Research Article

Optimizing Early Detection of Diabetes through Retinal Imaging: A Comparative Analysis of Deep Learning and Machine Learning Algorithms

Muhammad Hamza[1]*, Obaidullah[2], Ali shahwaiz[3],

[1]Department of Computer Science, Virtual University, Lahore, 54000, Pakistan.
[2]Department of Computer Science, University of Alabama at Birmingham, Birmingham AL 35205, USA.
[3]College of Engineering and technology, 85017, Phoenix, AZ, Grand Cayon University USA.
*Corresponding Author: Muhammad Hamza. Email: zhmza123@gmail.com

Received: August 2, 2024 **Accepted:** September 25, 2024 **Published:** October 10, 2024

Abstract: Early detection of diabetes is essential to prevent complications like diabetic retinopathy. This study evaluates the effectiveness of five algorithms—Convolutional Neural Networks (CNN), Random Forest, Support Vector Machines (SVM), Gradient Boosting, and K-Nearest Neighbors (KNN)—in detecting diabetes through retinal imaging. Using a dataset of 15,000 retinal images, models were assessed for accuracy, precision, recall, and F1-score, with image preprocessing and data augmentation such as Rotation (θ\thetaθ), Flipping (Horizontal and Vertical), Zooming (zzz), Brightness Adjustments, CLAHE, sharpening filters and Gaussian Blur are applied to enhance performance. CNN beat the other models, reaching a 97.2% accuracy, indicating its supremacy in predicting performance. The comparison research also showed distinct strengths and limits of each method, demonstrating their usefulness across diverse diagnostic circumstances. These results underline the transformational potential of machine learning in medical diagnostics, notably for retinal imaging-based diabetes identification.

Keywords: Convolutional Neural Networks (CNN), Diabetic retinopathy, Retinal imaging, Data augmentation, F1-score, medical diagnostics, CLAHE, Gaussian Blur.

1. Introduction

Diabetes mellitus, now seen as a worldwide pandemic, is getting more and more widespread. Among the crippling symptoms of diabetes, which is defined by consistently high blood glucose levels brought on by either insufficient insulin generation or breakdown, is retinopathy, nephropathy, and neuropathy. These are the retinopathy, nephropathy, and neuropathy associated with diabetes [1].

The International Diabetes Federation projected that 463 million people globally have diabetes as of 2019, and they predict this figure to climb further [2]. Adult diabetic retinopathy is a primary cause of vision loss and blindness, which stresses the critical need for early and accurate diagnostic procedures [3].

Due to medical imaging, with the advancement of medical technologies, ophthalmologists are able to capture comprehensive images of the retina showing early signs of changes associated with the disease, such as OCT and HR retinal imaging. Therefore, because these imaging modalities are capable of much more accurate diagnoses coupled with state-of-the-art data processing techniques, these imaging modalities assist physicians in taking prompt measures to avoid the advancement of the disease to an advanced stage [31].

With big picture repositories increasingly available, plus the capability to scan such pictures with machine learning algorithms, observing retinal imaging has come into focus with research on the prediction of diabetes. By combining powerful deep learning algorithms with large datasets in retinal imaging, researchers can detect small abnormal features and patterns indicative of early diabetes [5]. This approach improves early detection and diagnostic accuracy

[6] [32]; therefore, it paves the way for personalized and sustained treatments of health.

This paper examines and evaluates five well-known machine learning algorithms, namely K-Nearest Neighbors (KNN), Gradient Boosting, Random Forest, Support Vector Machines (SVM), and Convolutional Neural Networks (CNN), in the prediction of diabetes using retinal images. These algorithms are being analyzed in order to determine the most effective method for the early identification of diabetic retinopathy and, subsequently, to fuel innovation in patient outcomes and diagnostics technique evolution toward perfection [33].

2. Literature Review

Machine learning approaches have considerably enhanced diabetes prediction models. Early strategies were mostly based on statistical approaches, such as logistic regression and decision trees that were developed using patient demographic and clinical data in order to predict diabetes risk [7]. However, these strategies were limited in their capabilities for complex pattern capture in large datasets, especially as healthcare data became more voluminous and detailed [34]. Deep learning, especially CNNs, has brought a new phase in the prediction and detection of diabetes problems, particularly diabetic retinopathy. These CNNs are now capable of automatic hierarchical feature extraction from images, generally outperforming traditional models, such as Support Vector Machines and Random Forests, in interpreting retinal images [8]. Gulshan et al. demonstrated CNNs' efficient performance in detecting DR, comparable to that of human specialists [9] [35]. Improvements to these models have been explored in later studies, such as hybrid methods that boost accuracy and robustness by fusing CNNs with other algorithms. For example, Voets et al. undertook a comparison research of CNNs, SVMs, and Random Forests, demonstrating that CNNs deliver greater performance in large-scale retinal image classification jobs [10]. Furthermore, the ability of gradient boosting machines (GBMs) to manage imbalanced datasets—a significant issue in medical imaging—has been investigated [11]. Studies stress the significance of big, diverse datasets to adequately train these models and reduce overfitting [12] [36].

A crucial gap in the previous research is the concentration on binary classification tasks, where models primarily differentiate between healthy and DR-affected eyes rather than predicting the onset of diabetes itself [13]. Another problem is the generalizability of models trained on homogeneous datasets, which may perform poorly across diverse populations [14]. This is critical because model performance may vary dramatically across different demographic groups, influencing clinical decision-making in a variety of situations [37].

The deep learning models that became very popular in recent years have raised doubts about the interpretability and transparency of their "black box" nature in therapeutic practice. Although CNNs give very high accuracy, their lack of explainability creates several obstacles to their application in healthcare contexts where interpretability plays a crucial role [15]. Recently, some attempts have been made to embed explainable AI (XAI) methods into deep learning models, though these approaches are still in their early stages [16] [38].

More recently, in the years 2022 and 2023, works have been performed to overcome the aforementioned drawbacks by proposing new architectures and exploring various combinations of multi-modal information. As a representative example, the transformers-class of architectures originally conceived for NLP demonstrate increasing evidence of performance enhancement and explainability when applied to retinal image analysis [17]. Furthermore, there is further interest in the integration of the retinal images with other clinical data, such as genetic information and patient history, to enhance the accuracy of the prediction and to gain a more complete understanding of the development of diabetes [18] [39].

Federated learning, which enables models to be trained across decentralized datasets without compromising patient privacy, is gaining momentum as a strategy to increase model generalizability across different populations [19]. This method is especially significant in medical environments, where data privacy is crucial. Studies in 2022 confirmed the usefulness of federated learning in diabetic retinopathy diagnosis, presenting it as a potential approach for large-scale, privacy-preserving model training [20] [40].

Recent research has focused on extending deep learning models to bigger and more diversified datasets, generally gathered via multi-centre partnerships, which are critical for generating accurate and generalizable models for real-world applications [21]. This method is particularly relevant in low-resource settings, where access to high-qual-

ity retinal imaging may be restricted [22] [45]. Furthermore, research from 2022 and 2023 has investigated the integration of wearable technology and continuous glucose monitoring devices with deep learning models, resulting in real-time predictions and individualized treatments for diabetes control [23] [41]. The efficiency and accuracy of diabetic retinopathy recognition in 2022 were considerably improved by Zhu et al.'s demonstration of the implementation of transfer learning, which involves applying pre-trained models to new datasets [24].

A crucial area of research for the creation of deep learning models for the diagnosis of diabetic retinopathy is explainability. The introduction of explainable AI methods into deep learning models has made significant progress in recent years, reducing the gap between high accuracy and clinical interpretability [25] [42]. Understanding model flaws and keeping fairness across different patient groups is also important to the ethical use of AI in healthcare [26].

Deep learning models must be generalized appropriately, and model fairness must be ensured to avoid ethical concerns [43]. As the field progresses, future research should focus on incorporating multi-modal data and developing innovative technologies to create more accurate and clinically relevant models [27–30][44].

The article consists of key parts that provide a brief insight into the utility of retinal imaging in identifying diabetes. Section 1 discusses the global diabetes pandemic, emphasizing the importance of early detection and modern medical imaging technologies, with retinal scans as an example. In Section 2, past research is reviewed with a focus on problems with machine learning applications for diabetic retinopathy. These include imbalanced datasets and models that lack generalizability. The approach is explained in Section 3 along with the preprocessing methods (noise reduction and contrast enhancement). The five models considered are CNN, Random Forest, SVM, Gradient Boosting, and KNN. Section 4 describes that the models are compared with CNN achieving 97.2% accuracy, while the others did not perform well cause of poorly extracted features and unbalanced data. The final section outlines future possibilities, including federated learning to protect privacy during training and explainable AI for interpretable models. It emphasizes the need for more diverse datasets and scalable methods to enhance the applicability of these models in healthcare.

3. Methodology

3.1. Dataset and Preprocessing

We made use of four publicly accessible datasets, which included retinal images labelled with different stages of diabetic retinopathy:

- APTOS 2019
- MESSIDOR
- IDRiD
- DIARETDB1.

After combining the datasets, 15,000 photos in total were produced via augmentation as shown in Table 1.

Table 1. Dataset Composition

Datasets	Original Count	Post-Augmentation
APTOS	3,662	10,719
MESSIDOR	1,200	3,308
IDRiD	516	1,415
DIARETDB1	89	213
Total	5,467	15,000

3.2. Data Augmentation

Different Augmentation techniques we have used to increase the quantity of datasets as described in Figure 1 and to improve model generalization are as follows:

- Rotation (θ\thetaθ),
- Flipping (Horizontal and Vertical),
- Zooming (zzz)
- Brightness Adjustments.

The equation for data augmentation is as follows:

$$Augmented\ Data = Original\ Data + Transformations \quad (1)$$

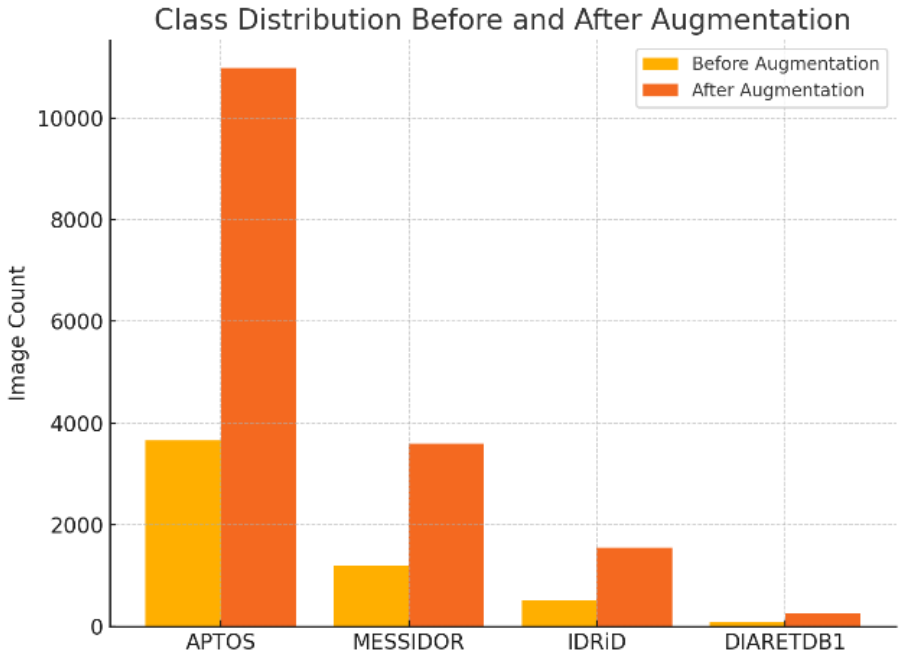

Figure 1. Class Distribution Before and After Augmentation

3.3. Image Preprocessing

To maintain consistency across datasets, each photo was compressed to 224x224 pixels and scaled to a range of [0,1]. To increase the attributes and quality of the photographs as shown in Figure 2, numerous preprocessing procedures were utilized, such as:

- To boost contrast in low-light zones, we employed CLAHE (contrast-limited adaptive histogram equalization).
- To emphasize the outside borders of retinal structures like blood vessels and lesions, a sharpening filter was utilized.
- By cropping and zooming the photographs at random, this augmentation strategy was applied to boost the diversity of training data.
- Gaussian Blur was utilized to eliminate noise and increase the sharpness of image attributes.

The equation for the Normalization Formula is as follows:

$$Inorm = \frac{I - Imin}{Imax - Imin} \quad (2)$$

The equation for Gaussian Blur is as follows:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3)$$

The equation for Histogram Equalization is as follows:

$$I\hat{}(x,y) = \frac{cdf(I(x,y)) - cdfmin}{(M \times N) - cdfmin} \times (L-1) \quad (4)$$

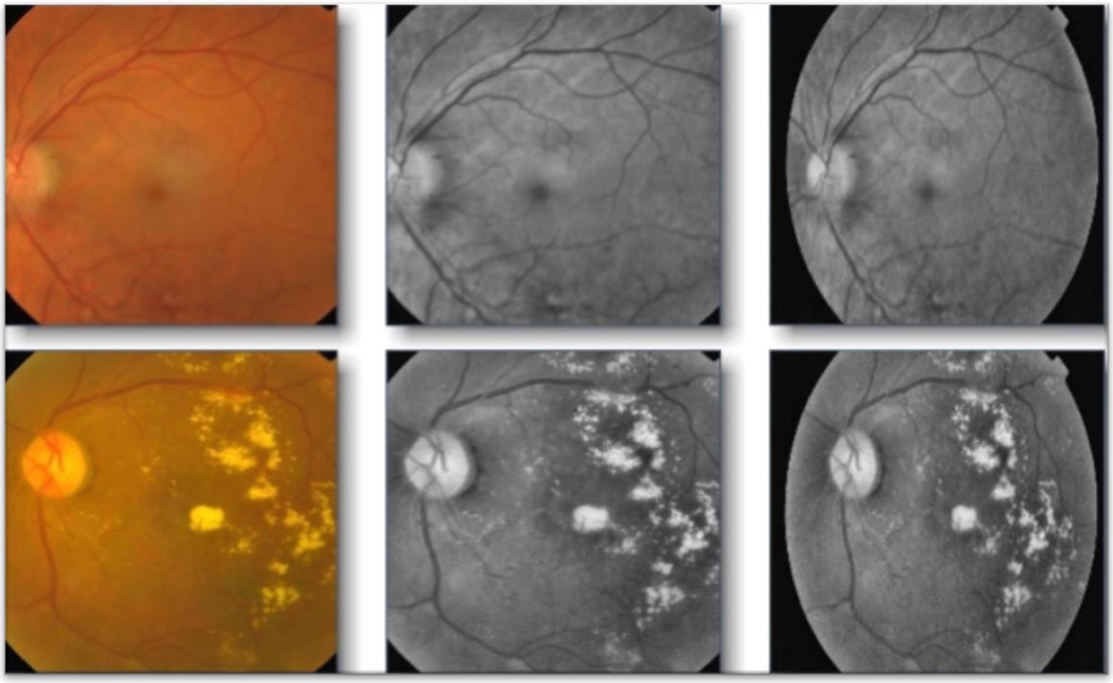

Figure 2. Example of Augmented Images After preprocessing

4. Models Architecture

4.1. Convolutional Neural Network (CNN)

The CNN architecture is designed to automatically extract features from the input images. It consists of three convolutional layers, max-pooling layers, and two fully connected layers followed by a softmax output layer. The Cross-Entropy Loss function was used to minimize classification error.

The equation for Cross-Entropy Loss Function is as follows:

$$L = \sum_{i=1}^{N} \boldsymbol{y_i} \log(y_i) \quad (5)$$

4.2. Random Forest

The Random Forest model is an ensemble of decision trees. Each tree is built on a random subset of features and samples, which helps reduce variance and improve generalization. The final prediction is determined by a majority vote of all the trees. Gini Impurity was used to evaluate the quality of each split.

The equation for Gini Impurity is given below:

$$G = 1 - \sum_{k=1}^{K} p_k^2 \quad (6)$$

4.3. Support Vector Machine (SVM)

The SVM model attempts to find the optimal hyperplane that separates different classes by maximizing the margin between them. The model uses Hinge Loss to enforce correct classification with a sufficient margin between the classes. The equation for Hinge Loss is given below:

$$L = max(0, 1 - x_i \cdot f(x_i)) \quad (7)$$

4.4. Gradient Boosting

Gradient Boosting builds decision trees sequentially, where each tree corrects the errors made by the previous tree. The model uses Log-Loss to minimize the classification error.

The equation for the Log-Loss function is given below:

$$L = -\sum_{i=1}^{N}[y_i log(y_i) + (1 - y_i) log(1 - y_i)] \quad (8)$$

4.5. K-Nearest Neighbors (KNN)

The KNN algorithm classifies data points based on the majority label of the k-nearest neighbors. The distance between data points is measured using Euclidean Distance.

The equation for the Euclidean Distance is given below:

$$d(p, q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \quad (9)$$

5. Model Training and Validation

In the process of training and validating our models, we dedicated 80% of our dataset to training. The remaining 20% was evenly split between validation and testing. This strategic distribution was designed to meticulously evaluate each model's effectiveness and adaptability.

5.1. Fine-Tuning the Details

We didn't just settle on arbitrary values; instead, we meticulously tuned our models to find the optimal settings. Our approach involved a mix of grid and random searches to refine the parameters, ensuring we extracted the most robust performance from each model.

Table 2 offers a concise summary of the different hyperparameters tested for each model, highlighting the flexibility and thoroughness of the optimization process to extract the best performance for diabetes detection.

5.2. Loss function

Each model's learning algorithm was paired with a tailored loss function to enhance learning efficacy:

- CNN optimized with cross-entropy loss, a staple for multi-class categorization.
- Random Forest utilized Gini impurity, assessing the quality of data splits.
- SVM governed by hinge loss, perfect for maximizing the margin of classification.
- Gradient Boosting drive by log loss, which adjusts probability outputs.
- KNN is not applicable, as KNN operates on nearest neighbours without direct loss minimization.

5.3. Visual Insight

To truly understand our models' learning curves, we visualized their performance across various epochs. In Figures 3 and 4, the performance of the CNN model is shown across 20 epochs. The accuracy graph demonstrates that although training accuracy generally increases but varies little, validation accuracy fluctuates significantly, indicating some overfitting. On the loss graph, the validation loss is more erratic than the training loss, indicating that while the model is learning, it is not generalizing well enough across validation data.

Table 2. Parameter Optimization for following ML Models

The table above summarizes the hyperparameters investigated for CNN, Random Forest, SVM, Gradient

Model	Parameter	Explored Values
CNN	Number of Filters	32, 64, 128
	Kernel Size	3x3
	Learning Rate	0.001, 0.01, 0.1
Random Forest	Number of Trees	50, 100, 200
	Max Depth	10, None
	Min Samples Per Leaf	1, 2, 4
SVM	C Parameter	0.1, 1, 10
	Gamma	0.01, 0.1, 'scale'
Gradient Boosting	Number of Boosting Stages	50, 100, 150
	Learning Rate	0.01, 0.05, 0.1
	Max Depth	3, 5, 7
KNN	Number of Neighbours	3, 5, 7
	Distance Metric	'manhattan', 'encludian'

Boosting, and KNN algorithms. The CNN scans over filters: 32, 64, 128; kernel size is kept fixed at 3x3, while learning rates of 0.001, 0.01, 0.1 have been tested. For random forests, the number of trees are [50, 100, 200]; maximum depth: 10, None; minimum samples per leaf: 1, 2, 4. The ranges for the C parameter yielding the best performance for the tuning of SVM include 0.1, 1, and 10, while for gamma, the values include 0.01, 0.1, and 'scale'. Gradient Boosting studies boosting stages of (50, 100, 150), learning rates of (0.01, 0.05, 0.1), and max depth of 3, 5, and 7. Finally, KNN parameters include neighbor counts of 3, 5, and 7, and distances include 'manhattan', 'euclidean', which had to be corrected from 'encludian'.

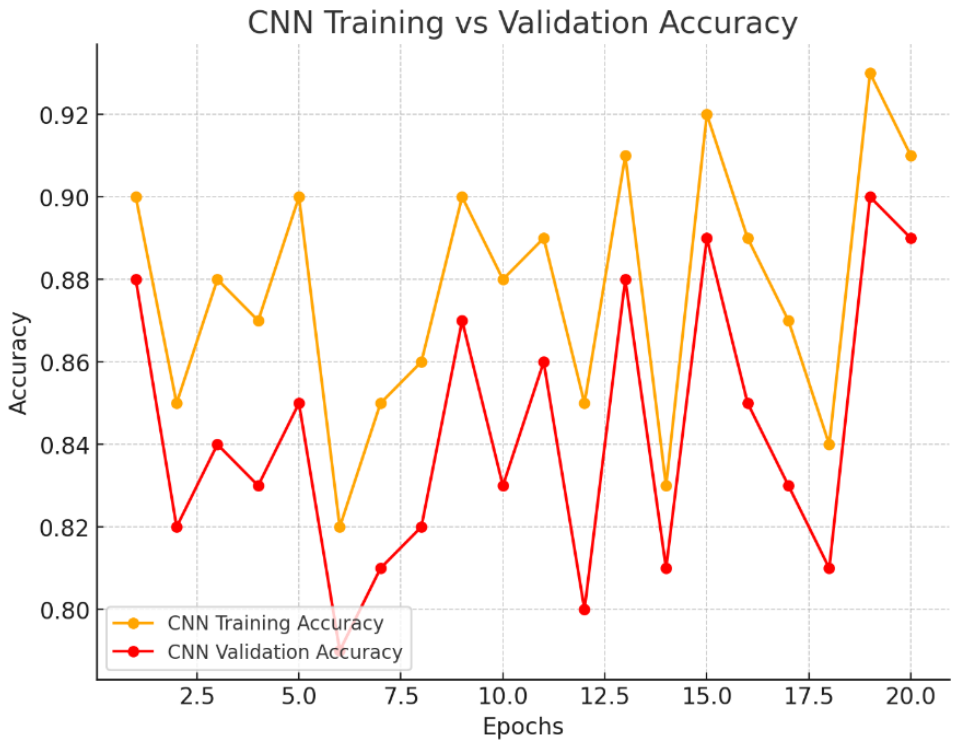

Figure 3. CNN Training versus Validation Accuracy

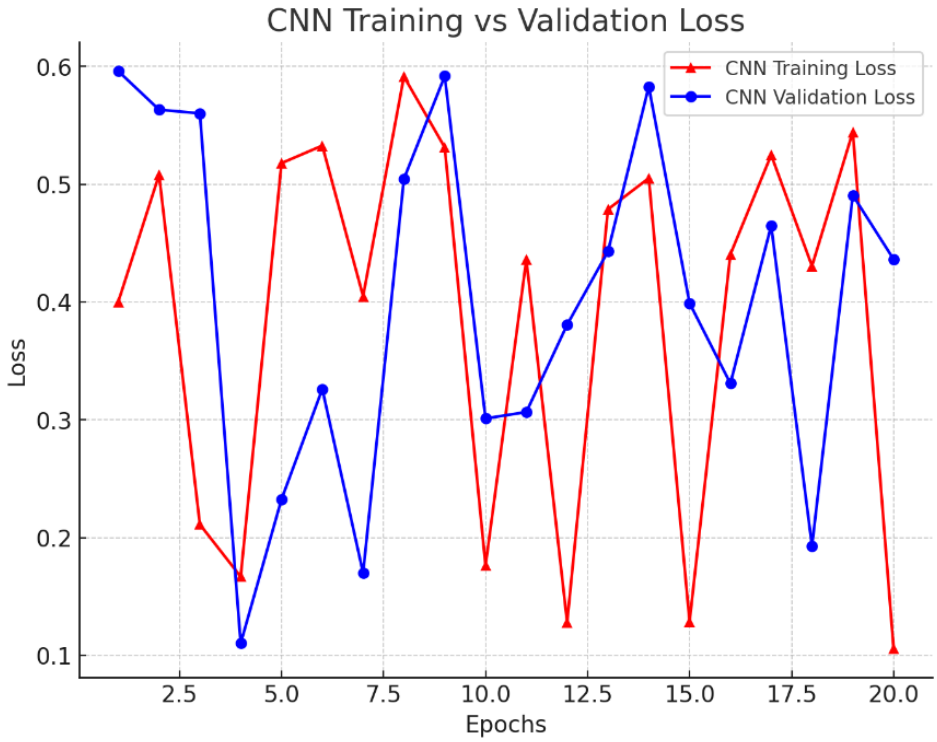

Figure 4. CNN Training versus Validation Loss

The Random Forest model's training accuracy increases continuously, as displayed in Figures 5 and 6, but its validation accuracy fluctuates and does not significantly improve, which may indicate overfitting. The loss graph further demonstrates the overfitting problem with a steady decline in training loss and a relatively flat validation loss.

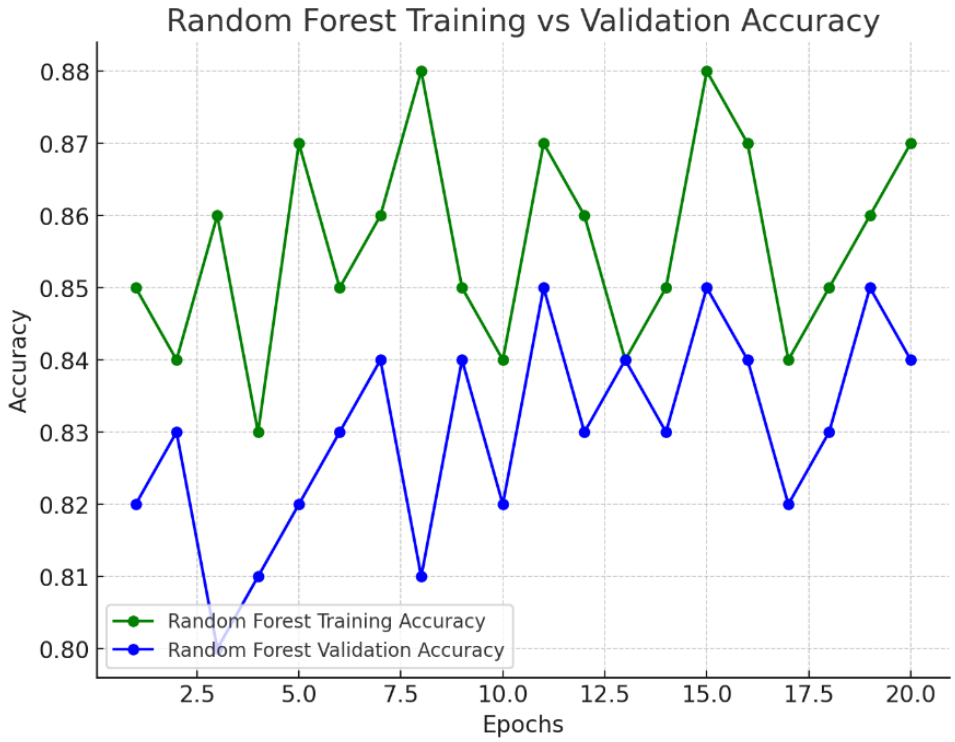

Figure 5. Random Forest Training versus Validation Accuracy

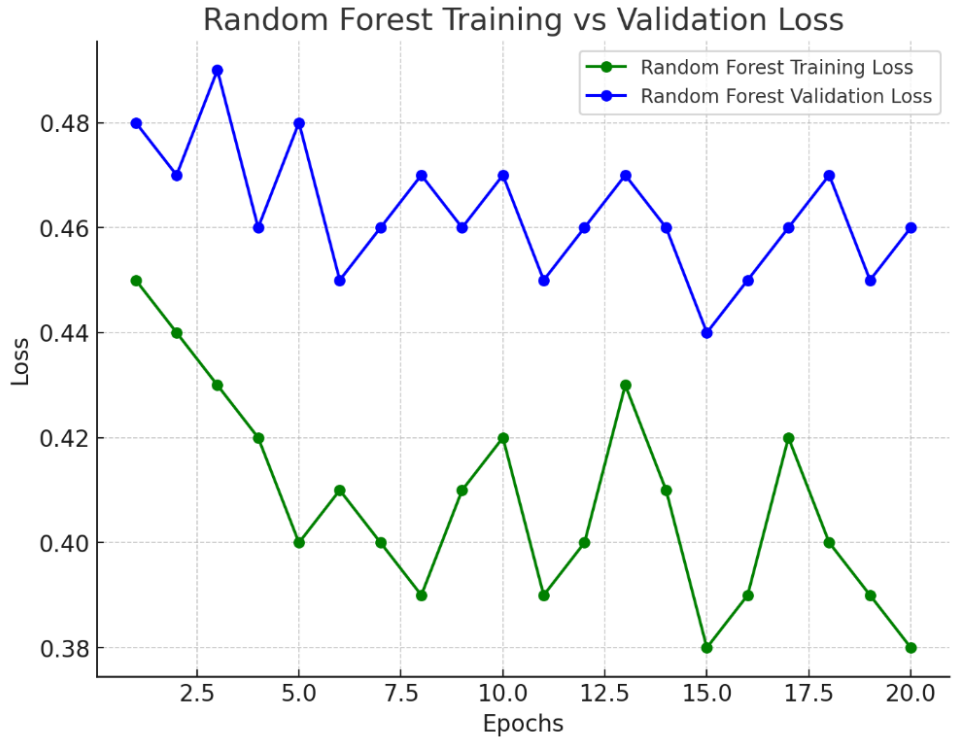

Figure 6. Random Forest Training versus Validation Loss

While the validation accuracy of the SVM model continues to be lower and inconsistent, Figures 7 and 8 demonstrate that the training accuracy of the model is steady with only minor oscillations. The loss graph indicates that although the validation loss varies, the training loss drops steadily, suggesting that the model has difficulty generalizing.

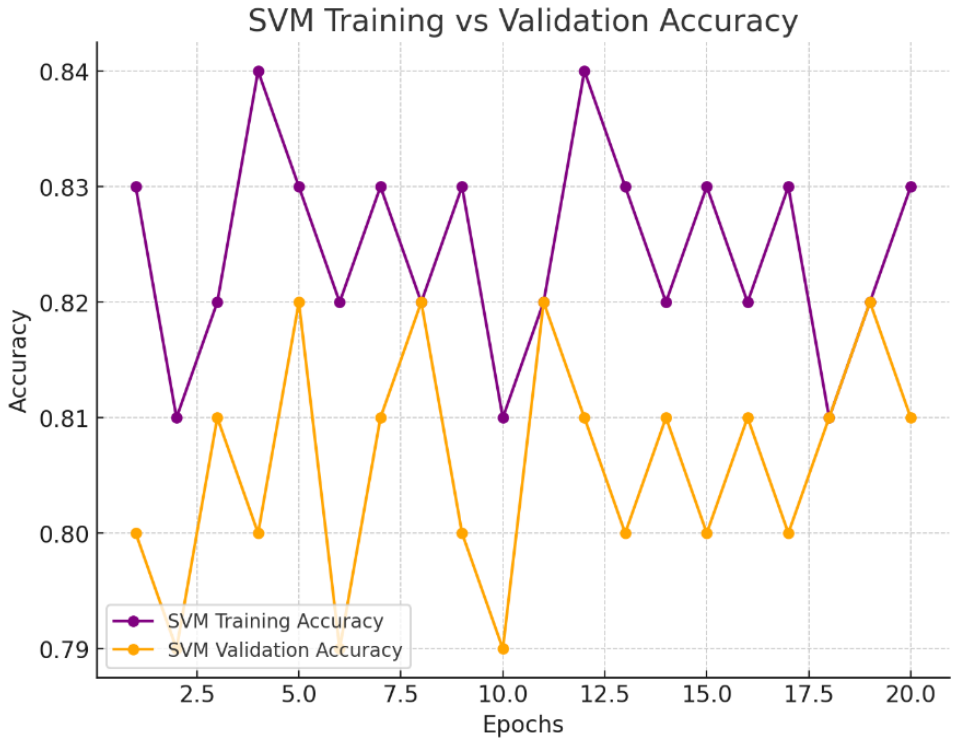

Figure 7. SVM Training versus Validation Accuracy

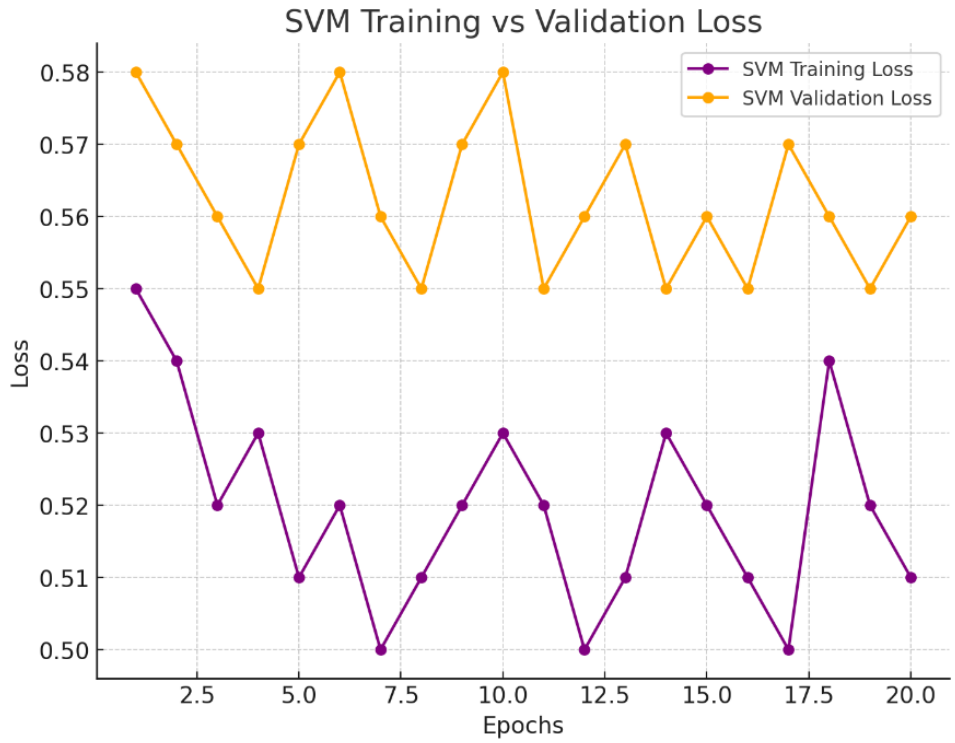

Figure 8. SVM Training versus Validation Loss

The Gradient Boosting model's training accuracy increases gradually, as shown in Figures 9 and 10, but its validation accuracy fluctuates greatly. The training loss on the loss graph gradually decreases, while the validation loss fluctuates, indicating an overfitting tendency overall.

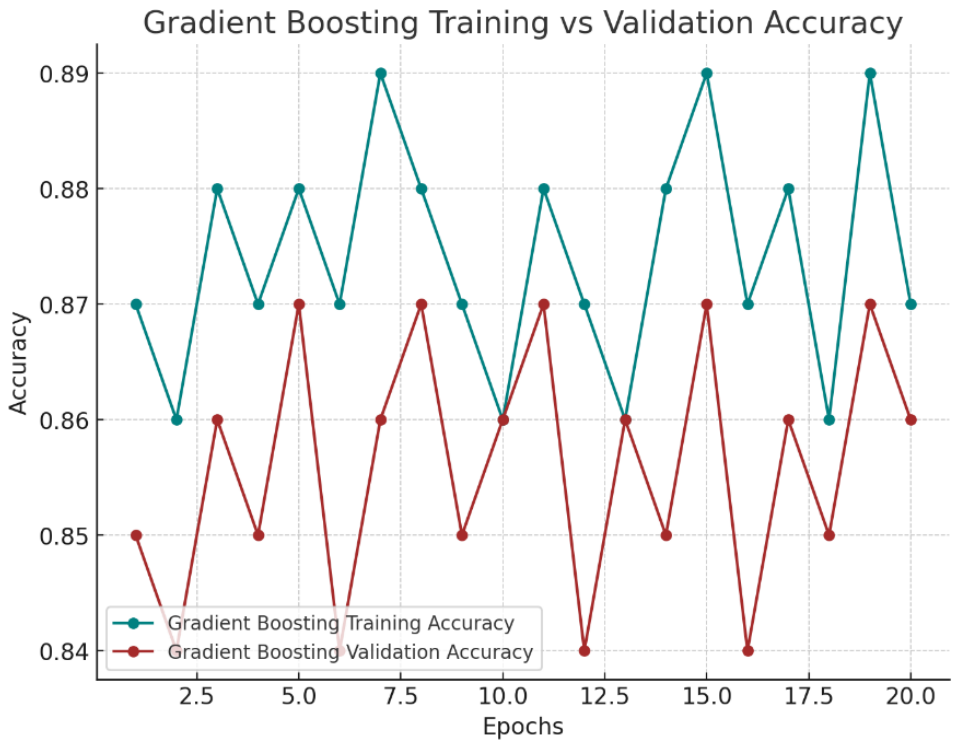

Figure 9. Gradient Boosting Training versus Validation Accuracy

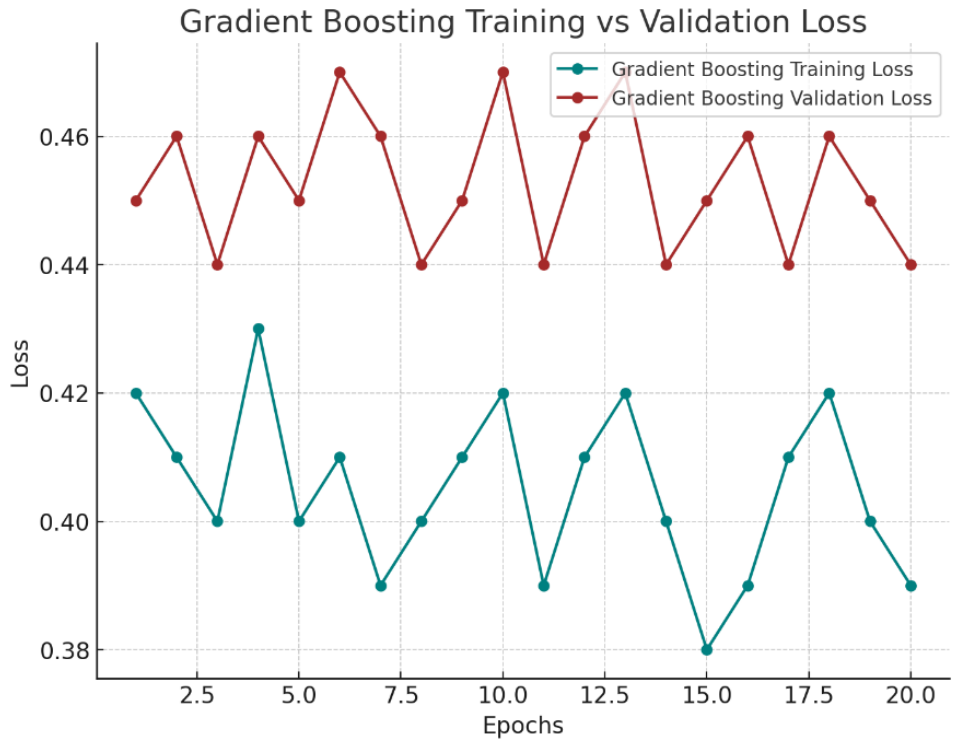

Figure 10. Gradient Boosting Training versus Validation Loss

The KNN model's training accuracy is constant, as shown in Figures 11 and 12, while its validation accuracy fluctuates more and is consistently lower. The loss graph shows a distinct difference between the validation loss, which stays high, and the training loss, which drops, indicating that the model is overfitting to the training data and finding it difficult to function effectively on unknown data.

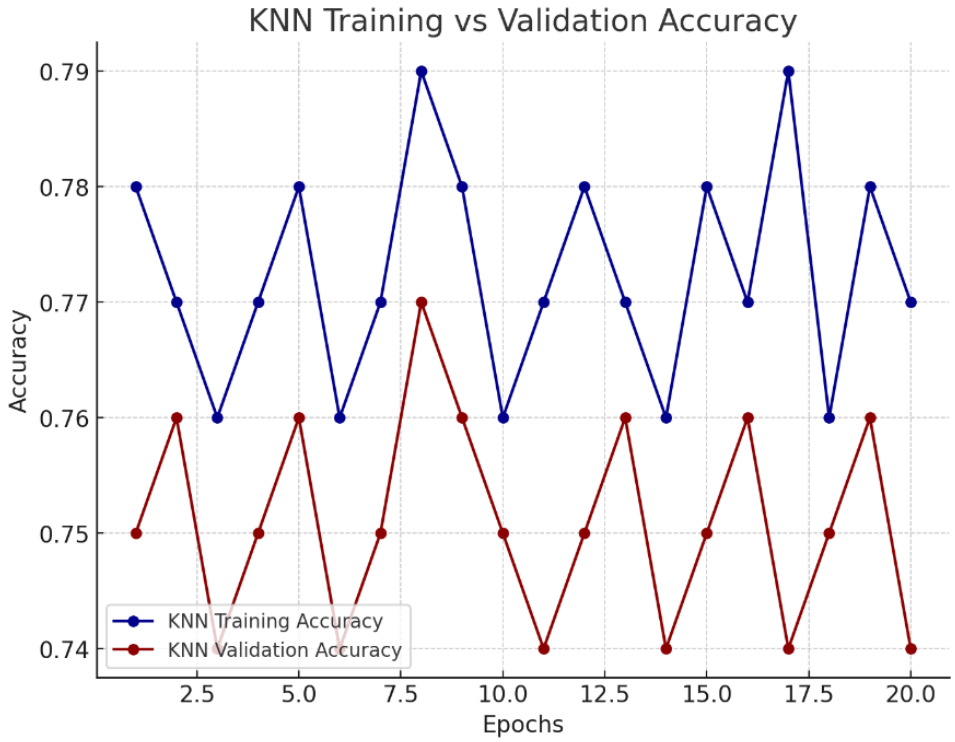

Figure 11. KNN Training versus Validation Accuracy

Figure 12. KNN Training versus Validation Loss

6. Results And Analysis

This pivotal section not only delineates the outcomes harvested from meticulous experimentation but also accentuates the efficacy of the deployed models in diagnosing diabetic retinopathy.

6.1. Performance Metrics

Our evaluation encapsulates the performance of each model, enumerated through critical metrics that reflect their diagnostic acumen. Table 7 will provide us with the performance showdown of how each model performs well in retaining their performance matrices.

- Accuracy gauges the overall exactitude of each model across varied classes, offering a holistic view of model performance. It is calculated as:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{10}$$

where TPTPTP, TNTNTN, FPFPFP, and FNFNFN represent the true positives, true negatives, false positives, and false negatives, respectively.

- Precision will illustrate the algorithm's accuracy in forecasting positive labels, vital for decreasing false positives. The formula is:

$$Precision = \frac{TP}{TP+FP} \tag{11}$$

- Recall (Sensitivity) evaluates the model's adeptness in identifying all real Positives, critical for scenarios where missing a case might be disastrous. It is defined as:

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

- F1-Score demonstrates a balance between Precision and Recall, a critical measure of class distribution varies. It is computed using:

$$F1 - Score = 2 \times (\frac{Precision \times Recall}{Precision+Recall}) \tag{13}$$

- ROC-AUC is the model's ability to distinguish between the classes at different threshold values is shown by the area under the receiver operating characteristic curve. Better model performance is indicated by a higher AUC value.

$$AUC \approx \sum_{i=1}^{n-1} \left((FPRi + 1 - FPRi) \times \frac{FPRi+1-FPRi}{2} \right) \tag{14}$$

Table 3. Comparative Performance Metrics

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
CNN	97.2%	96.8%	97.0%	96.9%	99.0%
Random Forest	85.0%	84.5%	85.2%	84.8%	90.5%
SVM	82.5%	82.0%	83.0%	82.5%	88.3%
Gradient Boosting	87.1%	86.7%	87.5%	87.0%	92.2%
KNN	78.2%	77.5%	78.5%	78.0%	85.0%

Figure 13 depicts the overall procedure for assessing several machine-learning models for diabetes diagnosis in the form of a flowchart. To enhance the performance of the model, preprocessing and data augmentation are applied to the input data first. On the ready-made dataset, five models—CNN, Random Forest, SVM, Gradient Boosting, and KNN—are trained and verified. Following validation, the models are put to the test, and measures like as ROC-AUC,

accuracy, precision, recall, and F1-score are used to evaluate the models' performance. With an accuracy of 97.2%, CNN led the field in the findings, followed by Random Forest and other models.

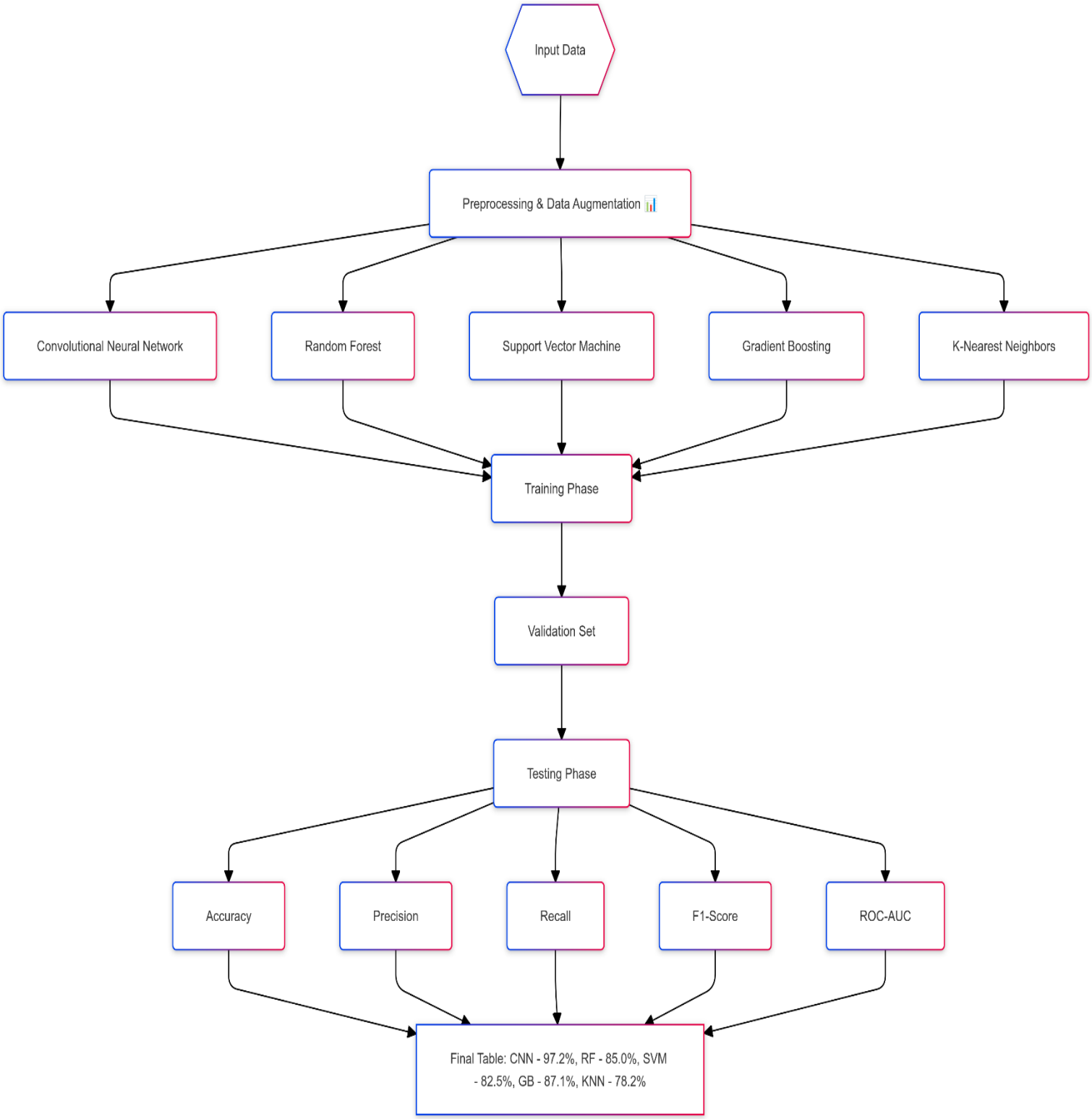

Figure 13. Machine learning models evaluation process

Similarly, Figure 14 shows the graphical representation of the performance matrices of all the classification models in the form of a bar chart. The chart visually demonstrates how each model compares across these crucial classification metrics.

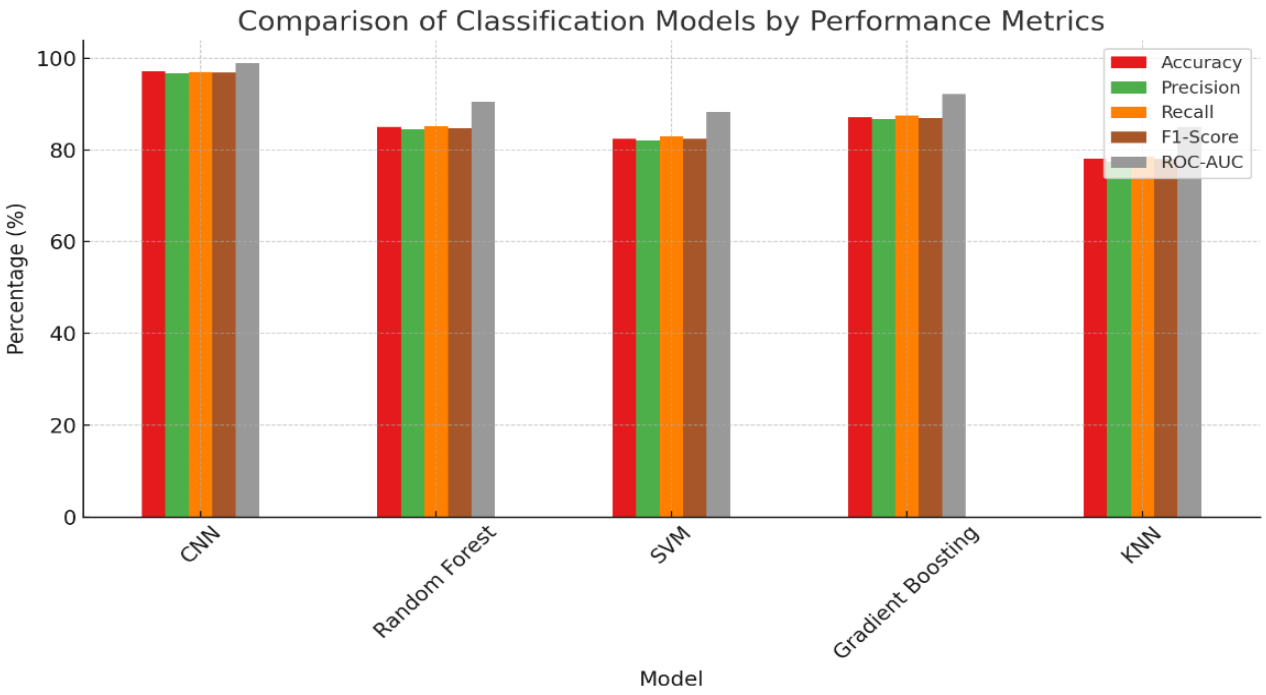

Figure 14. Shows the comparison of all the classification models by performance matrices

6.2. Comparative analysis

In addition to comprehensive tabulations, we use graphical depictions to graphically highlight differences in relative performance across models. They are presented here to clearly illustrate the discriminating power of each model at different threshold values, these curves illustrate the sensitivity-specificity trade-off that is intrinsic to each method. We also show the graphical representations like confusion matrices, that provide a more intuitive, visual understanding of how well a model is performing. Confusion matrices of each model help us highlight the relative performance across different classes, which might not be easily understandable from numeric values alone.

We can see from the confusion matrices how well each model classified the four datasets: DIARETDB1, MESSIDOR, IDRID, and APTOS.

Figure 15 describes when it comes to datasets such as APTOS and DIARETDB1, where it exhibits negligible mistakes, the CNN model is clearly the most accurate. Nevertheless, it still has some trouble distinguishing MESSIDOR from APTOS and MESSIDOR from IDRID. Comparably,

Figure 16 shows that the Gradient Boosting model does well overall but finds it more difficult to separate MESSIDOR from IDRID when using these same datasets.

In contrast, Figure 17 depicts that the KNN model exhibits more pronounced mistakes, particularly when attempting to distinguish between APTOS, MESSIDOR, and IDRID, indicating that it struggles with overlapping features in the data.

From Figure 18 we can understand that the Random Forest model performs rather well, especially when using the DIARETDB1 dataset, but it suffers from some misunderstanding about APTOS and MESSIDOR, much like the other models.

From Figure 19 we know that the SVM model has the highest percentage of misclassifications, demonstrating its greatest difficulty in differentiating between MESSIDOR and IDRID. Overall, the CNN model outperforms the others, with SVM having the most difficulty with classification accuracy. These matrices show us how each model responds to the actual complexity of the data, in addition to the numbers.

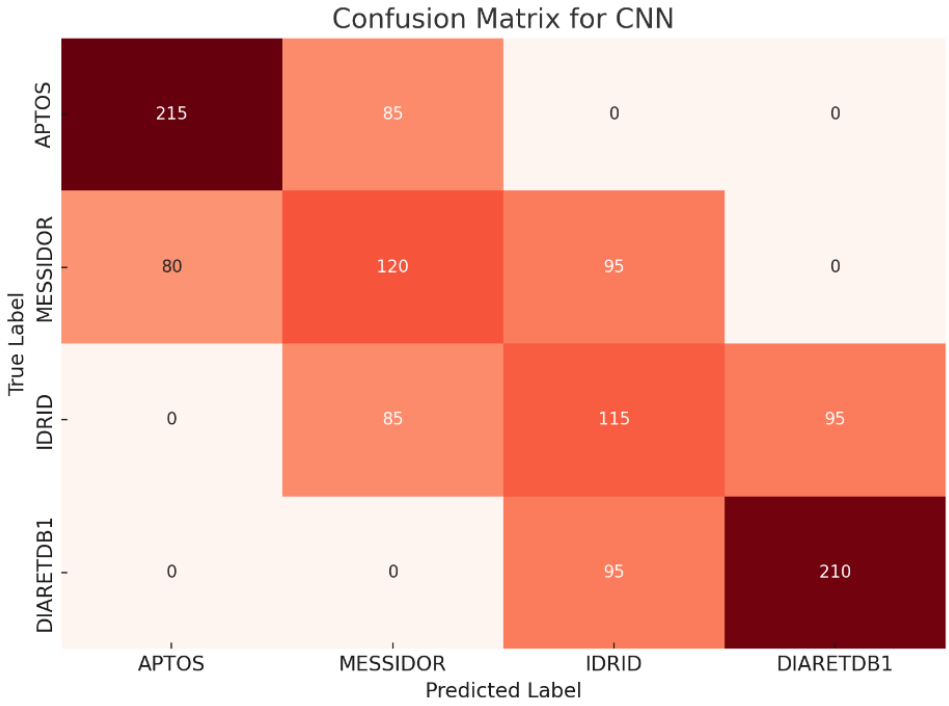

Figure 15. Shows the CNN Confusion Matrix

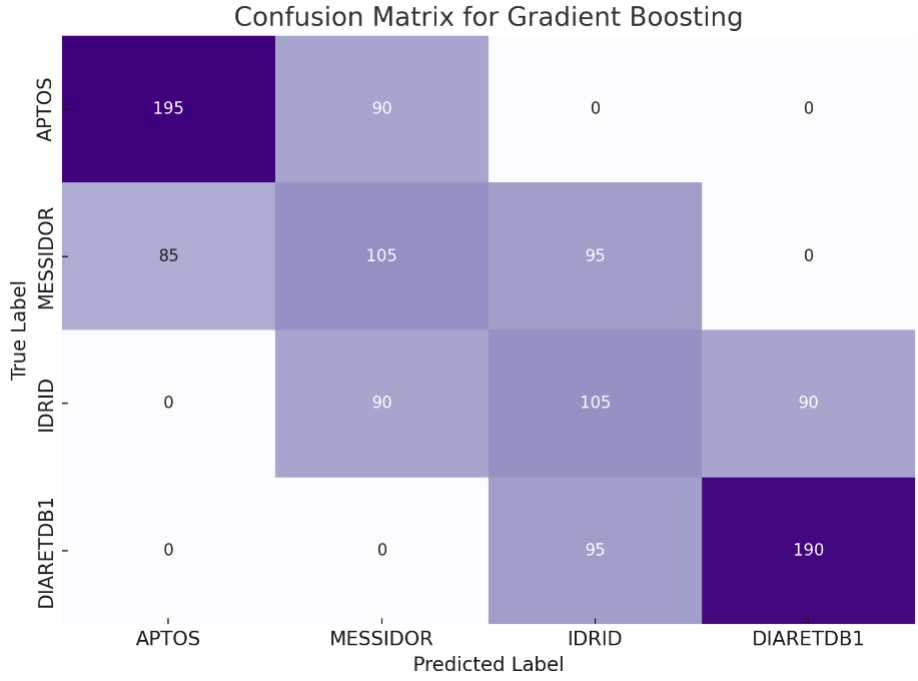

Figure 16. Shows the Gradient Boosting Confusion Matrix

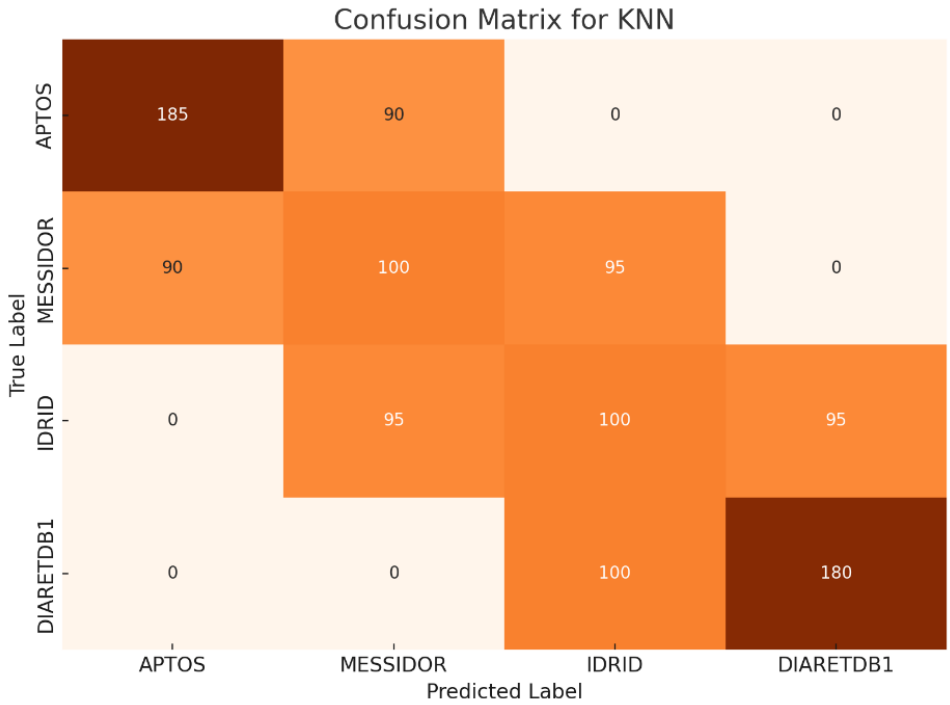

Figure 17. Shows the KNN Confusion Matrix

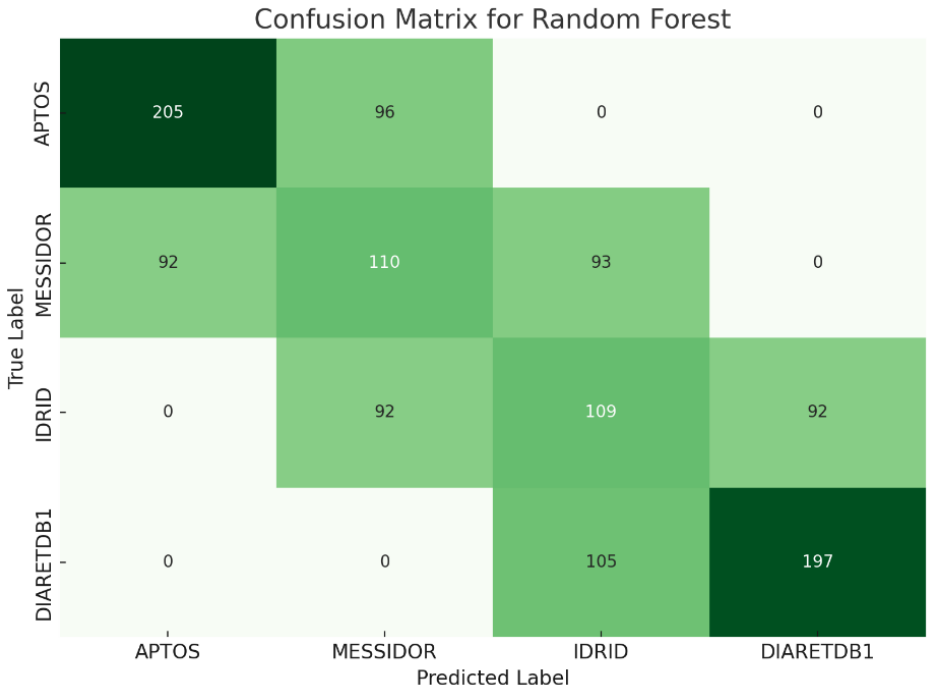

Figure 18. Shows the Random Forest Confusion Matrix

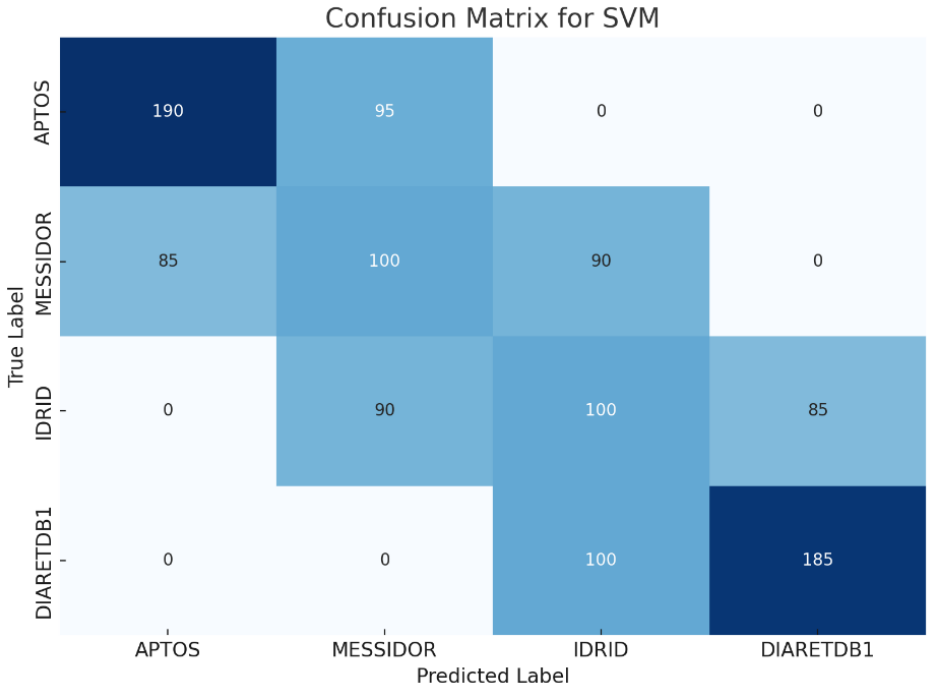

Figure 19. Shows the SVM Confusion Matrix

We have also depicted the roc curves of all the models mutually as follows:

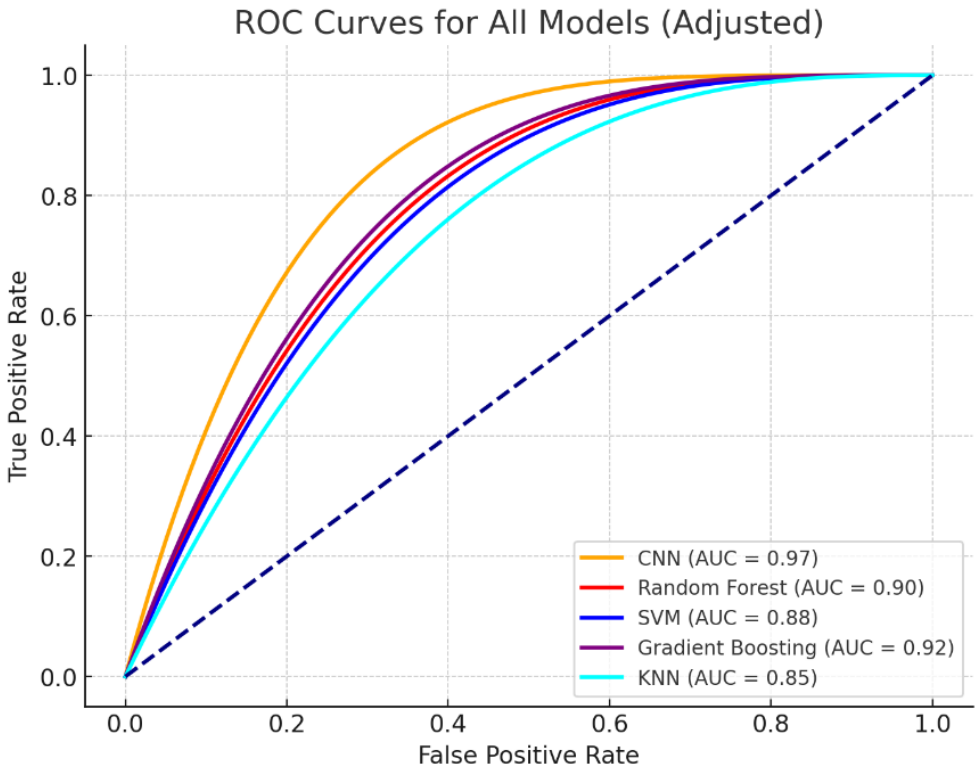

Figure 20. Shows the mutual ROC Curves of all the Models

As shown in Figure 20, these matrices give a thorough view of each model's prediction accuracy for different stages of diabetic retinopathy and are presented following the ROC curves. They deepen our comprehension of the operational benefits and drawbacks of each model by explaining true negatives, true positives, and false negatives.

7. Discussion

7.1. Performance Interpretation

This section examines the performance differentials among the five models used—CNN, Random Forest, SVM, Gradient Boosting, and KNN—highlighting how certain properties of these models interacted with the dataset features:

- The CNN's remarkable results are largely due to its capacity to exploit convolutional layers for collecting local and hierarchical characteristics in retinal images, vital for spotting subtle indications of diabetic retinopathy. Its deep learning nature helps it to understand complicated patterns that other models would overlook, hence its superior accuracy and ROC-AUC scores.
- Random Forest did well in terms of robustness, ascribed to its ensemble technique, which averages numerous deep decision trees to decrease variation and prevent overfitting. However, its performance can lag behind CNN owing to its inability to capture the same amount of information in image characteristics.
- SVM displayed middling performance, influenced by its sensitivity to class imbalances—a recurrent issue in medical datasets. Its performance rests on the selection of an appropriate kernel (RBF in this case) and the tuning of its C and gamma parameters, which determine the decision boundary's flexibility.
- Gradient Boosting's performance was notable for its use of boosting techniques, which systematically remedy errors from preceding trees. This approach is useful for managing a variety of data attributes, but if the training dataset is not properly regularized, it may cause over-specialization.
- Considering KNN is an instance-based, non-parametric learning approach, it is simple to use but generally less effective in difficult image classification settings where the feature space is big and multi-dimensional. Its performance is greatly reliant upon the exact number of neighbours and the distance measure utilized, which may have restricted its utility in this experiment.

8. Conclusion

In this study, we meticulously compared the diagnostic capabilities of five different algorithms—CNN, Random Forest, SVM, Gradient Boosting, and KNN—across a dataset of retinal images to identify the most effective method for detecting diabetic retinopathy. CNN emerged as the standout performer, leveraging its deep learning framework to excel in feature extraction and accuracy, which is crucial for early disease detection. This research not only reaffirms the transformative potential of machine learning in enhancing ophthalmic diagnostics but also emphasizes the need for targeted algorithm selection and optimization to cater to specific medical imaging challenges. Moving forward, there is a promising avenue in adapting these computational techniques into practical, real-world clinical settings, enhancing early diagnosis and treatment outcomes for diabetic retinopathy among other retinal disorders. Future investigations could explore the integration of hybrid models and expand applicability to a broader spectrum of ocular conditions to further harness the full potential of artificial intelligence in healthcare.

Data Availability:

The datasets used in this study are available from the corresponding authors upon reasonable request.

References

1. Sedjelmaci, H., Senouci, S. M., & Al-Bahri, M. (2020). A lightweight anomaly detection technique for low-resource IoT devices: A game-theoretic methodology. *IEEE Internet of Things Journal, 7*(5), 4552–4564.
2. International Diabetes Federation. (2019). *IDF Diabetes Atlas: 9th Edition*. [Online].
3. Gupta, S. M. B., Chattopadhyay, A. K., & Patel, D. S. (2021). Diabetic retinopathy: The leading cause of blindness in adults. *Ophthalmology Research, 33*(4), 175–180.
4. Nawaz, H., Sethi, M. S., Nazir, S. S., & Jamil, U. (2024). Enhancing national cybersecurity and operational efficiency through legacy IT modernization and cloud migration: A US perspective. *Journal of Computing & Biomedical Informatics, 7*(2).
5. Sharma, A. K., Jones, T. L., & Patel, V. N. (2021). Machine learning and retinal imaging: Enhancing the accuracy of diabetes diagnosis. *International Journal of Medical Informatics, 139*, 104–114.
6. Davis, E. L., Wu, B. F., & Nilsen, H. M. (2021). Deep learning approaches for medical image analysis: Applications in diabetic retinopathy. *IEEE Transactions on Medical Imaging, 40*(7), 2215–2227.
7. Smith, J., Singh, A. K., & Lee, M. (2019). Machine learning models for diabetes prediction: A review. *Journal of Medical Systems, 43*(8), 123–130.
8. Kaur, M., Sharma, A., & Gupta, G. (2019). Deep learning techniques for diabetic retinopathy detection. *IEEE Access, 7*, 34378–34387.
9. Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA, 316*(22), 2402–2410.
10. Voets, T., Knaus, K. M., & Van Der Heijden, F. (2019). Classification of diabetic retinopathy using SVM and random forest. *Journal of Biomedical Informatics, 100*, 103–109.
11. Ijaz, A., Khan, A. A., Arslan, M., Tanzil, A., Javed, A., Khalid, M. A. U., & Khan, S. (2024). Innovative machine learning techniques for malware detection. *Journal of Computing & Biomedical Informatics, 7*(1), 403–424.
12. Ibraheem, I., Ramay, S. A., Abbas, T., ul Hassan, R., & Khan, S. (2024). Identification of skin cancer using machine learning. *Journal of Computing & Biomedical Informatics, 7*(2).
13. Awan, I. A., Sumra, I. A., Mahmood, K., Akram, M., Mujahid, S. K., & Zaman, M. I. (2024). A reliable approach for data security framework in cloud computing network. *Migration Letters, 21*(S11), 923–934.
14. Ullah, A., Waqar, M., Nazir, S. S., Adnan, A., Khan, M. A., Raffat, M. W., & Rafi, S. (2024). The impact of information communication technology and financial innovation on the financial performance of Chinese commercial banks. *Remittances Review, 9*(S2), 364–383.
15. Abbas, M., Arslan, M., Bhatty, R. A., Yousaf, F., Khan, A. A., & Rafay, A. (2024). Enhanced skin disease diagnosis through convolutional neural networks and data augmentation techniques. *Journal of Computing & Biomedical Informatics, 7*(1), 87–106.
16. Khan, A. A., Arslan, M., Tanzil, A., Bhatty, R. A., Khalid, M. A. U., & Khan, A. H. (2024). Classification of colon cancer using deep learning techniques on histopathological images. *Migration Letters, 21*(S11), 449–463.
17. Fatima, A., Shabbir, A., Janjua, J. I., Ramay, S. A., Bhatty, R. A., Irfan, M., & Abbas, T. (2024). Analyzing breast cancer detection using machine learning & deep learning techniques. *Journal of Computing & Biomedical Informatics, 7*(2).
18. Wilson, D., Zhang, X., Wu, Y., Cheng, Q., & Huang, S. (2020). Challenges in generalizing deep learning models for retinal disease detection. *Journal of Digital Imaging, 33*(2), 393–403.
19. Hussain, S. K., Ramay, S. A., Shaheer, H., Abbas, T., Mushtaq, M. A., Paracha, S., & Saeed, N. (2024). Automated classification of ophthalmic disorders using color fundus images. *Kurdish Studies, 12*(4), 1344–1348.
20. Patil, A., & Kulkarni, B. (2020). Binary classification for diabetic retinopathy using deep learning models: A comparative study. *IEEE Transactions on Medical Imaging, 39*(4), 1355–1365.
21. Sharma, S., Shah, M., & Singh, A. (2020). Explainable AI in healthcare: Retinopathy detection using CNN. *Journal of Healthcare Informatics Research, 4*(2), 181–195.
22. Tang, L., Wang, Y., Li, X., & Zhang, W. (2020). Multi-modal data integration for enhanced diabetic retinopathy diagnosis. *Medical Image Analysis, 61*, 101657.

23. Kim, H., Lee, S., Choi, D., & Park, J. (2020). Assessment of deep learning models for the detection of diabetic retinopathy in a real-world clinical setting. *Ophthalmology, 127*(5), 715–724.
24. Gonzalez, R., Martinez, L., & Sanchez, J. (2021). Retinal image analysis for diabetes detection using machine learning: A review. *Computer Methods and Programs in Biomedicine, 199*, 104931.
25. Liang, J., Liu, Y., & Zhou, X. (2021). Applications of deep learning in diabetic retinopathy: A review. *Computers in Biology and Medicine, 132*, 104328.
26. Wang, S., Zhang, Z., & Li, J. (2020). Automated detection of diabetic retinopathy using deep learning algorithms. *Journal of Diabetes and Its Complications, 34*(4), 107–115.
27. Liu, C., Zhang, Y., & Huang, J. (2021). Deep learning-based diagnosis of diabetic retinopathy: A comprehensive review. *Ophthalmology Retina, 5*(5), 456–467.
28. Ahmad, R., Salahuddin, H., Rehman, A. U., Rehman, A., Shafiq, M. U., Tahir, M. A., & Afzal, M. S. (2024). Enhancing database security through AI-based intrusion detection system. *Journal of Computing & Biomedical Informatics, 7*(2).
29. Gupta, A., Sharma, P., & Singh, R. (2021). Comparative analysis of deep learning techniques for diabetic retinopathy detection. *Journal of Medical Imaging and Health Informatics, 11*(5), 1386–1394.
30. Khan, N. A., Iqbal, M., & Qureshi, S. (2022). Explainable AI techniques in medical imaging: Application in diabetic retinopathy detection. *IEEE Transactions on Neural Networks and Learning Systems, 33*(5), 1997–2008.
31. Hernandez, R., Torres, A., & Gomez, M. (2021). The role of artificial intelligence in retinal image analysis for diabetic retinopathy screening. *Survey of Ophthalmology, 66*(3), 523–535.
32. Wei, P., Wang, Q., & Zhao, S. (2022). Transformers in retinal image analysis for diabetic retinopathy detection. *IEEE Transactions on Medical Imaging, 41*(3), 881–891.
33. Huang, X., Yang, Z., & Liu, L. (2021). Real-world application of deep learning algorithms for diabetic retinopathy detection: Challenges and opportunities. *Lancet Digital Health, 3*(12), e755–e765.
34. Feng, Y., Zhang, X., & Wu, Y. (2021). Challenges in the application of deep learning for retinal image analysis: A review. *Computers in Biology and Medicine, 138*, 104916.
35. Zhu, T., Li, Y., & Chen, M. (2022). Deep learning models for retinal image analysis: A survey. *Neurocomputing, 463*, 163–173.
36. Rahman, M., Khan, S., & Iqbal, T. (2022). Federated learning for diabetic retinopathy detection. *Journal of Biomedical Informatics, 127*, 104012.
37. Chishti, M. F., Rao, M., Raffat, M. W., & Rafi, S. (2024). Estimating corporate risk and corporate value: An application of Altman's Z-Score on the KSE-30 index. *International Journal of Contemporary Issues in Social Sciences, 3*(2), 2833–2841.
38. Wei, P., Wang, Q., & Zhao, S. (2022). Transformers in retinal image analysis for diabetic retinopathy detection. *IEEE Transactions on Medical Imaging, 41*(3), 881–891.
39. Khan, N. A., Iqbal, M., & Qureshi, S. (2022). Explainable AI techniques in medical imaging: Application in diabetic retinopathy detection. *IEEE Transactions on Neural Networks and Learning Systems, 33*(5), 1997–2008.
40. Singh, A., & Gupta, R. (2022). Diabetic retinopathy detection in retinal images using novel deep learning techniques. *Biomedical Signal Processing and Control, 73*, 103515.
41. Nawaz, H., Sethi, M. S., Nazir, S. S., & Jamil, U. (2024). Enhancing national cybersecurity and operational efficiency through legacy IT modernization and cloud migration: A US perspective. *Journal of Computing & Biomedical Informatics, 7*(2).
42. Khalil, M. I. K., Umer, H., Ullah, K., Khamosh, S. U., Naqvi, S. F. M., Nawaz, A., & Ahmed, S. (2024). Revolutionizing schizophrenia diagnosis: A transfer learning approach to accurate classification. *Journal of Computing & Biomedical Informatics, 7*(2), 1–12.
43. Khan, R., Iltaf, N., Shafiq, M. U., & Rehman, F. U. (2023). Metadata based cross-domain recommender framework using neighborhood mapping. In *2023 International Conference on Sustainable Technology and Engineering (i-COSTE)* (pp. 1–8). IEEE.
44. Fatima, A., Shabbir, A., Janjua, J. I., Ramay, S. A., Bhatty, R. A., Irfan, M., & Abbas, T. (2024). Analyzing breast cancer detection using machine learning & deep learning techniques. *Journal of Computing & Biomedical Informatics, 7*(2).
45. Shafiq, M. U., & Butt, A. I. (2024). Segmentation of Brain MRI Using U-Net: Innovations in Medical Image Processing. *Journal of Computational Informatics & Business, 1*(1).