An Enhanced Convolutional Neural Network (CNN) based P-EDR Mechanism for Diagnosis of Diabetic Retinopathy (DR) using Machine Learning

Munawar Hussain
Department of Computer Science, Faculty of Computer Science & IT, Superior University, Lahore, Pakistan
su92-mscsw-f22-012@superior.edu.pk

Hassan A. Ahmed
Information Systems, Cleveland State University, Ohio, USA
h.a.ahmed@csuohio.edu

Muhammad Zeeshan Babar
Heriot-Watt University, UK
m.babar@hw.ac.uk

Arshad Ali
Faculty of Computer and Information Systems, Islamic University of Madinah, Al Madinah Al Munawarah, Saudi Arabia
a.ali@iu.edu.sa

H. M. Shahzad
Faculty of Computer Science and Information Technology, Superior University, Lahore, Pakistan
shahzad.dar@gmail.com

Saif ur Rehman
Faculty of Electrical Engineering and Technology, Superior University, Lahore, Pakistan
saifurrehman@superior.edu.pk

Hamayun Khan
Department of Computer Science, Faculty of Computer Science & IT, Superior University, Lahore, Pakistan
hamayun.khan@superior.edu.pk (corresponding author)

Abdulaziz M. Alshahrani
Faculty of Computer and Information Systems, Islamic University of Madinah, Al Madinah Al Munawarah, Saudi Arabia
alshahrani@iu.edu.sa

Received: 30 August 2024 | Revised: 23 September 2024 | Accepted: 27 September 2024

 DOI: https://doi.org/10.48084/etasr.8854

ABSTRACT

This study focuses on Diabetic Retinopathy (DR), a disease caused by diabetes that affects the retina of the eye and eventually leads to blindness. Diabetes development progresses to retinopathy and must be addressed at an early stage for effective treatment. Currently, DR is classified as Non-Proliferative DR (NPDR) and Proliferative DR (PDR). This study proposes an Enhanced DR (P-EDR) method based on CNN using a high-resolution dataset benchmark of retinal images. Initially, the data were preprocessed by normalization, augmentation, and resizing to improve image quality and feature extraction. Evaluation was based on accuracy, specificity, sensitivity, and AUC-ROC. The proposed CNN-based P-EDR outperformed advanced ML strategies such as Support Vector Machine (SVM), Random Forest (RF), Probabilistic Neural network (PNN), and Gradient Boosting Machine (GBM) that were executed and compared to diagnose and classify DR. The proposed P-EDR extracts features such as a hemorrhage of the NPDR retina image to identify the disease using image processing for classification. P-EDR provides significant features from images in detection and classification, making it a successful model for diagnosing DR with improved accuracy of 93%, sensitivity of 92%, specificity of 94%, and AUC-ROC of 0.97%. These results highlight the potential of a P-EDR-based machine learning model to support ophthalmologists with the early and precise detection of DR, eventually helping with appropriate treatment and prevention of vision loss.

Keywords-diabetic retinopathy; machine learning; convolutional neural networks; support vector machines; random forest; gradient boosting machines; medical image analysis; Non-Proliferative Diabetic Retinopathy (NPDR); Proliferative Diabetic Retinopathy (PDR)

I. INTRODUCTION

Diabetic Retinopathy (DR) is a severe complication of diabetes that affects the eyes by impacting the blood vessels within the retina, leading to vision disability and possibly visual impairment if not treated promptly [1]. The worldwide prevalence of diabetes has expanded over the past few decades, increasing the chance and frequency of DR. Early discovery and convenient intervention are significant in protecting vision [1]. Manual examination by ophthalmologists is time-consuming and subject to change. Several Machine Learning (ML) techniques have been proposed to automate and increase the accuracy of DR determination by analyzing retinal pictures [2-8]. These techniques can handle large volumes of data and distinguish designs that are not apparent to the human eye, thus encouraging early and exact detection. This study examines ML methods to detect DR, promoting a reliable system for early discovery and treatment to avoid vision impairment, using Convolutional Neural Networks (CNN), Support Vector Machines (SVM), Random Forests (RF), and Gradient Boosting Machines (GBM). A dataset of high-resolution retinal images from the Kaggle DR detection competition was used. The images were preprocessed by normalization, expansion, and resizing. A CNN technique was the best-performing model in improving DR detection, based on crucial execution measurements, including accuracy, sensitivity, specificity, and AUC-ROC.

II. RELATED WORKS

In recent years, the application of ML in medical image analysis, especially for diagnosing DR, has gained noteworthy attention. Several studies have illustrated the suitability of different ML algorithms in this space. For instance, in [9], a DL algorithm was proposed, which achieved high sensitivity and specificity in identifying DR from retinal fundus images, highlighting the potential of CNNs in therapeutic diagnostics. CNNs are very effective in image classification tasks because they can extract significant features from raw images [9]. In [10], a CNN model was used to classify retinal images into different DR stages, achieving an accuracy of more than 90%. In [11], a CNN-based model was compelling and proficient in detecting DR. SVM have also been used for DR detection. SVM are compelling in high-dimensional spaces and have been shown to perform well with restricted training data [8, 12-16].

RF and GBM are ensemble learning strategies to improve accuracy [17-19]. RF-based models have shown their viability in handling imbalanced datasets. GBM, on the other hand, builds models in a consecutive way, where each unused model strives to rectify the errors of its predecessor, improving classification performance [16, 18]. Recent advances in ML have improved models in restorative image investigation. Pre-trained models on massive image datasets can be used to improve accuracy on particular assignments such as DR detection. Despite these advances, challenges remain, such as the need for expansive and different datasets and the interpretability of ML models in clinical settings. Addressing these challenges is vital for the selection of ML procedures in DR diagnosis [19].

III. METHODOLOGY

A. Data Collection and Preprocessing

This study used the Diabetic Retinopathy Detection dataset, which contains thousands of high-resolution retinal images categorized into five classes: No DR, Gentle, Direct, Severe, and Proliferative DR [20]. The preprocessing steps to enhance the quality and features of the retinal images included normalization, expansion, and resizing. Normalization included altering the pixel values to a standard scale to ensure consistency in the dataset. Augmentation was utilized to produce additional images through revolutions, zooming, and flipping, expanding the dataset's differing qualities and enhancing the model's generalization capacity. Finally, resizing standardized the images to 224×224 pixels, ensuring uniform input sizes for the ML models.

```
Algorithm 1: Proposed CNN-based P-EDR
Input: Retinal images dataset
Output: Classification of retinal images
  into different DR stages
// Step 1: Data Preprocessing
images =
  load_images('retinal_images_dataset')
images = normalize_images(images)
images = augment_images(images,
  techniques=['rotation', 'zoom', 'flip'])
images = resize_images(images, size=
  (224, 224))
// Step 2: Model Initialization
model = Sequential()
model.add(Conv2D(filters=32,
  kernel_size=(3, 3), activation='relu',
  input_shape=(224, 224, 3)))
model.add(MaxPooling2D(pool_size=(2, 2)))
// Step 3: Adding Layers
model.add(Conv2D(filters=64,
  kernel_size=(3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Flatten())
model.add(Dense(units=128,
  activation='relu'))
model.add(Dropout(rate=0.5))
// Step 4: Output Layer
model.add(Dense(units=5,
  activation='softmax'))
// Step 5: Model Compilation
model.compile(optimizer='adam',
  loss='categorical_crossentropy',
  metrics=['accuracy'])
// Step 6: Training
train_data, val_data, test_data =
  split_data(images, train_ratio=0.7,
  val_ratio=0.15, test_ratio=0.15)
// Step 7: Evaluation
Evaluation = model.Evaluate(test_data)
print('Test accuracy:', evaluation[1])
// Step 8: Prediction
Predictions = model.Predict(new images)
```

The models were prepared utilizing the training set and hyperparameter tuning was performed using the validation set. The test set was used for the final evaluation. The following ML models were trained, tested, and compared.

B. Convolutional Neural Network (CNN)

The CNN model achieved the best performance among the models examined. Its design included different convolutional layers taken after max-pooling and completely associated layers. The ultimate softmax layer yields the probabilities for each class. The convolution operation is given by:

$$(I * K)(i,j) = \sum_m \sum_n I(m,n)K(i-m,j-n) \quad (1)$$

where I is the input image, K is the kernel, and (i,j) are the pixel positions. The cross-entropy loss for CNN is given by:

$$L = -\frac{1}{N}\sum_{i=1}^{N} [y_i \log(p_i) + (1-y_i)\log(1-p_i)] \quad (2)$$

C. Support Vector Machine (SVM)

The SVM model was prepared using features extracted from the preprocessed images. The Radial Basis Function (RBF) was used for non-linear classification.

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right) \quad (3)$$

where γ defines the influence of a single training example

D. Gradient Boosting Machine (GBM)

GBM was built using 500 boosting cycles. Each cycle focused on the hardest-to-classify samples, amending the errors of the previous cycle.

$$F_m(x) = F_{m-1}(x) + \eta \sum_{i=1}^{N} \nabla L\left(y_i, F_{m-1}(x_i)\right) \quad (4)$$

where $F_m(x)$ is the model at iteration m, η is the learning rate, ∇L is the gradient of the loss function, and y_i is the actual label. The GBM model showed a commendable performance in diagnosing DR, achieving an accuracy of 86%. The sensitivity, or true positive rate of 85% outlined the model's capacity to accurately distinguish patients with DR. The specificity, or true negative rate, was 87%, reflecting the model's efficiency in precisely detecting individuals without the disease. In addition, the Area Under the Curve of the Receiver Operating Characteristic (AUC-ROC) was 0.90, highlighting the model's efficiency in detection positive and negative cases. These results highlight its performance in detecting DR.

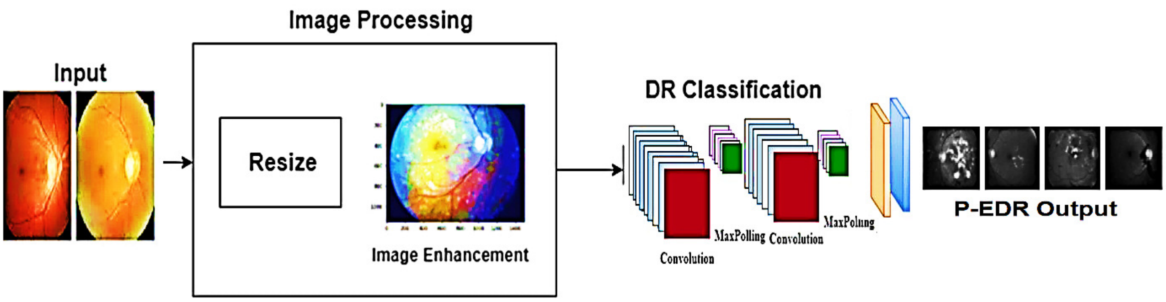

Fig. 1. Proposed experimental framework for P-EDR.

Figure 1 presents the framework for P-EDR based on a CNN along with image processing and the classification features that are used to diagnose DR images.

IV. RESULTS AND DISCUSSION

A. *Performance Metrics*

The models were evaluated based on accuracy, precision, specificity, and AUC-ROC. The CNN outperformed the other models with an accuracy of 93%, precision of 92%, specificity of 94%, and AUC-ROC of 0.97.

$$Accuracy\ (A) = \frac{TP+TN}{TP+FP+TN+FN} \tag{8}$$

$$Precision = \frac{TP}{TN+FP} \tag{9}$$

$$Recall = \frac{TP}{TP+FN} \tag{12}$$

$$F1\text{-}Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{15}$$

B. *Results*

The CNN model achieved an accuracy of 93%, a high sensitivity of 92%, and a specificity of 94%. AUC-ROC was 0.97, underscoring the model's remarkable capacity to distinguish the positive and negative classes. These measurements collectively highlight the effectiveness of this model in precisely diagnosing DR from retinal images, as shown in Figure 2.

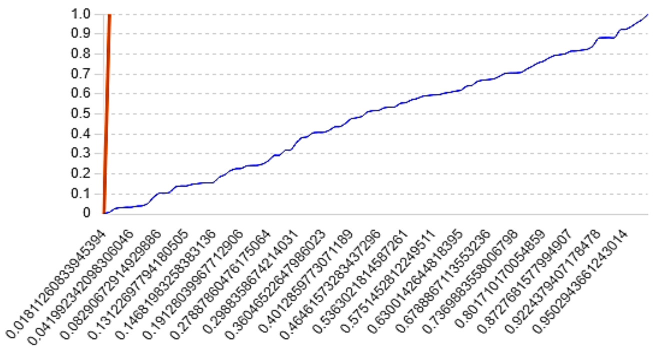

Fig. 2. ROC curve for the CNN model.

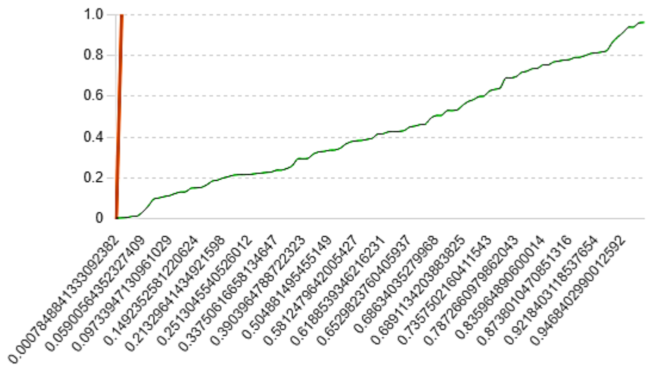

Fig. 3. ROC curve for the SVM model.

The SVM model achieved an accuracy of 88%, a precision of 87%, and a specificity of 89%. AUC-ROC was 0.9 and its ROC curve is shown in Figure 3. The RF model comprised 100 decision trees. Figures 4 and 5 show the ROC curves for the RF and GBM models.

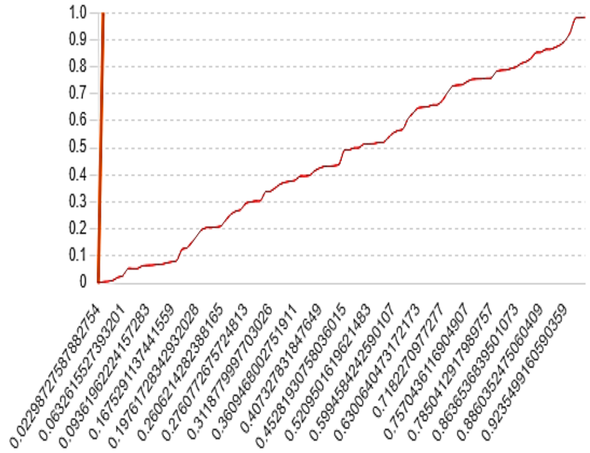

Fig. 4. ROC curve for RF.

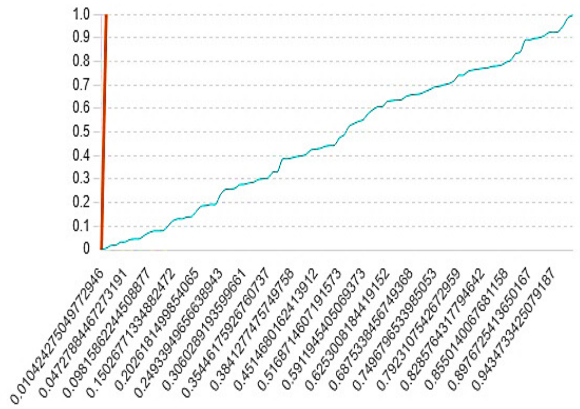

Fig. 5. ROC vurve for the GBM model.

Table I summarizes the performance evaluation metrics for the ML techniques examined. Figure 6 illustrates a comparative analysis of the models examined. The results show that CNN-based P-EDR was more efficient than other models in diagnosing DR from retinal images, outperforming SVM, RF, and GBM.

TABLE I. PERFORMANCE OF PEDR-CNN, SVM, RF, AND GBM

Metric	PEDR-CNN	SVM	RF	GBM
Accuracy	93%	88%	85%	86%
Sensitivity	92%	87%	84%	85%
Specificity	94%	89%	86%	87%
AUC-ROC	0.97	0.91%	0.89%	0.90%

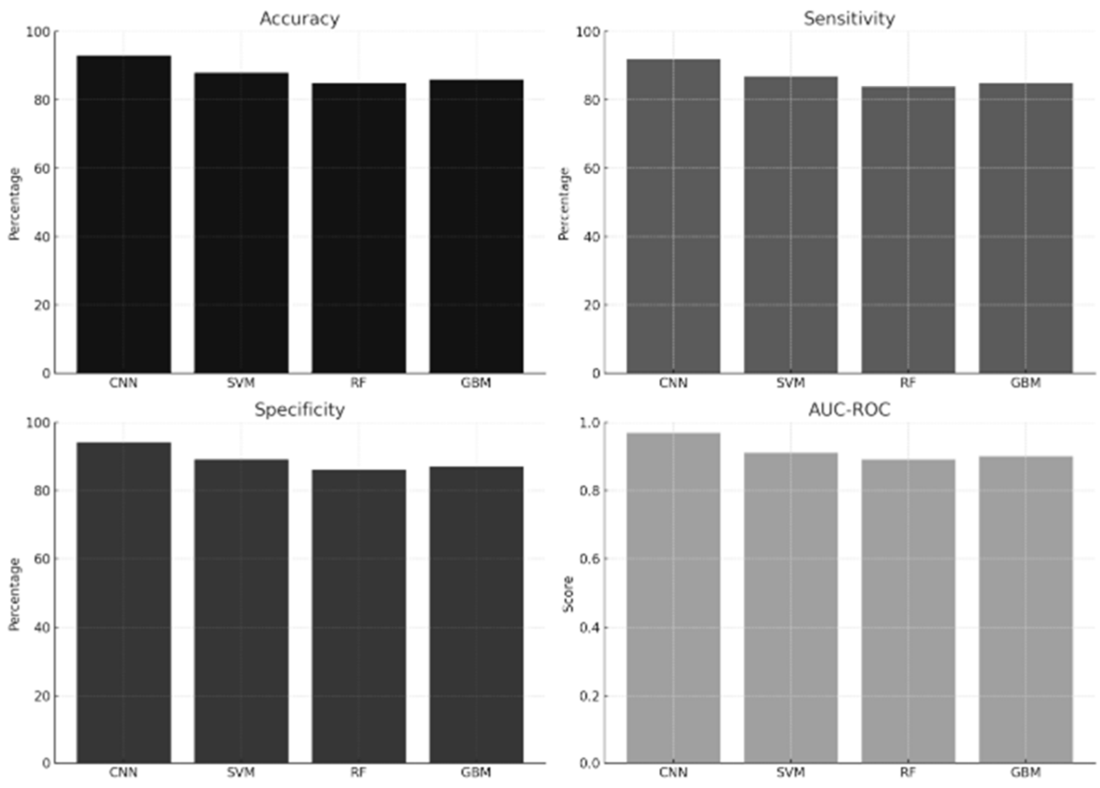

Fig. 6. Comparative analysis.

V. CONCLUSION

This study focused on a disease caused by diabetes that affects the retina of the eye. Long-term DR leads to complete blindness and continuous itching. This study presented an enhanced DR (P-EDR) detection mechanism using CNN on a high-dimensional dataset for retinal images. Initially, data were preprocessed by normalization, augmentation, and resizing to improve image quality and feature extraction. The proposed model with CNN outperformed other ML techniques such as SVM, RF, and GBM. The proposed CNN-based PEDR model can detect DR from retinal images, achieving the best results with an accuracy of 93%, a sensitivity of 92%, a specificity of 94%, and an AUC-ROC of 0.97. These results highlight the potential of CNNs to help ophthalmologists with the early and accurate detection of DR, enabling timely treatment and preventing vision loss. Future work can focus on numerous homogeneous and heterogeneous class images by utilizing other standard benchmark datasets using DL techniques and transfer learning to enhance accuracy.

REFERENCES

[1] R. Ramesh and S. Sathiamoorthy, "A Deep Learning Grading Classification of Diabetic Retinopathy on Retinal Fundus Images with Bio-inspired Optimization," *Engineering, Technology & Applied Science Research*, vol. 13, no. 4, pp. 11248–11252, Aug. 2023, https://doi.org/10.48084/etasr.6033.

[2] C. Nithyeswari and G. Karthikeyan, "An Effective Heuristic Optimizer with Deep Learning-assisted Diabetic Retinopathy Diagnosis on Retinal Fundus Images," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14308–14312, Jun. 2024, https://doi.org/10.48084/etasr.7004.

[3] G. Swapna, R. Vinayakumar, and K. P. Soman, "Diabetes detection using deep learning algorithms," *ICT Express*, vol. 4, no. 4, pp. 243–246, Dec. 2018, https://doi.org/10.1016/j.icte.2018.10.005.

[4] S. Gujral, "Early Diabetes Detection using Machine Learning: A Review," *International Journal for Innovative Research in Science & Technology*, vol. 3, no. 10, pp. 54–62, 2016.

[5] T. Sharma and M. Shah, "A comprehensive review of machine learning techniques on diabetes detection," *Visual Computing for Industry, Biomedicine, and Art*, vol. 4, no. 1, Dec. 2021, Art. no. 30, https://doi.org/10.1186/s42492-021-00097-7.

[6] A. Mujumdar and V. Vaidehi, "Diabetes Prediction using Machine Learning Algorithms," *Procedia Computer Science*, vol. 165, pp. 292–299, Jan. 2019, https://doi.org/10.1016/j.procs.2020.01.047.

[7] B. Farajollahi, M. Mehmannavaz, H. Mehrjoo, F. Moghbeli, and M. J. Sayadi, "Diabetes diagnosis using machine learning," *Frontiers in Health Informatics*, vol. 10, no. 1, Feb. 2021, Art. no. 65.

[8] R. Katarya and S. Jain, "Comparison of Different Machine Learning Models for diabetes detection," in *2020 IEEE International Conference on Advances and Developments in Electrical and Electronics Engineering (ICADEE)*, Coimbatore, India, Dec. 2020, pp. 1–5, https://doi.org/10.1109/ICADEE51157.2020.9368899.

[9] M. Warke, V. Kumar, S. Tarale, P. Galgat, and D. J. Chaudhari, "Diabetes Diagnosis using Machine Learning Algorithms," *International Research Journal of Engineering and Technology*, vol. 06, no. 03, 2019.

[10] T. Zhu, K. Li, P. Herrero, and P. Georgiou, "Deep Learning for Diabetes: A Systematic Review," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 7, pp. 2744–2757, Jul. 2021, https://doi.org/10.1109/JBHI.2020.3040225.

[11] A. H. Syed and T. Khan, "Machine Learning-Based Application for Predicting Risk of Type 2 Diabetes Mellitus (T2DM) in Saudi Arabia: A Retrospective Cross-Sectional Study," *IEEE Access*, vol. 8, pp. 199539–199561, 2020, https://doi.org/10.1109/ACCESS.2020.3035026.

[12] A. Sarwar, M. Ali, J. Manhas, and V. Sharma, "Diagnosis of diabetes type-II using hybrid machine learning based ensemble model,"

International Journal of Information Technology, vol. 12, no. 2, pp. 419–428, Jun. 2020, https://doi.org/10.1007/s41870-018-0270-5.

[13] H. Lai, H. Huang, K. Keshavjee, A. Guergachi, and X. Gao, "Predictive models for diabetes mellitus using machine learning techniques," *BMC Endocrine Disorders*, vol. 19, no. 1, Oct. 2019, Art. no. 101, https://doi.org/10.1186/s12902-019-0436-6.

[14] A. S. Alenizi and K. A. Al-karawi, "Machine Learning Approach for Diabetes Prediction," in *Proceedings of Eighth International Congress on Information and Communication Technology*, London, UK, 2024, pp. 745–756, https://doi.org/10.1007/978-981-99-3043-2_61.

[15] G. Bhuvaneswari and G. Manikandan, "A novel machine learning framework for diagnosing the type 2 diabetics using temporal fuzzy ant miner decision tree classifier with temporal weighted genetic algorithm," *Computing*, vol. 100, no. 8, pp. 759–772, Aug. 2018, https://doi.org/10.1007/s00607-018-0599-4.

[16] L. Math and R. Fatima, "Adaptive machine learning classification for diabetic retinopathy," *Multimedia Tools and Applications*, vol. 80, no. 4, pp. 5173–5186, Feb. 2021, https://doi.org/10.1007/s11042-020-09793-7.

[17] T. Sharma and M. Shah, "A comprehensive review of machine learning techniques on diabetes detection," *Visual Computing for Industry, Biomedicine, and Art*, vol. 4, no. 1, Dec. 2021, Art. no. 30, https://doi.org/10.1186/s42492-021-00097-7.

[18] U. M. Butt, S. Letchmunan, M. Ali, F. H. Hassan, A. Baqir, and H. H. R. Sherazi, "Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications," *Journal of Healthcare Engineering*, vol. 2021, no. 1, 2021, Art. no. 9930985, https://doi.org/10.1155/2021/9930985.

[19] S. S. Bhat, V. Selvam, G. A. Ansari, M. D. Ansari, and M. H. Rahman, "Prevalence and Early Prediction of Diabetes Using Machine Learning in North Kashmir: A Case Study of District Bandipora," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, 2022, Art. no. 2789760, https://doi.org/10.1155/2022/2789760.

[20] E. Dugas, J. Jared, and W. Cukierski, "Diabetic Retinopathy Detection." https://kaggle.com/diabetic-retinopathy-detection.