IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

DIABETIC RETINOPATHY DETECTION USING MULTI-TASK LEARNING

[1] Mr. S. Satishkumar, [2]Niranjan V, [3]Ashwak A, [4]Sunil A, [5]Rogan Aravinth V

[1]Assistant Professor, [2,3,4,5] UG Scholar

Department of Information Technology,

Adhiyamaan College of Engineering, Hosur, India

Abstract: Diabetic retinopathy (DR) is a serious eye disease that affects individuals with diabetes and can lead to vision loss and blindness. Preventing vision loss requires timely recognition and therapy of DR. Currently, DR detection is performed by trained ophthalmologists using fundus images of the eye. However, due to the increasing number of individuals with diabetes, the demand for DR screening exceeds the availability of trained professionals, making it necessary to develop automated DR detection systems. We provide a unique technique to DR detection based on multitask learning in this study. The proposed model is trained on a large dataset of fundus images to perform multiple tasks simultaneously, including DR classification, gaussian blur, and grey scaling. By jointly optimizing all tasks, the model is able to make more informed predictions and improve its performance compared to traditional single-task models. Experimental results on a publicly available DR dataset show that our multitask learning approach outperforms state-of-the-art DR detection models in terms of accuracy and F1-score. Additionally, the gaussian blur capabilities of the model enable it to provide visual explanations for its predictions, which can be valuable for clinicians to understand the basis for DR diagnosis. In conclusion, our proposed multi-task learning approach for DR detection demonstrates promising results and has the potential to be used as an efficient and effective screening tool for DR in a clinical setting.

Index Terms **- Multi-task learning, Diabetic Retinopathy, Convolutional Neural Networks, Annotated fundus images.**

I. INTRODUCTION

Diabetic Retinopathy (DR) is a common microvascular complication of diabetes that can lead to vision loss and blindness. It affects millions of people worldwide and is one of the leading causes of blindness in adults. Early detection and prompt treatment of DR are essential to prevent vision loss. However, DR is often asymptomatic in its early stages and can go undetected until the disease has progressed to an advanced stage. The visual changes in DR eye is shown in figure 1.

Traditionally, DR has been detected by ophthalmologists using various diagnostic tools such as fundus photographs, fluorescein angiography, and optical coherence tomography. These diagnostic tools have limitations, including the need for specialized equipment, trained personnel, and a clinical setting. This makes DR detection challenging, particularly in resource-limited settings where access to healthcare is limited. Recently, there has been a growing interest in developing computer-aided diagnostic (CAD) systems for DR detection.

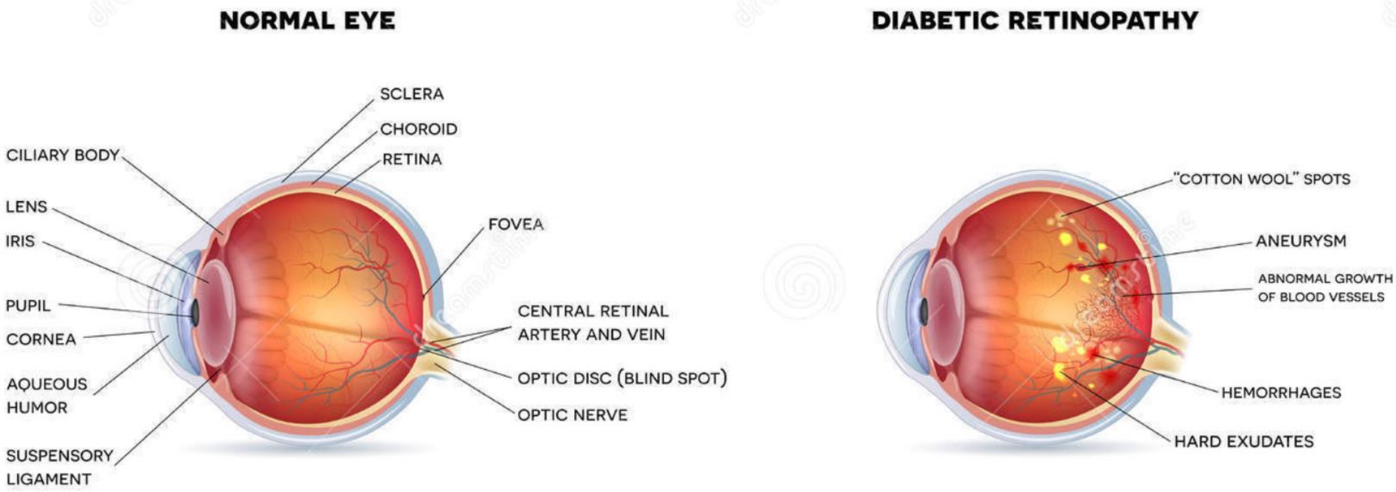

Figure 1: Visual representation of DR eye

A single model is trained to carry out several related tasks at once using the multi-task learning (MTL) machine learning technique. MTL has been shown to be effective in a variety of medical imaging applications, including DR detection. In DR detection, MTL can be used to perform multiple related tasks, such as DR classification, severity grading, and lesion localization. By performing these tasks simultaneously, MTL can improve the accuracy of DR detection and provide additional information that can be used to guide treatment decisions.

In this paper, we discuss the application of MTL in DR detection, including the methods and algorithms used, the datasets used for training and evaluation, and the results achieved. We also highlight the challenges and limitations of MTL in DR detection and discuss potential future directions for research. The Diabetic retinopathy is classified into five stages:

- Level 0 – *No Diabetic Retinopathy*
- Level 1 – *Mild Diabetic Retinopathy*
- Level 2 – *Moderate Diabetic Retinopathy*
- Level 3 – *Proliferative Diabetic Retinopathy*
- Level 4 – *Severe Diabetic Retinopathy*

Level 0 is a stage that denotes that the person is not affected by DR. Level 1 diabetic retinopathy appears as minute regions of swelling (microaneurysms) in the blood vessels of the retina. In level 2, some blood vessels that nourish the retina become blocked, leading to decreased blood supply to parts of the retina. This can result in the growth of abnormal blood vessels in the retina. In level 3, many blood vessels become blocked, leading to even less blood supply to the retina. This can cause the growth of new blood vessels, which are fragile and can bleed, leading to vision loss. In level 4, new blood vessels continue to grow, but they are also accompanied by scar tissue. The scar tissue can contract and pull on the retina, leading to retinal detachment and severe vision loss.

II. LITERATURE SURVEY

Fiaz Gul Khan[5] et al. (2019), describes "A deep learning ensemble approach for diabetic retinopathy detection,''. This method utilizes a combination of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to classify retinal images into different stages of DR. The proposed ensemble approach consists of two CNN models, two RNN models, and a decision-making module that combines the outputs of these models to make the final classification. The CNN models are designed to extract visual features from the retinal images, while the RNN models are used to capture temporal dependencies between different image patches. The author evaluated his approach on a publicly available dataset and achieved promising results, demonstrating the effectiveness of their ensemble approach for DR detection. The proposed method has the potential to aid in the early detection and diagnosis of DR, which is a leading cause of blindness in diabetic patients.

K. Wang[4] et al. (2020), proposed "a deep learning model called CABNet (Category Attention Block Network) for imbalanced diabetic retinopathy grading." The author addressed the issue of imbalanced data in the dataset, where some classes have very few samples compared to others, by introducing a category attention block that allows the model to focus on important regions of the image for each category. The CABNet architecture consists of a backbone network, which extracts features from the input image, and a category attention block, which selectively attends to the features relevant to each class. The category attention block consists of a convolutional layer that generates a category-specific attention map, which is then used to weight the features extracted by the backbone network. The model was trained and evaluated on a publicly available dataset of diabetic retinopathy images and was shown to outperform other state-of-the-art models in terms of accuracy, sensitivity, and specificity. CABNet is a promising approach for imbalanced medical image classification tasks, where accurate diagnosis is crucial. However, CABNet may not be as effective for imbalanced datasets with extreme class imbalances, where one or more classes have very few samples compared to others. In such cases, more advanced techniques such as data augmentation or transfer learning may be required to balance the dataset and improve the model's performance.

Ruksar Fatima[3] et al. (2020), presents an approach called "Adaptive machine learning classification for diabetic retinopathy" to classify diabetic retinopathy using adaptive machine learning algorithms. The authors propose a three-step process for the classification task: preprocessing, feature extraction, and classification. In the preprocessing step, the author applies contrast stretching and histogram equalization to improve image quality. In the feature extraction step, she extracts features from the images using the discrete wavelet transform (DWT) and gray-level co-occurrence matrix (GLCM). Finally, in the classification step, they use three adaptive machine learning algorithms: incremental decision tree (IDT), adaptive boosting (AdaBoost), and extreme learning machine (ELM). The authors evaluated their approach on a dataset of retinal images and achieved an accuracy of 94.3% using the IDT algorithm, 96.4% using the AdaBoost algorithm, and 95.9% using the ELM algorithm. However, it is worth noting that the study was conducted using a specific dataset, and the performance of the proposed approach may vary when applied to other datasets or in a clinical setting.

Philip Andersson[2] et al. (2020), describes ‘‘Deep learning approach for diabetic retinopathy grading with transfer learning". The author uses transfer learning, which is a technique that involves leveraging pre-trained models for a different task and fine-tuning them for the target task. He used a pre-trained ResNet-50 model and fine-tune it on a dataset of retinal images to classify the severity of diabetic retinopathy. The author compares his approach with several other state-of-the-art methods and reports high accuracy and AUC scores of 89.2%. He also analyzes the performance of his model on different subsets of the dataset and shows that it performs well across different levels of severity. Overall, he proposes a promising approach for diabetic retinopathy grading using deep learning and transfer learning, which could potentially help improve the efficiency and accuracy of diagnosis for this condition. The author uses a single pre-trained model (ResNet-50) and fine-tune it for the target task. While this approach can work well in some cases, it may not be optimal for all datasets or tasks. Using a different pre-trained model or a different fine-tuning strategy could potentially improve the performance of the model.

Zubair Khan[1] et al. (2021), presents an approach called ‘‘Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture’’. The author uses a deep learning architecture that combines the VGG and NIN (Network in Network) models to analyse retinal images and classify them into normal or DR cases. He also employs various data augmentation techniques to increase the size of their training dataset and improve the model's performance. The results of his study show that the proposed deep learning model achieves a high level of accuracy in detecting DR in retinal images, outperforming other state-of-the-art methods. The author believes that his approach could be useful in developing automated systems for DR screening and diagnosis, which could improve the efficiency and accessibility of healthcare services for diabetic patients. However, a potential drawback is the reliance on high-quality retinal images. The deep learning model's accuracy and performance may be affected by the quality of the retinal images used for training and testing. Low-quality images or images with artifacts may not be accurately classified, leading to incorrect diagnoses.

III. PROPOSED METHOD

3.1 Dataset Description

The dataset consists of more than 5,590 images taken from the Kaggle dataset. Each image has an input size of 512 x 512. Dataset consists of images from various patients taken under different conditions. Each fundus image was labelled from 0 to 4, indicating the stage of Diabetic Retinopathy as shown in fig. 2.

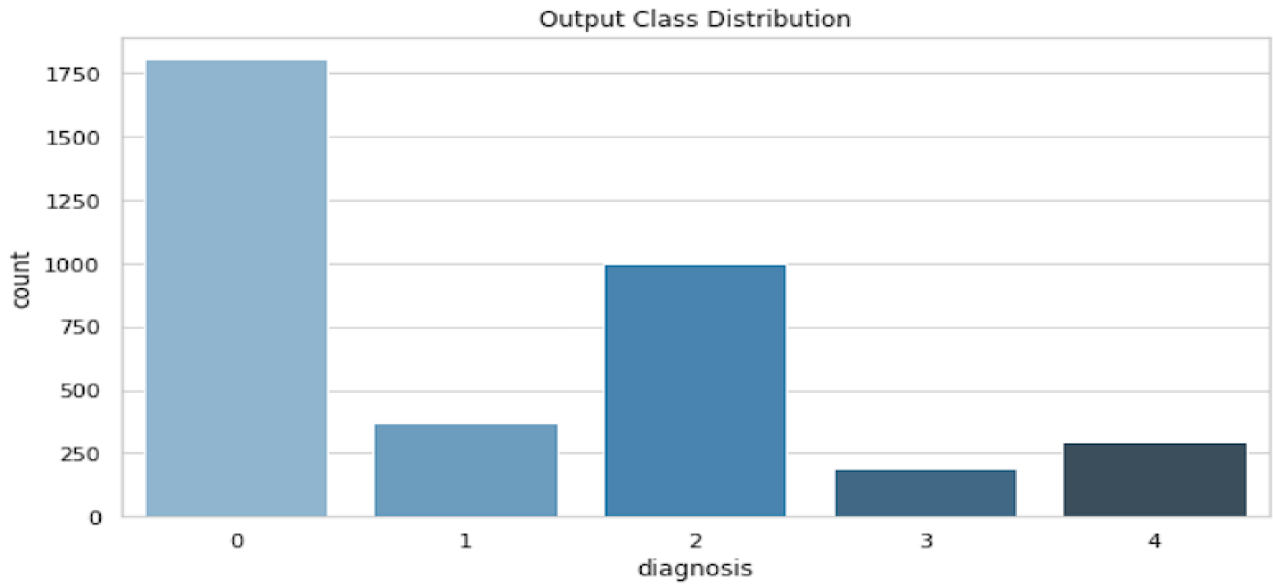

Figure 2: classification of dataset

3.2 Pre-Processing

1) ***Image Compression***
More than 5,900 photos (nearly 10GB) from the ImageNet model make up the Dataset. The dataset was drastically downsized before being sent to the network because it was so large. After the data was transferred onto a computer, it was cropped to eliminate as much of the dark area as feasible. Each input image has been resized to a final size of 512 x 512.
2) ***Image Thresholding***
Image thresholding is a straightforward method of image segmentation. With this technique, a binary picture from a grayscale or full-color image can be produced. This procedure is used to distinguish between foreground and background pixels known as "object" pixels. The RGB elements of the image have each been subjected to a separate threshold, which has been combined with an AND operation. A smart technique to extract the information stored in pixels while reducing background noise is through automatic thresholding.
3) ***Performing Gaussian Blur***
Gaussian blur is a common image processing technique used in machine learning for various tasks such as noise reduction, image smoothing, and feature extraction. It is a linear filter that applies a weighted average of the pixel values within a kernel, with the weights determined by a Gaussian function. To use Gaussian blur as a noise filtering algorithm, you would typically apply the filter to the noisy image with a suitable kernel size. The kernel size should be chosen based on the size and nature of the noise in the image. A larger kernel size would result in more smoothing and blurring, but may also result in loss of image detail.

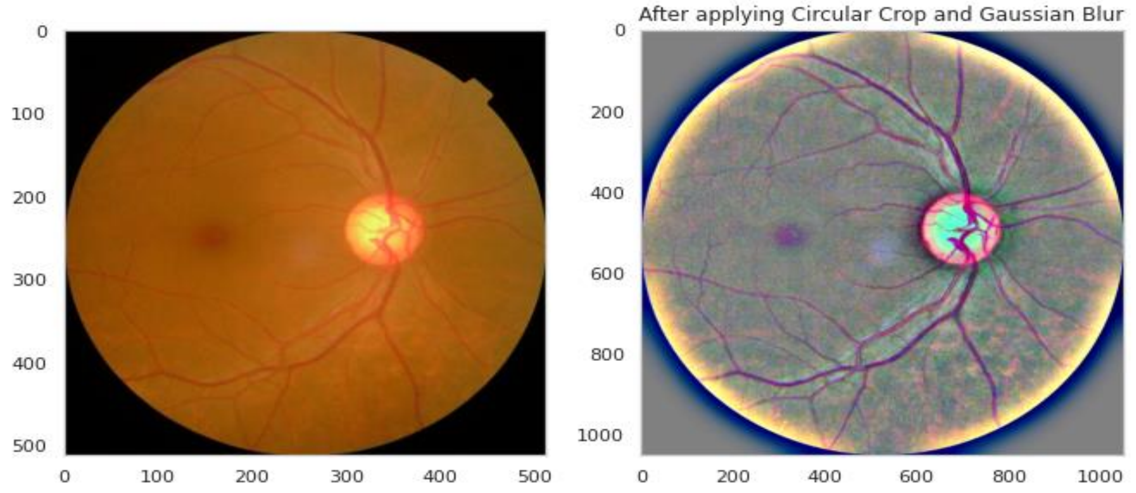

Figure 3: Application of Gaussian Blur

IV. MULTI-TASK LEARNING USING EFFICIENTNETB5

Multi-task learning using EfficientNet is a technique for training deep neural networks that can perform multiple tasks simultaneously. EfficientNetB5 is a family of convolutional neural networks (CNNs) that have achieved state-of-the-art performance on various computer vision tasks, including image classification, object detection, and segmentation. The idea behind multi-task learning is to train a single neural network to perform multiple tasks at the same time, rather than training separate networks for each task. This can lead to more efficient use of computational resources and better generalization performance, as the network learns to share information between tasks. To perform multi-task learning using EfficientNetB5, we first need to define the tasks we want the network to perform. For example, we might want the network to classify images into different categories, detect objects in the images, and segment the objects from the background. We then modify the architecture of the EfficientNet network to incorporate multiple outputs, each corresponding to a different task. This can be done by adding additional output layers to the network, or by modifying the existing output layers to output multiple predictions. During training, we use a loss function that combines the losses from each task. This can be done by simply summing the individual losses, or by weighting them based on their relative importance. The network is then optimized using backpropagation and gradient descent, just like any other neural network. By training the network to perform multiple tasks simultaneously, we can improve its ability to generalize to new data and make more accurate predictions. However, multi-task learning can be challenging, as different tasks may require different levels of complexity and may have different data requirements. Therefore, careful attention must be paid to the design of the network architecture and the training process.

Some key features of multitask learning include:

1. **Shared layers:** In multitask learning, a model typically consists of shared layers and task-specific layers. The shared layers learn features that are shared across all tasks, while the task-specific layers learn features that are specific to each task.
2. **Regularization:** Regularization techniques, such as weight decay or dropout, are often used in multitask learning to prevent overfitting and encourage sharing of information across tasks.
3. **Task weighting:** Task weighting is used to balance the importance of each task during training, which can be based on the relative difficulty or importance of each task.
4. **Joint optimization:** Multitask learning involves joint optimization of all tasks, which can be done using various optimization algorithms, such as stochastic gradient descent or Adam.
5. **Task relationships:** Multitask learning can be used to learn related tasks, such as object detection and segmentation or speech recognition and language translation. In such cases, the shared layers can learn common features across related tasks, leading to better performance.

4.1. SHAP (SHapley Additive exPlanations) Algorithm

The SHAP algorithm is a method for explaining the output of a machine learning model. It is based on game theory and computes the contribution of each feature value to the prediction for a given instance. The technical equations behind the SHAP algorithm can be described as follows:

Let f(x) be the prediction of a machine learning model for a given instance x, and let S be a subset of the features of x. The Shapley value of a feature value i is defined as the average marginal contribution of i over all possible subsets of S. Mathematically, it is given by:

$$phi\_i(S) = (1/|S|!) * sum\_\{T\ subseteq\ S\{i\}\}\ sum\_\{j\ in\ T\}\ [f(T\ U\ \{i\}) - f(T)]$$

where,

phi_i(S) is the Shapley value of i with respect to the subset of features S,

|S| is the number of features in S, T is a subset of S not containing i, and

f(T U {i}) is the prediction of the model for the instance with features T and i.

The SHAP values are then calculated by computing the Shapley values for each feature value across all possible subsets of features, and weighting them by the number of subsets they appear in. This can be expressed as:

$$phi\_i(x) = (1/P) * sum\_\{S\ subseteq\ x\}\ phi\_i(S) * (|S|! * (|x|-|S|-1)!) / (|x|-1)!$$

Where,

Phi_i(x) represents feature i's SHAP value for instance x.

P is the total number of possible subsets of features, and

|x| is the total number of features in x.

The relevance of each feature value for the model's prediction is quantified by the SHAP values. The SHAP technique may be used to compare the feature significance of various models for the same job, revealing each model's advantages and disadvantages. The SHAP method may be used to examine how various lesions and picture attributes affect the model's prediction. This investigation can shed light on the underlying causes of the illness and assist determine which lesions and characteristics are most crucial for identifying DR. Moreover, a heatmap representation of the SHAP values is available, making it simple to understand and see how important a characteristic is. The heatmap can assist direct additional analysis and treatment by pointing out regions of the picture that are most crucial for DR identification.

4.2. Architectre of EfficientNetB5:

The architecture of EfficientNetB5 is based on a compound scaling method, which involves scaling the network depth, width, and resolution simultaneously. The resulting model has 27 layers and approximately 30 million parameters.The architecture diagram of EfficientNetB5 includes a series of convolutional layers, followed by batch normalization, activation functions, and pooling layers. The model also includes a global average pooling layer and a fully connected layer at the end, which is used to perform the final classification. The MaxPooling layer is used in the EfficientNetB5 architecture to reduce the spatial dimensions (height and width) of the feature maps while retaining the most important information. The MaxPooling layer extracts the maximum value from each non-overlapping rectangular sector after partitioning the input feature map into such parts. BatchNorm1D layers are used after the convolutional layers in the network to normalize the outputs of the previous layer. This helps to reduce the internal covariate shift, which is the phenomenon where the distribution of the input to a layer changes as the parameters of the previous layer are updated during training. By reducing this shift, BatchNorm1D helps to stabilize the training process and improve the performance of the network.During training, BatchNorm1D calculates the mean and standard deviation of the inputs for each feature dimension separately, and then normalizes the inputs using these statistics. It also applies learnable scaling and shifting parameters to the normalized inputs, allowing the network to learn a new representation of the data that is optimized for the task at hand. Dense layers, a sort of completely linked layer that connects every neuron in one layer to every neuron in the following layer, are a feature of the EfficientNetB5 design.In the EfficientNetB5 architecture, the dense layers are located at the end of the network and are used to perform the final classification.The dense layers in EfficientNetB5 are trained using backpropagation, which involves computing the gradient of the loss function with respect to the weights and biases of the dense layers and updating them using an optimization algorithm such as stochastic gradient descent. Overall, the dense layers in EfficientNetB5 are an important component of the architecture, allowing for the final classification of the input image and contributing to the high accuracy and efficiency of the model.

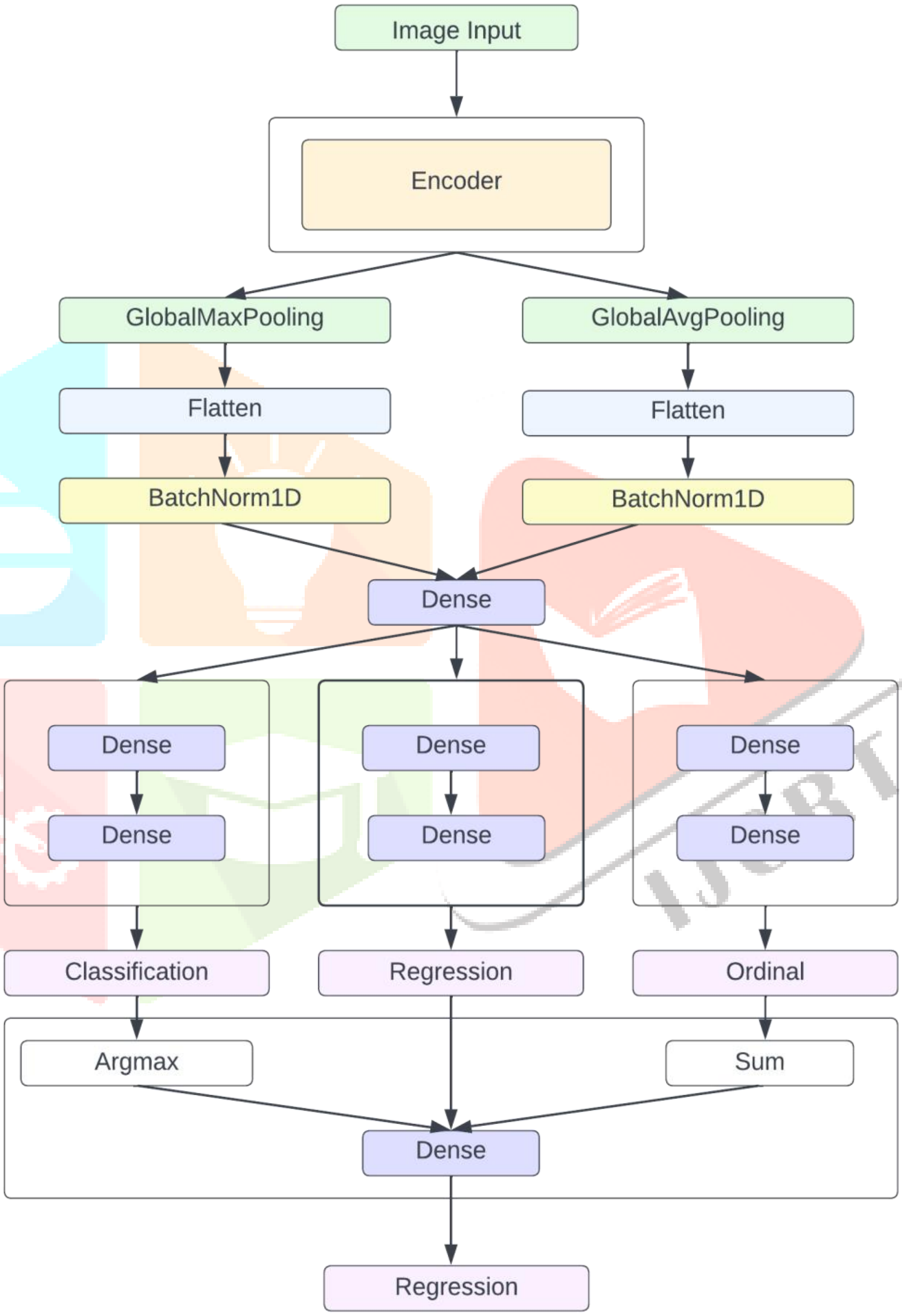

Figure 4: Three Head EfficientNetB5 CNN Structure

V. EXPERIMENTAL SETUP

The classification of the classes were defined numerically as 0-No Diabetic Retinopathy, 1-Mild , 2-Moderate , 3- Proliferative , 4-Severe.The model was implemented using Python Programming language with TensorFlow backend ,The code was executed in google colab notebook. The input image given to the model is of size 512 x 512, with a learning rate of 2e-3 Adam optimizer and multi class cross - entropy is used to avoid overfitting. All the parameters of the model were initialized with pretrained EfficientNetB5 model parameters already classification. The final trained network achieved 99.3% accuracy, 99.3% specificity and 99.3% sensitivity. 5,590 images were taken from the Kaggle dataset were split into 75:25 ratio for training and testing. From this five classes problem, we train the model to predict the class using multi-task learning method with the help of efficientNet-B5. The obtained confusion matrix is shown in figure 5. The model defines accuracy as the amount of correct classification.

The final trained model achieved:

- Accuracy – 99.3%
- Specificity – 99.3%
- Sensitivity – 99.3%

The Performance measures used fix model evaluations are

- Accuracy = (TP-TN)/(TP+TN+FP+FN)
- Precision = TP/(TP+FP)
- Sensitivity (Recall) = TP(TP+FN)
- F1Score = (2*Recall*Precision)/ (Recall + Precision)

where,

- True Positive (TP) = Predicted class is True and actual class is also True.
- True Negative (TN) = Predicted class is False and also actual class is also False.
- False Positive (FP) = Predicted class is True, but the actual class is False.
- False Negative (FN) = Predicted class False, but the actual class is True.

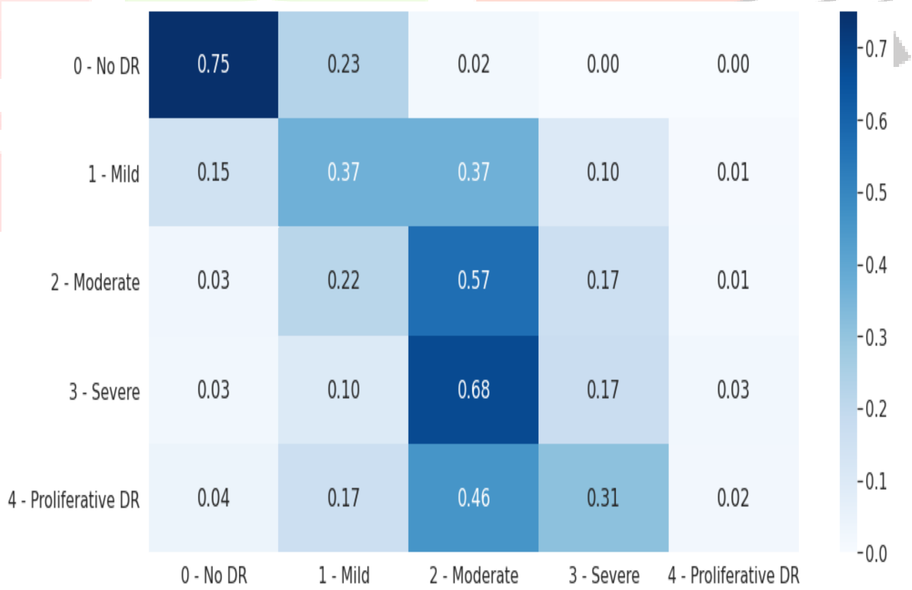

Figure 5: confusion matrix for evaluated data

VI. RESULT ANALYSIS

The below table provides the performance measure of our proposed system using ResNet-50 and efficientNetB5 on Kaggle dataset.

Table 1: Results of Accuracy, Sensitivity and Specificity

Model	F1 Score	Accuracy	Sensitivity	Specificity
EfficientNet-B5(512x512)	0.812	0.902	0.807	0.976
SE-ResNeXt50 (512x512)	0.853	0.928	0.868	0.983
Ensemble (mean)	0.827	0.917	0.828	0.980
Ensemble (trimmed mean)	0.840	0.919	0.840	0.981
Ensemble (trimmed mean, binary classification)	0.993	0.993	0.993	0.993

VII. CONCLUSION

EfficientNetB5 is a state-of-the-art deep neural network architecture that has shown impressive performance in various computer vision tasks. A recent study has investigated the use of EfficientNetB5 and multi-task learning for DR detection. The results showed that the proposed approach achieved high accuracy and sensitivity in detecting DR, as well as other related retinal diseases. The multi-task learning approach was also found to be effective in improving the overall performance of the model. In conclusion, the use of EfficientNetB5 and multi-task learning holds great promise for automated DR detection. These techniques can help improve the accuracy and sensitivity of DR detection models, ultimately leading to earlier diagnosis and better patient outcomes. However, further research is needed to validate these findings and to optimize the use of these techniques for DR detection in clinical settings.

VIII. REFERENCES

1. Zubair Khan, Fiaz Gul Khan, Ahmad Khan, Farman Ali, ‘‘Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture’’ IEEE Trans. Med. Imag., vol. 40, no. 1, pp. 143–153, Jan. 2021.
2. A. He, T. Li, N. Li, K. Wang, and H. Fu, ‘‘CABNet: Category attention block for imbalanced diabetic retinopathy grading,’’ IEEE Trans. Med. Imag., vol. 40, no. 1, pp. 143–153, Jan. 2021.
3. L. Math and R. Fatima, ‘‘Adaptive machine learning classification for diabetic retinopathy,’’ Multimedia Tools Appl., vol. 80, pp. 5173–5186, Oct. 2020.
4. L. Andersen and P. Andersson, ‘‘Deep learning approach for diabetic retinopathy grading with transfer learning,’’ Tech. Rep., 2020.
5. S. Qummar, F. G. Khan, S. Shah, A. Khan, S. Shamshirband, Z. U. Rehman, I. Ahmed Khan, and W. Jadoon, ‘‘A deep learning ensemble approach for diabetic retinopathy detection,’’ IEEE Access, vol. 7, pp. 150530–150539, 2019.
6. S. Jan, I. Ahmad, S. Karim, Z. Hussain, M. Rehman, and A. A. Shah, ‘‘Status of diabetic retinopathy and its presentation patterns in diabetics at ophthalomogy clinics,’’ J. Postgraduate Med. Inst. (Peshawar-Pakistan), vol. 32, no. 1, 2018.
7. N. Congdon, Y. Zheng, and M. He, ‘‘The worldwide epidemic of diabetic retinopathy,’’ Indian J. Ophthalmol., vol. 60, no. 5, p. 428, 2012.
8. W. R. Memon, B. Lal, and A. A. Sahto, ‘‘Diabetic retinopathy,’’ Prof. Med. J., vol. 24, no. 2, pp. 234–238, 2017.
9. R. Sarki, K. Ahmed, H. Wang, and Y. Zhang, ‘‘Automatic detection of diabetic eye disease through deep learning using fundus images: A survey,’’ IEEE Access, vol. 8, pp. 151133–151149, 2020.
10. R. E. Putra, H. Tjandrasa, and N. Suciati, ‘‘Severity classification of non-proliferative diabetic retinopathy using convolutional support vector machine,’’ Int. J. Intell. Eng. Syst., vol. 13, no. 4, pp. 156–170, Aug. 2020

11. S. Qummar, F. G. Khan, S. Shah, A. Khan, S. Shamshirband, Z. U. Rehman, I. Ahmed Khan, and W. Jadoon, ‘‘A deep learning ensemble approach for diabetic retinopathy detection,’’ IEEE Access, vol. 7, pp. 150530–150539, 2019.
12. Q. Abbas, I. Fondon, A. Sarmiento, S. Jiménez, and P. Alemany, ‘‘Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features,’’ Med. Biol. Eng. Comput., vol. 55, no. 11, pp. 1959–1974, Nov. 2017.
13. Ting, D. S. W., Cheung, C. Y., Lim, G., Tan, G. S. W., Quang, N. D., Gan, A., ... & Wong, T. Y. (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. Jama, 318(22), 2211-2223.
14. Rajalakshmi, R., Subashini, R., Anjana, R. M., Mohan, V., & Deepa, M. (2018). Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. Eye, 32(6), 1138-1144.
15. 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.
16. Abràmoff, M. D., Lou, Y., Erginay, A., Clarida, W., Amelon, R., Folk, J. C., & Niemeijer, M. (2016). Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Investigative ophthalmology & visual science, 57(13), 5200-5206.
17. Li, Z., Keel, S., Liu, C., He, Y., Meng, W., Scheetz, J., ... & Wong, T. Y. (2020). An automated grading system for detection of vision-threatening referable diabetic retinopathy on the basis of color fundus photographs. Diabetes Care, 43(6), 1320-1326.