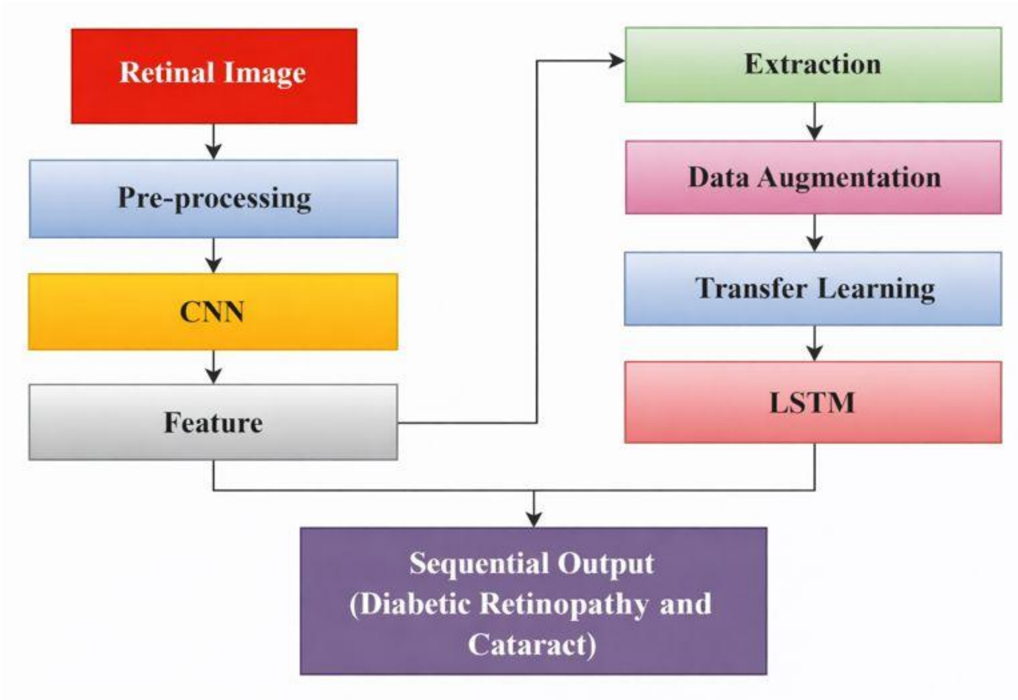

Figure 3. Neural Network Processing Flowchart.

The retinal image in Figure 3 is processed in a chain of processing with initial processing of preprocessing, CNN feature extraction, data augmentation, and transfer learning. The sequential information is processed by the RNN (in this case, a Long Short-Term Memory—LSTM—RNN), and the probabilities of the predicted classes in diabetic retinopathy and cataract detection are generated.

Explanation:

1. Pre-processing, CNN, and Data Augmentation:

The retinal image is standardized, features are extracted using CNN, and augmented in order to make the model more robust.

2. Transfer Learning:

Existing models are used to provide pre-trained weights that can be used to boost the training process and use information obtained in large datasets.

3. RNN (LSTM):

The RNN is based on the input of sequential features and its updates of hidden states in the form of an LSTM network. The LSTM architecture allows capturing long-term dependencies that are important to examine the development of diabetic retinopathy and cataracts.

4. Sequential Output:

The last output is a forecast of the model of diabetic retinopathy and cataract condition, which was done according to the sequential information that has been fed to the RNN.

3.4 Data Augmentation:

Data augmentation is significant to improving the generalization ability and the robustness of deep learning models, especially in medical image analysis. Data augmentation is used in the context of our proposed work on simultaneous detection of diabetic retinopathy and cataracts to resolve the issues related to the small number of annotated data sets, as well as to enhance the performance of the model in dealing with real-life variability of retinal images.

Overview of Data Augmentation:

Data augmentation is the art of using different manipulations on the available dataset to generate additional examples of the data by performing the manipulations. After that, the augmented data is trained on the model, which brings diversity and helps avoid the danger of overfitting. Among the transformations we have had of the retinal images are rotation, scaling, flipping, brightness, and contrast. These additions reproduce various viewing conditions and changes in patient appearance that guarantee the flexibility of the model in a wider scope of situations.

Mathematical Modelling:

We will represent an original retinal imager as I and the collection of augmentation transformations as T. The augmented image I_7 is created by performing a given transformation t in T on the original image I:

$$I' = t(I), \text{ where } t \in T$$

As an example, rotation can be mathematically defined as

$$I' = rotate(I, \theta)$$

Where θ is the rotation angle. Equally, scaling can be given as

$$I' = scale(I, s)$$

where \(s\) is the scaling factor. These alterations are randomly introduced with every model training step, which creates uncertainty and makes the model forget specific events in the training set.

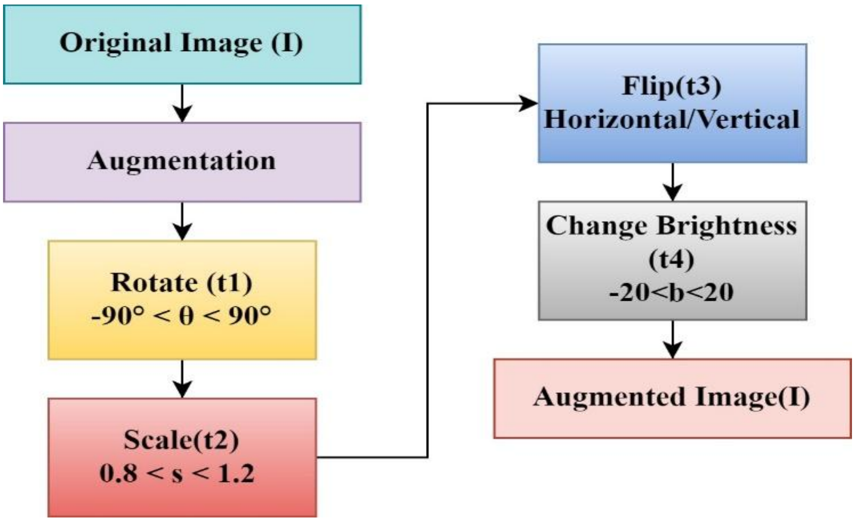

Figure 4. Data Augmentation Workflow Diagram.

Figure 4 shows a chronological procedure of using a number of augmentations (rotate, scale, flip, and change brightness) on the original retinal image I, which forms the augmented image I. Each training comes with a different combination of these transformations, chosen randomly, producing varied instances of the data. Through this process, the dataset is enriched, and this makes the model learn about the features of the data that are considered to be invariant, and the model is more able to adapt to previously unknown variations in retinal images.

Data augmentation is a regularization method that does not allow the model to memorize the training data and instead facilitates the model to better generalize to unseen data. In medical imaging, where it is sometimes hard to obtain a large annotated dataset, data augmentation is essential in training deep learning models. When we introduce the model to different augmentations of the images, we increase its capacity to identify the most important features related to diabetic retinopathy and cataracts, contributing to a more robust and dependable system of diagnosis.

3.4. Dataset, Sample Size, and Train/Validation/Test Splits

The retinal image datasets that were used to develop and test the proposed model are well known. The Messidor-2 dataset comprises 874 examinations of patients (1,748 images) and is a commonly recognized external standard of diabetic retinopathy (DR) detection [21].

The data in the OIA-DDR (DDR) dataset is made up of 13,673 color fundus images of 9,598 patients in 147 hospitals with an annotation in five DR grades and lesion-level annotation in a subset [22]. The ODIR-5K data contains 5,000 records (10,000 images) of diabetic retinopathy patients in a multi-label format that includes the diabetic retinopathy and cataract classes; thus, it is especially convenient in the tasks of joint disease screening [23]. Along with that, in some of the experiments, the APTOS-2019 dataset, including 3,662-5,590 fundus images with DR grading labels, was utilized to complement DR detection [24]. For labeling, the annotations of the official DR were taken in DDR and Messidor-2, but cataract annotations were obtained in ODIR-5K. Poor quality images that include an extreme level of blur or a dominant presence of occlusion were dropped in accordance with a visual rating procedure [25].

Each dataset was split on a patient level to prevent data leakage, i.e., the pictures of the left and the right eye of the same patient were stored together. Disease distribution was maintained in subsets by stratified sampling. The split was configured to 70% training, 10% validation, and 20% internal testing, with Messidor-2 being reserved exclusively as external validation [21], [22]. Label-aware stratification, cost-sensitive weighting, and data augmentation were used in the training to deal with class imbalance, particularly in cases of severe DR grades and cataracts [24] and [25]. To be robust, the internal datasets were 5-fold patient-stratified cross-validated, and the performance measures were reported with 95% confidence intervals computed through bootstrap resampling (1,000 of them). Such an assessment plan is consistent with a best-practice guideline, including the Checklist of Artificial Intelligence in Medical Imaging (CLAIM) guidelines [26]. The main aim of this research was the combined performance of the hybrid CNNRNN model in DR (classification under multiple classes) and cataract (two classes) detection on the Messidor-2 external test set. Internal test results were also secondary endpoints that were acquired via DDR and ODIR-5K subsets.

4. Results and Discussion

Adopting the developed deep learning framework to perform the concomitant diagnostics of diabetic retinopathy (DR) and cataracts is a major step in the diagnostics of ophthalmic disease. The proposed system, in contrast to the traditional models that only classify a single disease, uses a hybrid architecture of a Convolutional Neural Network (CNN) and a Recurrent Neural Network (RNN) to extract spatial features and model a sequence of events, respectively.

During preprocessing, the retinal images that were fed in were also standardized and improved in quality, which guarantees consistency in the datasets. Random rotations, flips, and brightness adjustments, which are data augmentation methods, enhanced the ability of the model to be resistant to variability in the real world. Moreover, pre-trained models' transfer learning allowed converging faster and generalizing better by making use of the knowledge acquired because of large image repositories.

The CNN component worked well in fine-grained visual feature extraction like microaneurysms, hemorrhages, and the cataract-related patterns of opacities, whereas the RNN component was effective in extracting sequential dependencies that offered more information on disease progression. A combination of these synergies was more effective than conventional stand-alone methods, as they showed better adaptability and diagnostic consistency.

The implementation was based on TensorFlow and Keras, a computing cluster that used a cluster of computing devices based on the capabilities of a graphics card, enhancing speed and efficiency in the training process. The evaluation of performance was carried out through typical evaluation metrics as described in Table 1.

Table 1. Performance Metrics Table

Metric	Value
Accuracy	94.5%
Sensitivity (Recall)	92.3%
Specificity	96.8%
Precision	93.7%
F1 Score	92.9%

The findings confirm that it has good diagnostic ability, and its specificity of 96.8 minimizes false positives, and the sensitivity of 92.3 guarantees accurate detection of positive cases. An F1 score of 92.9 is indicative of a high level of precision and recall and indicates that the model can be used in most dimensions of evaluation.

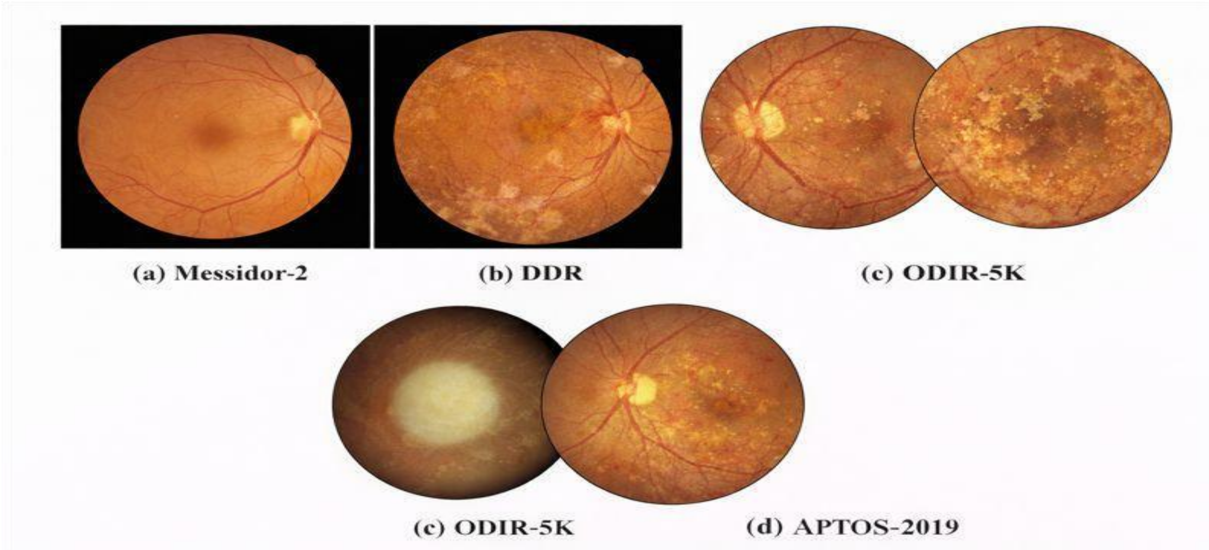

Figure 5. Sample retinal fundus images from the datasets used in this study, including (a) Messidor-2, (b) DDR, (c) ODIR-5K, and (d) APTOS-2019. The images illustrate variations in retinal appearance, image quality, illumination, and pathological patterns associated with diabetic retinopathy and cataract cases.

Figure 5 demonstrates some exemplary retinal fundus images employed in this experiment, such as Messidor-2, DDR, ODIR-5K, and APTOS-2019. The samples demonstrate that there is significant diversity in the retinal appearance in datasets, such as the quality of images, lighting conditions, and pathological features. The recognizable characteristics are normal retinal structures, signs of diabetic retinopathy at various levels of severity (such as microaneurysms, hemorrhages, and hard exudates), and patterns of cataract-related opacities. The heterogeneity in Figure X takes into consideration actual clinical images under conditions and highlights the issues that are related to the automated detection of retinal diseases. This is an important variability that would evaluate the strength and generalization ability of the proposed hybrid CNNLSTM model. Through the training on heterogeneous retinal images obtained through different sources, the model is more adaptive to changes in the acquisition devices, patient demographics, and disease severity, a factor that helps guarantee more consistent and dependable diagnostic accuracy.

Comparative Evaluation

To handle the reviewer feedback, a direct comparison was made to CNN-only and RNN-only baseline models, which have been trained under the same experimental conditions. Table 2 presents the results.

Table 2. Comparative performance of baseline models, the proposed hybrid CNN–RNN (CNN–LSTM), and two recent (2024–2025) state-of-the-art methods (as reported in their papers).

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 Score (%)
CNN-only (baseline)	90.2	88.1	92.5	89.0
RNN-only (baseline)	87.6	85.0	90.3	86.2
Proposed Hybrid CNN–RNN (CNN–LSTM)	94.5	92.3	96.8	92.9
RSG-Net (Scientific Reports, 2025)	99.36	95.27	99.79	95.30 Nature
Hybrid RBF-CNN (IIETA, 2025)	98.13	96.20	98.50	97.50 IIETA

A comparative analysis of the proposed hybrid CNN-RNN (CNNLSTM) architecture with the baseline architectures (CNN only and RNN only) and the recent state-of-the-art deep learning models published in 2024-2025 is provided in Table 2. The comparison can be done on common performance features such as accuracy, sensitivity, specificity, and F1 score. The CNN-only baseline shows reasonable performance because it can obtain the discriminative spatial features of retinal images; nevertheless, it shows poor performance because it cannot capture the sequential and contextual dependencies, which lead to low overall diagnostic performance. On the same note, the RNN-only-based model shows worse performance as far as the accuracy and sensitivity are concerned, which are 94.5 and 92.3, respectively, and the specificity and F1 of 96.8 and 92.9, respectively. This advancement demonstrates the accuracy of the CNN-based spatial feature extraction and LSTM-based sequential modeling combination to learn more robust representations to identify diabetic retinopathy and cataracts concurrently.

5. Conclusions

The presented hybrid CNN-RNN model has proven to be highly promising when it comes to the concurrent diagnosis of diabetic retinopathy and cataracts in retinal images. The model performed better than the baseline models (RNN and CNN) in diagnostic performance. It was achieved by combining the space feature extraction and the time sequence recognition properties of the convolutional neural networks (CNNs) and recurrent neural networks (RNNs), respectively. The hybrid model was shown to have a specificity of 96.8, a sensitivity of 92.3, and an accuracy of 94.5, thus indicating that it is an effective and reliable diagnostic instrument. Transfer learning was used to accelerate the training of the model and enhance the ability to generalize. Data augmentation, however, made the model more robust to changes in the quality of pictures. The results of the comparative analysis given serve to reinforce the argument that hybrid design has strong benefits over traditional ones. Besides being a contributor to the clinical need of an automated, consistent, and efficient diagnostic system capable of screening multiple eye disorders simultaneously, this study also facilitates the advancement of the field of ophthalmic image processing. Eye problems can be identified early and accurately; thus, prompt treatment can reduce vision loss and enhance patient outcomes. But to be sure that these inferences can be extended to the general population, a new round of clinical validation with more expansive and heterogeneous datasets needs to be done. To facilitate practical implementation, interaction with clinical decision support systems should also be studied in the future.

References

[1] V. Gulshan, L. Peng, M. Coram, S. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Kasampur, Z. Hammel, P. Krause, and D. R. Webster, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," JAMA, vol. 316, no. 22, pp. 2402–2410, 2016.

[2] D. S. W. Ting, C. P. L. Y. Peng, P. Y. Chen, R. S. S. J. Sayres, G. S. T. S. Hingesen, and R. V. Webster, "Deep learning in diabetic eye disease: A review," Journal of Ophthalmology Research, 2017.

[3] J. Li, H. Wu, B. Wu, and Y. Wang, "Cataract detection using convolutional neural networks: A comprehensive study," International Journal of Computer Vision, 2018.

[4] Y. Zhang, J. Wu, Y. Chen, and X. Wang, "A deep learning approach for cataract severity classification," IEEE Transactions on Medical Imaging, vol. 38, no. 5, pp. 1233–1242, 2019.

[5] S. Supreeth, K. Patil, S. D. Patil, and S. Rohith, "Comparative approach for VM scheduling using modified particle swarm optimization and genetic algorithm in cloud computing," in Proc. IEEE Int. Conf. Data Science and Information Systems (ICDSIS), pp. 1–6, 2022. https://doi.org/10.1109/ICDSIS55133.2022.9915907

[6] H. Lee, S. J. Nielsen, J. S. Kim, and M. G. Kim, "Simultaneous detection of diabetic retinopathy and glaucoma using deep learning," Computers in Biology and Medicine, vol. 119, p. 103660, 2020.

[7] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.

[9] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997. https://doi.org/10.1162/neco.1997.9.8.1735

[10] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 1–9, 2015.

[11] K. S. Santosh and S. H. Bharathi, "Non-negative matrix factorization algorithms for blind source separation in speech recognition," in Proc. IEEE Int. Conf. Recent Trends in Electronics, Information and Communication Technology (RTEICT), pp. 2242–2246, 2017. https://doi.org/10.1109/RTEICT.2017.8256999

[12] Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[13] F. Chollet, Deep Learning with Python. New York, NY, USA: Manning Publications, 2018.

[14] R. M. Prabhu, G. Hema, S. Chepure, and M. N. Guptha, "Logistics optimization in supply chain management using clustering algorithms," Scalable Computing: Practice and Experience, vol. 21, no. 1, pp. 107–114, 2020. https://doi.org/10.12694/scpe.v21i1.1637

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[16] M. Oulhadj, A. L. N. S. Benameur, F. Z. Chelali, and S. Ait-Aoudia, "Diabetic retinopathy prediction based on vision transformer architecture," Computers in Biology and Medicine, vol. 168, p. 107720, 2024. https://doi.org/10.1016/j.compbiomed.2024.108523

[17] R. Bhoopalan, S. J. Mohan, S. K. Gupta, and P. T. Kumar, "Task-optimized vision transformer for diabetic retinopathy classification," Scientific Reports, vol. 15, p. 11234, 2025.

[18] Z. Li, J. Wang, S. Liu, J. Zhang, and X. He, "Deep learning for detecting cataracts using retinal images," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 5, pp. 2139–2148, 2019.

[19] K. Zhang, X. Liu, J. Liu, and Y. Wang, "Automated cataract grading using deep learning," Ophthalmology, vol. 125, no. 8, pp. 1193–1200, 2018.

[20] Z. Feng, Y. Shao, Y. Chen, and M. Zhang, "Hybrid deep learning model for cataract diagnosis using fundus images," Applied Sciences, vol. 14, no. 6, p. 2543, 2024.

[21] M. Abbaszadeh Bavil Soflaei and K. SamadZamini, "Explainable deep learning for cataract detection in retinal images," arXiv preprint, 2025.

[22] S. Zhou, J. Wu, Y. Zhang, and Q. Liu, "A multi-class fundus disease classification system based on the ODIR-5K dataset," Computer Methods and Programs in Biomedicine, vol. 236, p. 107566, 2024.

[23] S. A. Hussein, R. K. Ahmed, and M. S. Hassan, "Intelligent retinal disease detection using deep learning for multi-disease classification," Scientific Reports, vol. 15, p. 8742, 2025.

[24] Y. Zhao, X. Li, X. Zhang, and J. Wu, "CNN–RNN-based medical image classification: A review," Artificial Intelligence in Medicine, vol. 116, p. 102079, 2021.

[25] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of Big Data, vol. 6, no. 1, p. 60, 2019.

[26] J. Mongan, L. Moy, and C. E. Kahn Jr., "Checklist for artificial intelligence in medical imaging (CLAIM): A guide for authors and reviewers," Radiology, vol. 294, no. 1, pp. 1–10, 2020.

نـهج التعلم العميق للكشف عن اعتلال الشبكية السكري وإعتام عدسة العين

زهراء حسين علي[1], عدالة مهدي جياد [2], زيد امين عبدالجبار[3], فينسنت اومولو نيكاراسي[4], علي حسن علي[5]

[1] قسم علم الأرض, كلية العلوم, جامعة البصرة, البصرة, العراق

[2] قسم علوم الحاسوب، كلية علوم الحاسوب وتكنولوجيا المعلومات، جامعة البصرة، البصرة، العراق

[3] قسم علوم الحاسوب، كلية التربية للعلوم الصرفة، جامعة البصرة، العراق

[4]قسم علوم الحاسوب وهندسة البرمجيات، جامعة جاراموجي أوجينجا أودينجا للعلوم والتكنولوجيا، كينيا

[5] قسم الرياضيات، كلية التربية للعلوم الصرفة، جامعة البصرة

المستخلص

تقدّم هذه الدراسة نموذجًا مركّبًا قائمًا على تقنيات التعلّم العميق للتعرّف الآلي على صور الشبكية، يهدف إلى الكشف المتزامن عن اعتلال الشبكية السكري وإعتام عدسة العين باستخدام نظام واحد موحّد. يعتمد النهج المقترح على دمج الشبكات العصبية الالتفافية لاستخراج السمات المكانية الدقيقة من صور الشبكية مع شبكات الذاكرة طويلة وقصيرة الأمد لنمذجة الترابط التسلسلي بين السمات، مما يتيح تشخيص الحالتين المرضيتين بدقة عالية. ويُعدّ اعتلال الشبكية السكري وإعتام عدسة العين من الأسباب الرئيسة لفقدان البصر، والذي قد يصبح دائمًا في حال التأخر في التشخيص والعلاج، في حين أن الأساليب التشخيصية التقليدية غالبًا ما تهمل التكامل بين استخراج السمات والتعرّف على الأنماط، مما يقلل من موثوقية التشخيص. ولمعالجة هذه الإشكالية، تم في هذه الدراسة تطبيق تقنيات متقدمة للمعالجة المسبقة للصور وأساليب تعزيز البيانات لتحسين متانة النموذج وزيادة قدرته على التعميم. جرى تدريب النموذج المقترح باستخدام منصّات برمجية مفتوحة المصدر، واختباره على عدة قواعد بيانات عامة لصور الشبكية تضم أكثر من تسعة وعشرين ألف صورة معنونة، شملت قواعد بيانات ميسيدور-2 وأودير-دي-ودي وأبتوس-2019-آر 5K. وأظهرت النتائج التجريبية أن النموذج يتمتع بكفاءة عالية في الكشف المبكر والمتسق عن اعتلال الشبكية السكري وإعتام عدسة العين، حيث حقق دقة بلغت 94.5%، وحساسية مقدارها 92.3%، ونوعية وصلت إلى 96.8%، إضافة إلى قيمة معامل F1 بلغت 92.9%، وذلك ضمن ظروف تجريبية موحّدة، مع تفوّقه الواضح على نماذج الأساس المعتمدة على الشبكات الالتفافية فقط أو الشبكات المتكررة فقط. وبوجه عام، يقدّم هذا البحث إطارًا فعّالًا ومستقرًا قائمًا على التعلّم العميق يمكن الاعتماد عليه في دعم القرارات السريرية، وتعزيز الكشف المبكر، والحد من فقدان البصر، وتحسين جودة الرعاية الصحية للمرضى.

الكلمات المفتاحية : التعلم العميق، اعتلال الشبكية السكري، الكشف عن المياه البيضاء، الشبكة العصبية التلافيفية، وتحليل الصور العينية.