Enhancing Proliferative Diabetic Retinopathy Detection: Leveraging Customized CNN and Lightweight Machine Learning Models

MSc Research Project
MSc in Artificial Intelligence for Business

Amey Dhoot
Student ID: x23144424

School of Computing
National College of Ireland

Supervisor: Prof. Brian Byrne

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Amey Dhoot
Student ID:	x23144424
Programme:	MSc in Artificial Intelligence for Business
Year:	2023-24
Module:	MSc Research Project
Supervisor:	Prof. Brian Byrne
Submission Due Date:	12/08/2024
Project Title:	Enhancing Proliferative Diabetic Retinopathy Detection: Leveraging Customized CNN and Lightweight Machine Learning Models
Word Count:	5560
Page Count:	20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	16th September 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Enhancing Proliferative Diabetic Retinopathy Detection: Leveraging Customized CNN and Lightweight Machine Learning Models

Amey Dhoot
x23144424

Abstract

Diabetic patients are often under the danger of getting an incurable eye disease called as Diabetic Retinopathy or DR which may lead to harmful effects on the eyes such as blindness. Therefore, early detection of this disease is mandatory as it may lead to permanent vision loss, if it is not treated on time. But detecting its symptoms early is not an easy task, as it involves proper medical screening along with qualified individuals, as there are limited medical practitioners available, there is a necessity of automated solution for DR detection. To facilitate early diagnosis of this problem, lightweight machine learning algorithms can be used. In this research we will do a comparative study of algorithms like Logistic Regression, Random Forest and Convolutional Neural Network (CNN). Furthermore, to evaluate the test accuracy, metrics like accuracy, sensitivity, confusion matrix, specificity and F1 score will be measured. Apart from these metrics, we will calculate another metric which studies the amount of carbon emitted by the machine when these algorithms are implemented. This will help us to understand and analyse the algorithms which are computationally less demanding but at the same time gives higher accuracy. By doing so, we will be able to provide a sustainable solution. The output shows that Random Forest gives better test accuracy of 96% which is the highest from the remaining two algorithms. This work helps us to identify the different stages of DR thus providing a reliable and quick automated solution for screening tasks.

Keywords: Diabetic Retinopathy (DR), Logistic Regression, Random Forest, Convolutional Neural Networks (CNN), Carbon Emissions

1 Introduction

There has been a drastic increase in the number of diabetic patients in recent times. There are two types of diabetes – Type 1 and Type 2, it is evident that most of the people are impacted by Type 2 diabetes throughout the world Akil et al. (2022). This disease occurs when our body does not produce enough insulin resulting in high blood sugar levels. The impact of this disease can be seen on various body parts like eyes, kidneys, nerves and even heart. But the most unrecognized part is the eyes. On becoming more severe, it damages the blood vessels present in the eyes which connect the retina with optic nerves and causes diabetic retinopathy whose long-term impact is permanent visual impairment. Therefore, it is quite important to undergo eye checkups from time to time to know if there are any early signs of damage. Thus, the rise in diabetes Gautam et al. (2019) has emphasized the necessity of developing effective solutions.

Due to recent advancements in technology, it is possible to detect DR right in its initial stages. Traditional techniques like pupil dilation or optical coherence tomography are very effective but at the same time they consume a lot of time and therefore the patient's eye could be in danger. Moreover, there are high chances that human intervention in the screening examination process could be susceptible to errors, specifically during the starting stages of DR. The use of Artificial Intelligence and Machine Learning help us to address this problem by providing innovative solutions. These techniques aid us to identify the retinal fundus images with high accuracy and precision. The algorithms are trained thoroughly so that they can identify the smallest potential change in DR making it more efficient as compared to the traditional techniques. This results in saving diabetic patients from visual impairment.

This study focuses on developing state of the art solutions by implementing machine learning algorithms like logistic regression, random forest and convolutional neural networks.It is important to undergo thorough research and evaluation of these ML models. To measure the accuracy of the model metrics like confusion matrix, F1 score, sensitivity, etc can be used. Moreover, to provide an environmentally friendly solution we will be measuring the carbon emitted by each machine learning model while maintaining high accuracy. This will help us to mitigate the environmental impact by reducing the carbon footprint. Thus, by adopting ML models we will be able to detect DR early and manage it effectively while improving patients' life, by decreasing the risk of blindness. This report will include a deep analysis of related work, methodology and its implementation along with the evaluation of the ML models.

1.1 Research Questions and Objectives

1.1.1 Research Questions:

1. In what ways can we enhance the screening and early detection of diabetic retinopathy by using lightweight machine learning models?

2. In terms of Accuracy Rate, Confusion Matrix, F1 Score, Specificity and Sensitivity, how would the Convolutional Neural Networks (CNN), logistic regression and random forest perform?

3. How do different algorithms contribute to building Green AI solutions by calculating the carbon emissions associated with all algorithms?

4. How do lightweight machine learning models offer greater accuracy and efficiency when compared with the traditional approaches?

1.1.2 Objectives:

The main goal of conducting this research is to build and analyse portable machine learning models to detect DR in its early stages. Primarily, it focuses on the following:

1. Implementing Machine Learning Models: Design and develop lightweight machine learning models such as Random Forest and Logistic Regression, for classifying DR in different stages.

2. Evaluating Performance of ML Models: To know about the performance of these models, metrics like confusion matrix, F1 score, specificity and sensitivity will be accessed.

3. ML Models Vs Traditional Approaches: Comparing the lightweight machine models with the traditional approaches which are used for screening and detection of DR. Also, analysing the three algorithms among themselves is also critical to obtain the most effective ML model.

4. Creating Environment Friendly Solution: By calculating the carbon emitted by all these three algorithms, we will be able to provide a robust and sustainable solution.

5. Enriching Early Detection & Treatment of DR: By providing more efficient lightweight machine learning models, we can easily manage the early detection of DR, thus decreasing the risk of blindness.

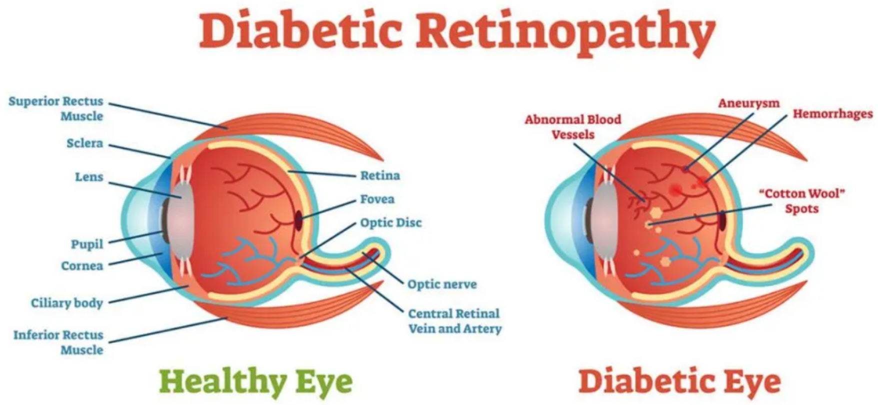

Figure 1: Diabetic Retinopathy asiaretina.com (2019)

2 Related Work

Various machine learning algorithms like Logistic Regression, Random Forest and Convolutional neural Network (CNN) are basically used for classification of images. Certain attributes are extracted from the images with the help of CNN, as shown in the study conducted for the detection of diabetic retinopathy. The layered architecture of CNN is used with other techniques like Adam optimizer and dropout layers to develop a robust algorithm Sugathan et al. (2023); Priya et al. (2023). For binary classification, Logistic Regression is used by combining techniques like sigmoid and loss functions for training purposes Azamen et al. (2023); Huang and Mao (2024). Lastly Random Forest Classifier is used to build a vast number of decision trees and by aggregating the decision made by individual trees, the model decides the output Nasir et al. (2022). All these algorithms help to increase the accuracy and predictive power to analyse the diabetic retinopathy problem.

This research illustrates the use of Convolutional Neural Networks to identify any early signs of DR detection based on the fundus images. It has been further classified into various categories from No DR to Severe DR (also included the stages in between the two). The research helps us to gain a deeper understanding of how vision impairment can be prevented for diabetes patients. This is a challenging task as it consists of complicated grading techniques along with subtle attributes. Here, usage of a highly computational graphics unit is made to train the ML model on the Aptos dataset and have achieved an accuracy of 80% based on the fundus. Also, to improve the accuracy further, data augmentation techniques were used to gain potentially improved ML models Sugathan et al. (2023).

Due to the recent advancements in the field of Artificial intelligence and Machine Learning, it has become evident that early detection of DR is possible using these technologies. In this research different medical departments are using ML models like Logistic Regression, Random Forest Algorithm, and Support Vector Machine to achieve high accuracy in the screening and detection process of DR. Among these models, the utmost accuracy was given by the logistic regression model achieving an accuracy of 83.80%. Moreover, along this research a web-designed application was suggested so that the medical experts can utilize these algorithms easily without the necessity of conducting any manual procedures Azamen et al. (2023).

It is proved that the manual process in the detection of DR is very time-consuming and subject to man-made errors. To address this problem, this research suggests creating a lightweight mobile application that includes features like capturing fundus images and automatically detecting the results based on it. Here, they have used the Random Forest Algorithm to do the multi-classification of DR in its different stages on the fundus images. Based on the proposed algorithm, the model achieved an accuracy of 76.27% along with an F1 score of 76.22, thus it outstands the previous models proposed on the same dataset Nasir et al. (2022).

In larger countries like India, where greenhouse gases are generated on a very large scale, it is evident that they have harmful effects on the environment leading to climatic changes, weather events and scarcity of natural resources. To address this problem, it is important to reduce the carbon emissions in the environment and India has managed to reduce its carbon emissions by 30-35% from 2005 to 2030. Highly Trained Machine Learning models utilize a lot of computational power and produce carbon continuously. As a result, it has become critical to take necessary actions to reduce the carbon footprint and keep the environment clean. Models like LSTM help us to calculate the CO2 emissions produced by the ML models Priya et al. (2023).

Using Artificial Intelligence, we can easily calculate the amount of carbon emitted by a particular algorithm. In this research, they have collected the relevant data in accordance with the carbon emissions, and by using the capability of machine learning they have developed optimized techniques that will aid us to reduce the carbon footprints. This analysis is done in real-time to maintain high operational efficiency and it adapts to the environmental change with ease. Thus, Artificial Intelligence not only helps us to identify the carbon emissions but also achieves highly accurate analysis based on it Huang and Mao (2024).

3 Methodology

3.1 Data Collection

The primary step is the collection of data. This research is based on the data collected from APTOS 2019 public dataset for Diabetic Retinopathy (Blindness) detection which is present on Kaggle website. In this dataset, retinal images of high quality are present, and they are further divided into 5 different parts based on the different stages of DR such as No DR, Mild, Moderate, Severe and Extremely Severe DR. It also comprises two different sets of images mainly used for training and testing purposes. The data present in the training set is labelled and its class is present in a separate CSV file whereas the testing data is unlabelled which will be used for evaluation purposes. In total, there are 3662 high resolution images classified into 5 different classes Mohanty et al. (2023).

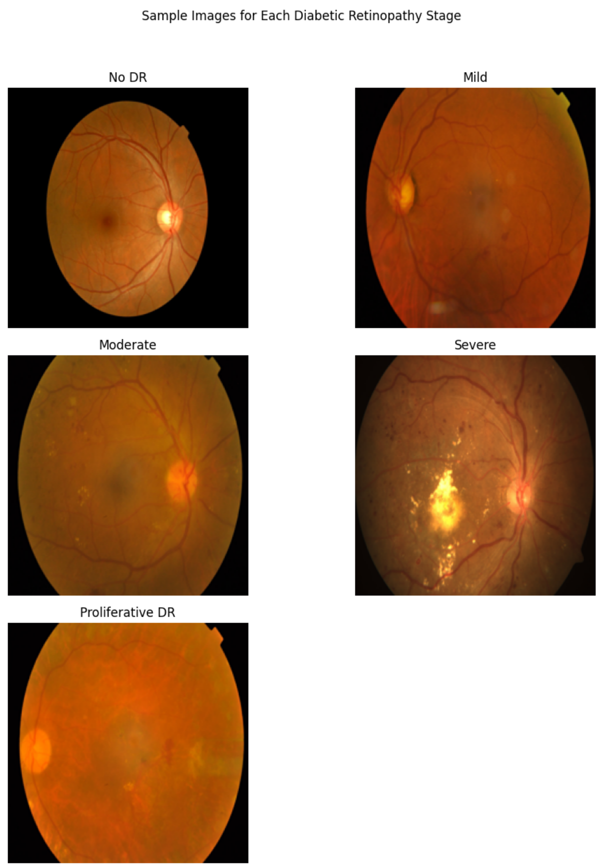

Figure 2: Sample Images for Diabetic Retinopathy Classes

This dataset is collected from various sources like clinics, hospitals etc which ensures that there is variability in the data under study. While collecting the data, vital patient information was removed to maintain data integrity. It was directly downloaded from Kaggle by maintaining the desired protocols and securely stored in the system. This dataset will help us to develop and implement machine learning models on these fundus images to gain accurate detection of diabetic retinopathy. Following diagram shows the distribution of the different classes in the dataset-

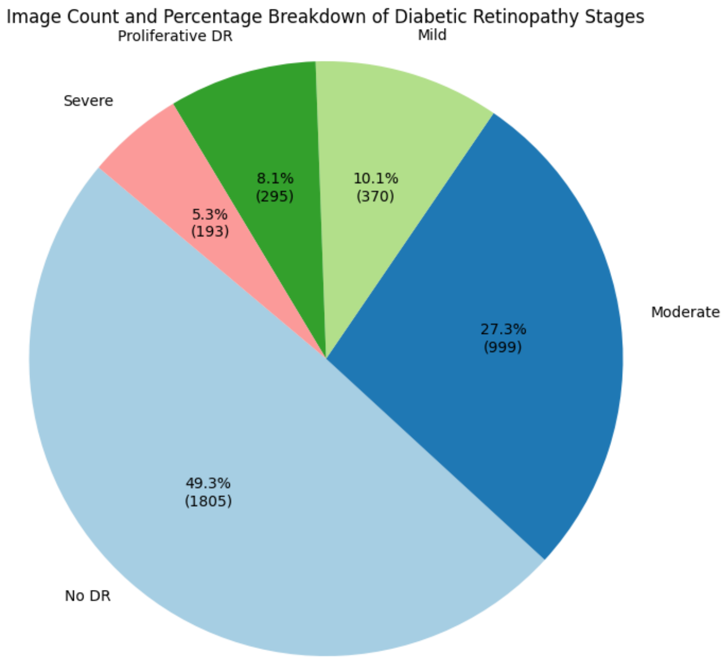

Figure 3: Image Distribution for Diabetic Retinopathy Dataset

3.2 Data Preprocessing/Transformation

There are various critical steps that need to be taken while preprocessing the data and making it ready for training and evaluation Maneerat, N., Thongpasri, T., Narkthewan, A. and Kimpan, C. (2020). For the Convolutional Neural Networks, images are read with the help of OpenCV python library and then they are resized to 256x256 pixels to maintain uniform data throughout the dataset. Further, every image is transformed into an array which ranges between 0 and 1. Based on the different stages of diabetic retinopathy, labels are assigned, and images are then amended to the respective lists and further they are converted to NumPy arrays. To preserve the class distribution, the dataset is divided into training and testing sets in the ratio 9:1. As this is a categorical binary classification, one-hot encoding is performed on the labels to maintain uniformity. Finally, data augmentation techniques like rotating the image, magnifying it, shearing or horizontal flipping is done using the Image data generator to make the model more robust to face the real-world images.

For effectiveness in the model training process, it is vital to follow certain important steps. The images need to be standardized to a specific size of 512x512 pixels, here we will generate two types of images, one would be coloured and the other will be grayscale Çinarer and Kiliç (2022). This is done to enhance the minute details present in the real-world data. To improve the contrast in the images, gaussian blurring is done and applied to all the images. To ensure that the area of study is focused on the retinal area, unwanted dark borders are removed. For the binary classification, the data consisting of images of different stages of diabetic retinopathy and the healthy images are labelled again. They are further resized to 64x64 pixels and then again converted to grayscale to maintain consistency. It is important to flatten these images into a 1-D array for model feeding. This data transformation will enable us to process the images uniformly and thus gaining better accuracy.

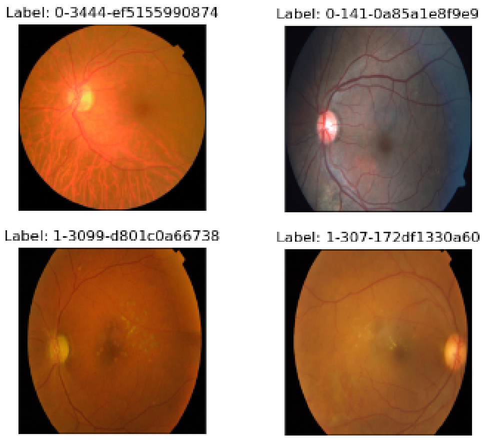

Figure 4: Before Image Preprocessing

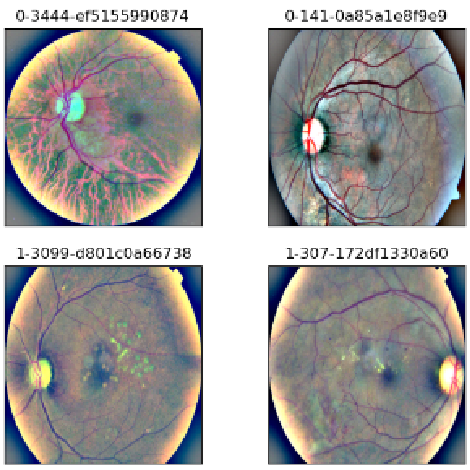

Figure 5: After Image Preprocessing

3.3 Data Modelling

In this research we will be implementing the ML models like Logistic Regression and Random Forest Classifier from scratch using the sklearn python library. This library will help us to initiate the weights and any bias in the data, perform backward and forward propagation. Parameter updates through gradient descent and ultimately make predictions based on the proposed model. To evaluate the model, we will need to calculate the train and test accuracy. Also, further we will show a comparison of these two algorithms showcasing which one is more efficient and can accurately predict the classes.

The deep learning algorithm that we will be using is the Convolutional Neural Networks which is generally used for image classification purposes. In tis study we will take 4 convolutional layers, wherein each layer is followed by the max-pooling strategy and dropout layers to avoid the risk of overfitting the dataset. Along with it, this model will consist of 2 dense layers and an output layer at the end which will include the sigmoid activation function for binary classification. On top of this, it also includes an Adam optimizer and handles the entropy loss. In the data modelling stage, data augmentation techniques are utilized with the adoption of early stopping techniques to avoid the overfitting problem.

3.4 Evaluating ML Models

To evaluate the accuracy and efficiency of the ML models, several steps are to be followed. Based on the predictions generated by these ML models, certain metrics such as confusion matrix, F1 score, precision and recall were calculated for each classification Moccia and Mattos (2018). The confusion matrix helps us to predict the models on the unlabelled data and thus define its accuracy. To represent the model's ability to differentiate the classes, ROC and AUC curves were plotted. Furthermore, a brief classification report was generated which gives a clear picture in the identification of positive and negative images. These metrics helped in determining the efficiency of these machine learning models accurately.

3.5 Analysing Carbon Footprint

When the model is evaluated using various metrics, I calculated the energy consumed by each machine learning model in terms of RAM, CPUs and GPUs. This analysis will help me to analyse to what extent a particular algorithm has an impact on the environment. Thereafter, each model's CO2 emissions were calculated to understand how sustainable the developed model is, in delivering accurate results along with being sustainable.

4 Design Specification

The below approach highlights the flow for the detection of diabetic retinopathy. Basically, it makes a comparison between the lightweight machine learning models like logistic regression or random forest to the deep learning model of convolutional neural networks. When implementing the lightweight models, data preprocessing steps like converting images to grayscale, resizing the images, flattening them and after the image is ready, applying models to evaluate them. Parallelly, in the implementation of convolutional neural networks different preprocessing is carried out such as resizing image pixels, normalizing and undergoing data encoding along with data augmentation and splitting the data for applying the CNN model. The aim of both models is to detect the diabetic retinopathy condition in patients accurately and with greater effectiveness.

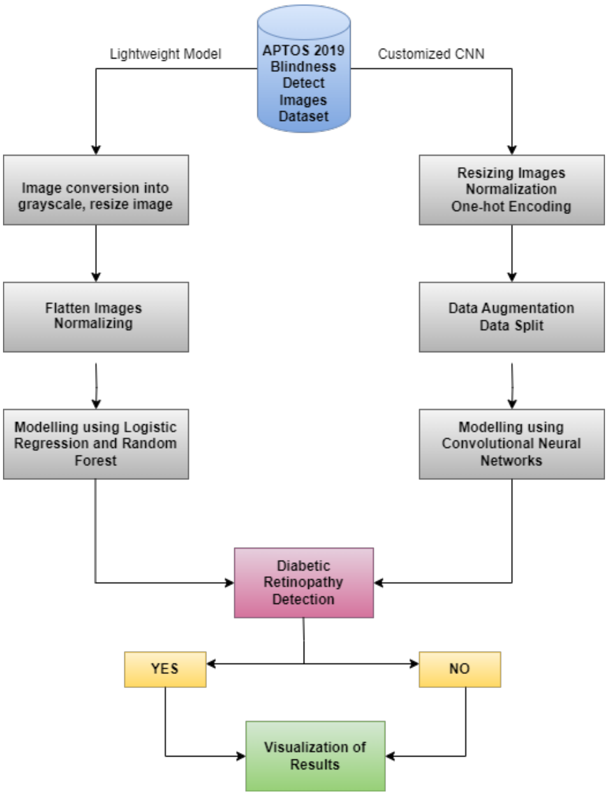

Figure 6: Design Flowchart for Diabetic Retinopathy

5 Implementation

In this research, a highly equipped Intel Core I7 processor is used with a RAM of 16GB. It also has a graphics card of 6GB Nvidia GeForce RTX 3060 and a storage of 512 GB. These configurations are well suited for performing highly computational tasks of processing the images using machine learning models like logistic regression, random forest and convolutional neural networks. The system has enough power to execute these complex algorithms. To implement the algorithms, python language is used in the Jupyter Notebook environment. The environment has all the necessary libraries installed for smoother execution.

5.1 Random Forest Algorithm

Random Forest Algorithm uses the ensemble learning approach for binary classification. In this approach, it creates as many decision trees as possible when undergoing the training stage and then predicts the output class based on it. The basic goal of this algorithm is to create a robust and generalized algorithm using multiple base estimators to improve the accuracy of the prediction geeksforgeeks (2024c).

- **Decision Tree Creation:** A decision tree is used to decide based on certain attributes where each attribute is represented by the node and the test outcome is stated with the help of branches. The leaf of the tree states the labels of a particular class.

- **Ensemble Technique:** The ensemble method is used in the building of random forest algorithms. It states that instead of relying on just a single tree to make decisions, it uses multiple trees and sums them up to make a particular decision.

- **Bootstrapping or Bagging for Aggregation:** In this algorithm, certain subsets of the dataset are created for the training purpose. Furthermore, to make the model more efficient and robust, each tree is trained based on the randomly created subsets of the data using the bootstrap sampling technique.

- **Selection Random Attributes:** When the decision tree is divided randomly, the random forest algorithm selects only the essential attributes instead of choosing all of them. This approach helps to reduce the correlation between the trees and thus obtain a more accurate solution.

The Random Forest Model is initialized with the help of 'RandomForestClassifier' and 100 trees are allocated to it with a fixed amount of randomness to ensure reproduction of the model. In the preprocessing stage, we have flattened the training data which is used in the training of the algorithm using the 'fit' function. In the training phase, 100 new trees are constructed, where each tree uses the bootstrapped data. At every node of this tree, where it is divided further, random features are selected to decrease the dependency and make the model robust.

After training the trees, flattened test data is used for prediction. Now, every tree votes for a specific class of the data and then the result is aggregated to make the ultimate prediction. Moreover, the model also finds the probability of predicting the positive class which can be used for analysis. This approach of adding multiple trees together to make an accurate decision, solves the problem of overfitting and makes the model more reliable.

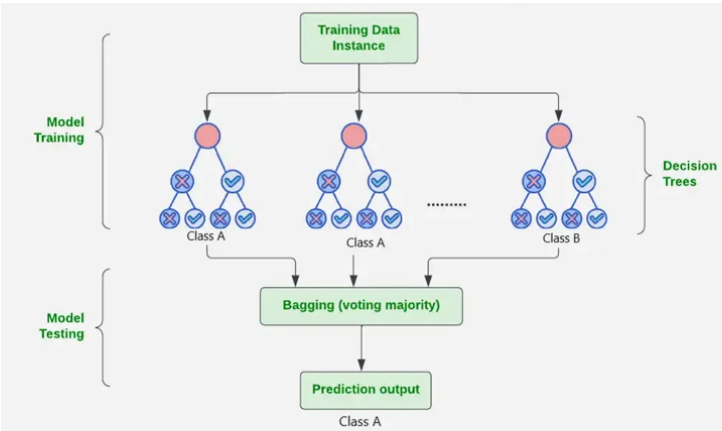

Figure 7: Random Forest geeksforgeeks (2024a)

5.2 Logistic Regression

The Logistic Regression model is mainly used for binary classification where the input data is divided into one or more classes to give a specific output. Here it is used to classify whether the patient is suffering from diabetic retinopathy or not. This algorithm defines the relationship between the dependent and one or more independent attributes.

- **Sigmoid Function:** The probability is mapped with the predicted class using the sigmoid function. It is interpreted as when the output is nearer to 1, the model has high probability of predicting the positive class whereas when it is near 0, it predicts the probability for the negative class.

- **Decision Boundary:** A specific threshold is applied at the output generated from the sigmoid function to define the input into multiple classes. It uses a threshold of 0.5, that means when it is greater than this value it belongs to the positive class or else it belongs to the negative class. In this case, we can lower the threshold value so that we don't miss any of the positive cases especially when there is a high cost of false negatives.

- **Cost Function:** The cost function is a check that compares the predicted outcomes to the actual labels. To improve the accuracy of the mode. The cost function should be decreased during the training phase. It uses the log loss technique where it gives more penalty on the incorrect output than the correct output. This is done using the gradient descent algorithm so that it reduces the weight constantly while reducing the cost.

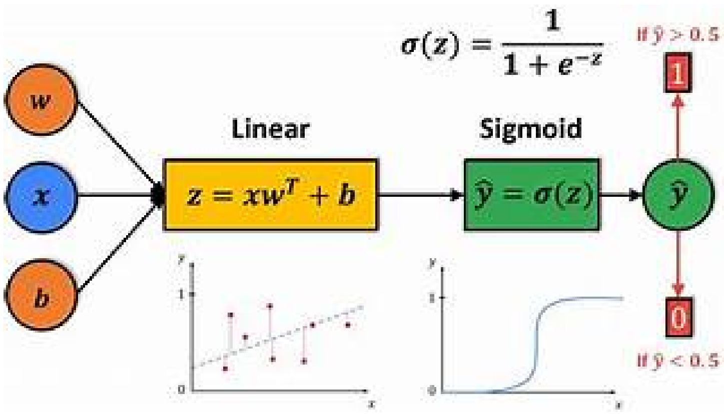

Figure 8: Logistic Regression yahoo.com (n.d.)

The Logistic Regression model is specifically implemented for classification purposes geeksforgeeks (2024b). While training the model, it tries to make an estimation of the optimum number of coefficients (150 in this case) for each input attribute to make a linear decision boundary so that the probability of predicting the output variable is more.

This decision boundary helps us to separate the data into two classes more effectively depending on the various attributes and the output variable. After training the model, the predict method is used to test the data and then store it. This algorithm offers a linear approach, hence there are very less chances of overfitting and it also has the capability to understand complex patterns in the dataset making it more flexible and reliable.

5.3 Convolutional Neural Networks

The Convolutional Neural Networks (CNN) Algorithm is often used to classify the images or doing binary classification like detecting diabetic retinopathy or not in this case. The architecture is designed in such a way that it identifies certain patterns in the images and based on that it distinguishes it in different classes Ouyang and Wu (2019). The primary step in the execution of this algorithm is to convert the images to a specific size for instance 256 X 256 pixels and then these are again normalized such that its values lie between 0 and 1. This preprocessing helps to fasten the training procedure for all the image data Touati et al. (2023a).

- **Convolutional Layers:** In CNN, the first layer is the convolutional layer where the feature extraction takes place. There are several filters in this layer which extracts the necessary information like textures and edges of the images. Every element follows an element-based multiplication which is then added together to generate a feature map. This is followed by the Rectified Linear Unit (RELU) activation function which converts negative pixels to zero to maintain linearity in the process. ReLU function helps to capture minute details about the images and its features.

- **Pooling Layers:** Pooling layers are often used to reduce the dimensionality of the feature maps. There is an approach called MaxPooling which signifies only the necessary information to consider while reducing the model complexity. This helps the model to remain robust for every small change in the images and enables it to recognize the key information even if it is present at different locations. This layer might decrease the size of the feature map but preserves the vital information at the same time.

- **Dropout Layers:** To solve the problem of overfitting, the model executes the dropout layer mechanism. It chooses certain highly dependent features from the input and maps that to 0, so that model doesn't rely on a single feature to train it. For instance, a dropout of 50% can be applied on the dataset in which only half of the data is considered for training while the rest is ignored.

- **Fully Connected Layers (Flattening):** The Images are flattened into a One-dimensional vector after the execution of pooling and dropout layers. This vector is given to the fully connected dense layers for performing feedforward neural networks. These layers help in identifying the classes based on the extracted features from the convolutional layers. To maintain non-linearity, the first dense layers make use of the ReLU activation function. While the final dense layers make use of softmax function to give a probability between the two classes making the output equal to 1.

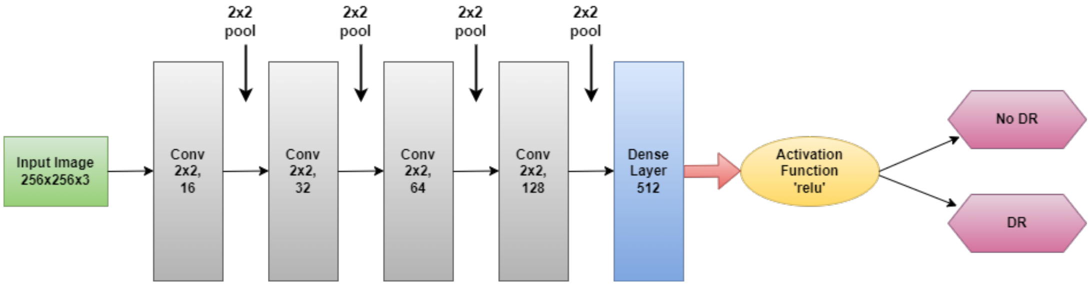

Figure 9: CNN Architecture

The model is then executed using the cross-entropy loss function which enables us to perform binary classification more efficiently. To handle the gradients in the images and adjust the learning rate, Adam optimizer is used. The training process is applied on the pre-processed images over many epochs by adopting certain techniques such as early stopping. To improve generalized input and prevent the model being overfitted, various data augmentation techniques are used. Lastly, after the training is completed, it is tested on the unknown data and various metrics for accuracy, f1 score, confusion metric is calculated to measure model efficiency.

6 Evaluation

Different metrics are used to check the effectiveness of the machine learning models. Accuracy is one of the measures that defines the correctness of predicting data based on its learning. Sensitivity gives the ratio of correctly predicted class to the total while specificity states the ratio in which incorrectly identified class is plotted against the total data. F-1 score is used to analyse the performance of each machine learning model. Following are the equations that are used to calculate these metrics Touati et al. (2023b):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$specificity = \frac{TN}{TN + FP} \tag{3}$$

6.1 Evaluation for Logistic Regression:

The confusion matrix is used to identify the model performance correctly. For Logistic Regression, it identifies 528 No DR images and 71 DR images correctly. Although it fails to identify 17 No DR images and 14 DR images correctly. The states where the algorithm misclassified the data is called False Positive and False Negative. These metrics aid us to calculate other metrics like precision, recall and f1 score which play a vital role in deciding the effectiveness of the algorithm Emon and Ohidujjaman (n.d.).

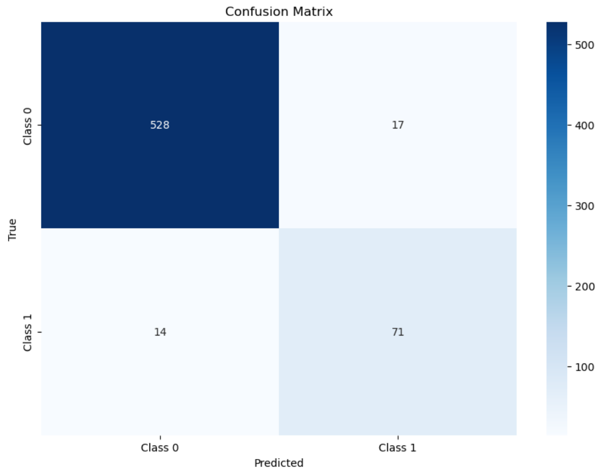

Figure 10: Confusion Matrix for Logistic Regression

The ROC and AUC curve gives a high model performance with an AUC value of 0.97. This states the model could identify the discrimination and differencing between the positive and negative images correctly. If we observe the ROC curve at the top-left corner, it defines that the model has a high true positive rate and low false positive rate.

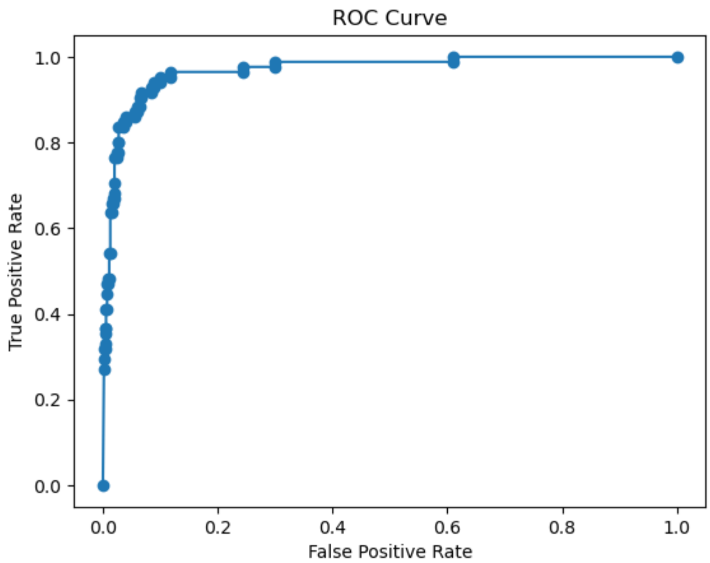

Figure 11: ROC Curve for Logistic Regression

The amount of carbon emitted by running the logistic regression model is quite low as it's a lightweight model. It emits around 1.0734 Kg of carbon. The following graph indicated the stability in the carbon emissions throughout its 150 iterations. Therefore, this can be treated as an environment friendly solution as it emits low carbon into the atmosphere making it more sustainable. Using a Logistic Regression model will help us to move towards Green AI.

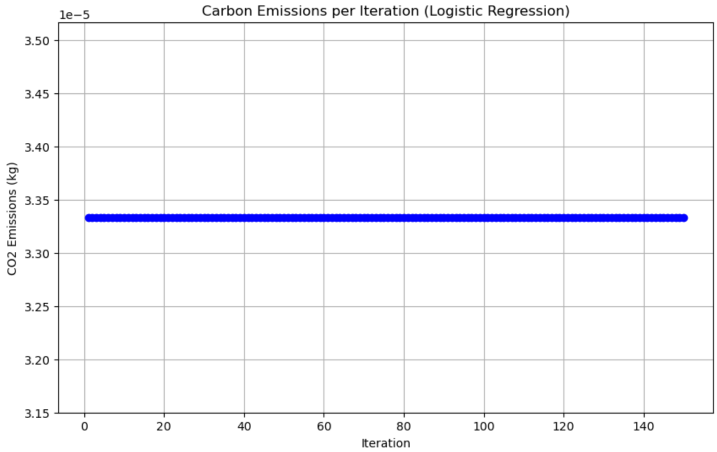

Figure 12: Carbon Emission for Logistic Regression

6.2 Evaluation for Random Forest:

Receiver Operating Characteristic or ROC curve is used to illustrate the performance of the random forest algorithm. It shows an AUC score of 0.96 which states the good accuracy of the model. Its proximity is quite near to the left-topmost corner which defines its strong behaviour to identify the two classes. Thus, it acts as a reliable solution to identify diabetic retinopathy due to its strong predictive power.

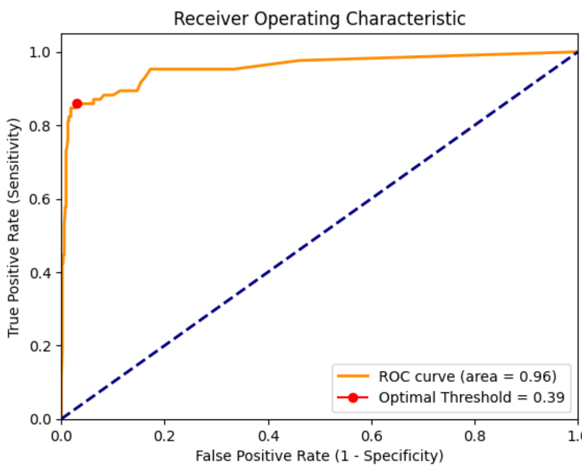

Figure 13: ROC Curve for Random Forest

Confusion matrix defines the performance of the random forest classifier in terms of true positive, true negative, false positive and false negative. This model shows that 535 No DR images and 70 DR images are correctly identified whereas the misclassification id for 10 No DR images and 15 DR images. This shows the model's ability to classify the images into correct classes efficiently.

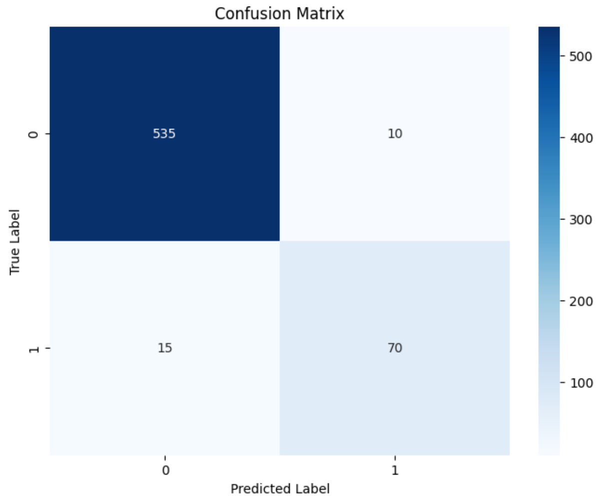

Figure 14: Confusion Matrix for Random Forest

6.3 Evaluation for Convolutional Neural Networks(CNN):

For the convolutional neural network, the area under the ROC curve is quite impressive. Its AUC value is 0.96 which demonstrates the ability to classify the dataset accurately. The curve at top-most left corner indicates a high true positive rate and low false positive rate. This algorithm is reliable and can be used to classify DR effectively.

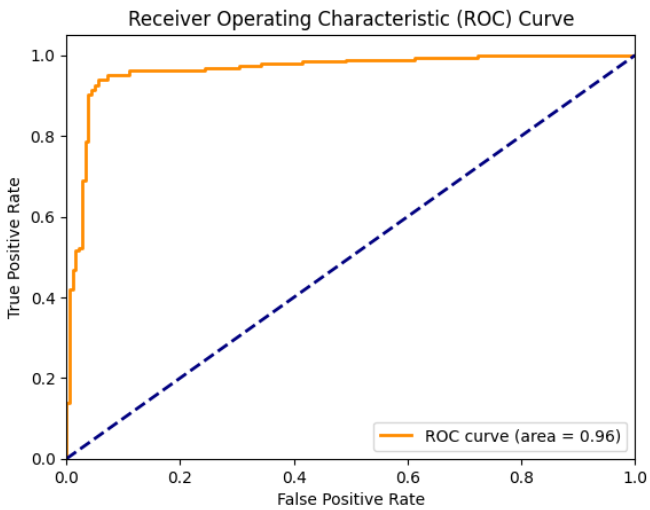

Figure 15: ROC Curve for CNN

The confusion matrix gives a model comparison in identifying the model accuracy. This algorithm predicts 173 DR and 173 No DR cases correctly and has only 8 false positive predictions. Here out of 173 No DR images, there are no false negatives, which indicates a strong model performance in identifying the No DR class correctly. This shows that CNN has a strong precision and recall rate in identifying DR as compared to other two classes based on the confusion matrix.

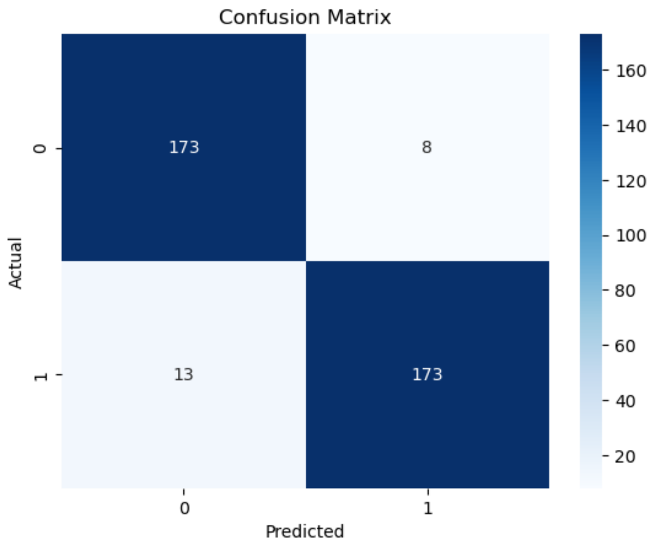

Figure 16: Confusion Matrix for CNN

The percentage of carbon generated by running the convolutional neural networks model is quite low despite of its complex architecture. It generates around 0.0004 Kg of carbon. Following graph shows the stability in the carbon emissions throughout its epochs. As a result, it is treated as an environmentally sustainable solution as it releases low carbon into the atmosphere. By using a CNN model we can march towards green AI.

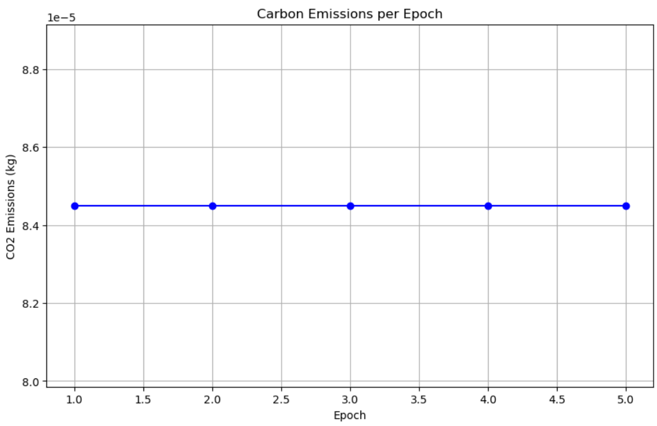

Figure 17: Carbon Emission for CNN

6.4 Discussion

According to the different metrics used for evaluating the algorithms like Logistic Regression, Random Forest and Convolutional Neural Networks, strengths of each algorithm are defined. For the binary image classification, the utmost accuracy is given by the Random Forest classifier with an accuracy score of 96%. Along with it, Logistic Regression also possessed a solid accuracy of 95% with precision and recall being highly balanced. In this way, these lightweight machine learning algorithms handle the large datasets and its complicated features in an effective way. Lastly the convolutional neural networks also possess a solid performance for the image classification of complex patterns and features with a bit lower accuracy of 94% though it gave higher costs for computation.

	Accuracy	CO2 Emission
Logistic Regression	95%	0.0734 Kg
Random Forest	96%	1.3724 Kg
Convolutional Neural Network	94%	0.0004 Kg

When we consider the carbon emitted by all these algorithms, Convolutional Neural Networks deep learning model possess patterns and complex architecture which has resulted in 0.0004 kg CO2, and for random forest it is 1.3724 kg CO2 emissions. Also, the carbon emitted by logistic regression is 0.0734 Kg. Among the three, Convolutional Neural Network has very low carbon emission of about 0.0004 kg. Moreover, if we run all these models in a single code script, the previously running background process can have an impact on the overall carbon emissions. This shows the comparison between the different models in terms of accuracy and carbon emissions focusing on the need of developing environmentally sustainable solutions.

7 Conclusion and Future Work

This research was particularly focused on the detection of diabetic retinopathy accurately by using machine learning models. For classification of the fundus images, models like Logistic Regression, Random Forest and Convolutional Neural Networks were used which yielded an accuracy score of 95%, 96% and 94% respectively. Out of these, Random Forest algorithm gave the most accuracy due to its robustness in identifying various features while handling large datasets. Logistic Regression also performed well with an accuracy score of 95%. It also showed its ability to classify the classes efficiently. Lastly, Convolutional Neural Network, with a bit lower accuracy than others, performed extremely well considering its complex architecture along with its difficult patterns and gave accuracy of 94%. These results show the comparison between all these three algorithms and help us choose a more accurate and reliable algorithm for analysis among them. In terms of accuracy score, Random Forest is the winner, but at the same time Logistic regression gives a solid baseline and CNN also offers critical insights despite its lower accuracy.

As part of the future work to improve the performance of these algorithms further, we can use ensemble learning approaches where the combination of two or more algorithms can be done considering their strengths to check if they perform well together and yield more accuracy in efficiently classifying DR. Moreover, the impact of tuning hyperparameters can be identified and more complex architectures like transfer learning can be implemented with the help of pre-trained models to improve the robustness of the models. Along with these improvements, we need to expand our dataset to include different and diverse images that will aid to generalize the performance of these algorithms.

References

Akil, H., Burgess, J., Nevitt, S., Harding, S. P., Alam, U. and Burgess, P. (2022). Early worsening of retinopathy in type 1 and type 2 diabetes after rapid improvement in glycaemic control: a systematic review, *Diabetes Therapy* **13**(1): 1–23.

asiaretina.com (2019). asiaretina.com, `https://asiaretina.com/wp-content/uploads/2019/12/diabetic-retinopathy-illustration-1080x522.jpg`. [Accessed 11-08-2024].

Azamen, N. T. N. N., Ali, A. M. and Aziz, N. A. A. (2023). Prediction of diabetic retinopathy based on risk factors using machine learning algorithms, *2023 4th International Conference on Artificial Intelligence and Data Sciences (AiDAS)*, pp. 308–312.

Çinarer, G. and Kiliç, K. (2022). Diabetic retinopathy detection with deep transfer learning methods, *Intelligent and Fuzzy Techniques for Emerging Conditions and Digital Transformation: Proceedings of the INFUS 2021 Conference, held August 24-26, 2021. Volume 2*, Springer, pp. 147–154.

Emon, M. U., Z. R. K. T. R. M. K. M. S. and Ohidujjaman (n.d.). Performance analysis of diabetic retinopathy prediction using machine learning models, 2021 6th International Conference on Inventive Computation Technologies (ICICT), pp. 1048–1052.

Gautam, A. S., Jana, S. K. and Dutta, M. P. (2019). Automated diagnosis of diabetic retinopathy using image processing for non-invasive biomedical application, *2019 international conference on intelligent computing and control systems (ICCS)*, IEEE, pp. 809–812.

geeksforgeeks (2024a). `https://media.geeksforgeeks.org/wp-content/uploads/20240701170624/Random-Forest-Algorithm.webp`.

geeksforgeeks (2024b). Logistic Regression in Machine Learning - GeeksforGeeks — geeksforgeeks.org, `https://www.geeksforgeeks.org/understanding-logistic-regression/`.

geeksforgeeks (2024c). Random Forest Algorithm in Machine Learning - GeeksforGeeks — geeksforgeeks.org, `https://www.geeksforgeeks.org/random-forest-algorithm-in-machine-learning/`.

Huang, R. and Mao, S. (2024). Carbon footprint management in global supply chains: A data-driven approach utilizing artificial intelligence algorithms, *IEEE Access* **12**: 89957–89967.

Maneerat, N., Thongpasri, T., Narkthewan, A. and Kimpan, C., title = Detection of hard exudate for diabetic retinopathy using unsupervised classification method, h. . t. (2020).

Moccia, S., D. M. E. E. H. S. and Mattos, L. S. . (2018). Blood vessel segmentation algorithms — review of methods,, datasets and evaluation metrics, Computer Methods and Programs in Biomedicine 158: 71–91.

Mohanty, C., Mahapatra, S., Acharya, B., Kokkoras, F., Gerogiannis, V. C., Karamitsos, I. and Kanavos, A. (2023). Using deep learning architectures for detection and classification of diabetic retinopathy, *Sensors* **23**(12): 5726.

Nasir, A. M., Abir, M. A.-I., Sadaf, Z. T., Shrestha, S. C., Shahnaz, C. and Subhana, A. (2022). Automatic diabetic retinopathy detection using random forest classifier, *2022 IEEE Region 10 Symposium (TENSYMP)*, pp. 1–6.

Ouyang, Y., W.-K. and Wu, S. (2019). Sar image ground object recognition detection method based on optimized and improved cnn, 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Vol. 1, pp. 1727–1731.

Priya, N., Srinidhi, K. and Kousalya, T. (2023). Carbon footprint monitoring system using machine learning and deep learning techniques, *2023 12th International Conference on Advanced Computing (ICoAC)*, pp. 1–8.

Sugathan, D., Sugumar, S. J., SamarasimhaReddy, B., Darshan, S., Naveen, R. and Sachin, K. (2023). Early detection of diabetic rethinopathy using machine learning, *2023 7th International Conference on Design Innovation for 3 Cs Compute Communicate Control (ICDI3C)*, pp. 243–248.

Touati, M., Nana, L. and Benzarti, F. (2023a). A deep learning model for diabetic retinopathy classification, *International Conference on Digital Technologies and Applications*, Springer, pp. 159–168.

Touati, M., Nana, L. and Benzarti, F. (2023b). A deep learning model for diabetic retinopathy classification, *in* S. Motahhir and B. Bossoufi (eds), *Digital Technologies and Applications*, Springer Nature Switzerland, Cham, pp. 159–168.

yahoo.com (n.d.). 004 Machine Learning - Logistic Regression Models - Master Data ... — in.images.search.yahoo.com, https://in.images.search.yahoo.com/images/view;_ylt=AwrtgxWanLhm4wc58Oe9HAx.;_ylu=c2VjA3NyBHNsawNpbWcEb2lkA2ViZDQ1Yzg4YTRkMDE5Y2FkNGQyOTQzMzc5MDdlNDc0BGdwb3MDMTkEaXQ
back=https%3A%2F%2Fin.images.search.yahoo.com%2Fsearch%2Fimages%3Fp%3Dlogistic%2Bregression%2Binfo%26type%3DE210IN714G0%26fr%3Dmcafee%26fr2%3Dpiv-web%26tab%3Dorganic%26ri%3D19&w=1555&h=879&imgurl=media5.datahacker.rs%2F2021%2F01%2F83.jpg&rurl=https%3A%2F%2Fdatahacker.rs%2F004-machine-learning-logistic-regression-model%2F&size=119.1KB&p=logistic+regression+info&oid=ebd45c88a4d019cad4d294337907e474&fr2=piv-web&fr=mcafee&tt=%23004+Machine+Learning+-+Logistic+Regression+Models+-+Master+Data+...&b=0&ni=21&no=19&ts=&tab=organic&sigr=RyRn5jQbZqG3&sigb=jUBUPcd2rZbW&sigi=rCz4_hzqaeBd&sigt=Q2LanGIfW2WT&.crumb=3STUpDwfDYy&fr=mcafee&fr2=piv-web&type=E210IN714G0.