Diabetic Retinopathy Detection

Devansh Arora Aayush Kapoor Krishnam Omar

1 Problem Statement

Diabetic retinopathy is a leading cause of blindness worldwide, affecting over 93 million people. Almost half of Americans with diabetes have some stage of the disease, which can be difficult to detect until it's too late for effective treatment. Current detection methods require trained clinicians and are resource-intensive, which can limit their effectiveness in areas where diabetes rates are high. Automated detection systems using image classification, pattern recognition, and machine learning have made progress, and this competition aims to push them to their limit for maximum impact on improving DR detection. Winning models will be open-sourced.

2 Related Work

The article reviews the use of deep learning, particularly convolutional neural networks, in the detection and classification of diabetic retinopathy (DR) from color fundus images. The authors analyze recent state-of-the-art methods and review available datasets for DR detection. They also discuss challenging issues that require further investigation. Deep learning has shown to be highly effective in medical image analysis, and the authors highlight its potential in DR detection and classification.[1] The study presents a diabetic retinopathy detection system that uses ultra-wide-field fundus photography and deep learning, which is more efficient than conventional fundus photography used in most automatic systems. The researchers used the early treatment diabetic retinopathy study 7-standard field image extracted from ultra-wide-field fundus photography and found that it outperformed the optic disc and macula-centered image in detecting diabetic retinopathy in experiments. The results suggest that ultra-wide-field fundus photography can be a useful tool for automated screening and grading of diabetic retinopathy.[2] The article proposes a deep learning method for detecting diabetic retinopathy that uses a regression activation map (RAM) to provide interpretable features. The RAM is added after the global averaging pooling layer of the convolutional neural network (CNN), allowing the model to identify and localize discriminative regions of the retina image and show the specific region of interest in terms of its severity level. This approach provides insights into why the learning model works and is highly desired in practice, as users are interested not only in high prediction performance but also in understanding the insights of DR detection. [3]

3 Dataset Description

The dataset we use was created by Google AI and EyePACS, a non-profit organization that aims to prevent blindness through the early detection and treatment of eye diseases. The dataset was released as part of a Kaggle competition in 2015. The dataset contains over 35,000 retinal images, of which approximately 75% have been labeled for diabetic retinopathy severity on a scale of 0 to 4. The images are in JPEG format and have a resolution of 224 x 224 pixels. The images have been preprocessed to remove patient-identifying information and resized to a standard resolution. We worked on a subset of the given dataset as the dataset was too big. The class distribution of our dataset can be visualised from figure 1 and figure 2.

We plot the Image Histogram(figure 3) over our dataset. The histograms are created for each channel of the images to see the distribution of pixel values for each class.

[1] https://www.sciencedirect.com/science/article/pii/S2352914820302069

[2] https://www.nature.com/articles/s41598-021-81539-3

[3] https://paperswithcode.com/paper/diabetic-retinopathy-detection-via-deep

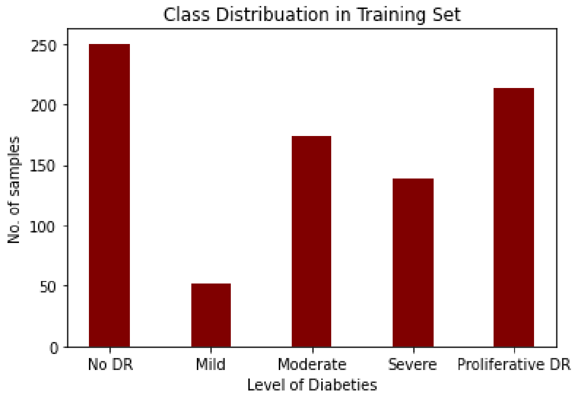

Figure 1: Class Distribution over our Training set

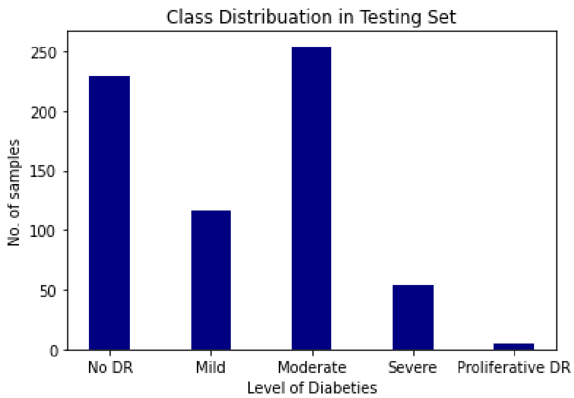

Figure 2: Class Distribution over our Testing set

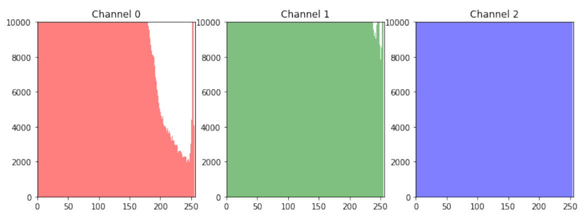

Figure 3: Color distribution over all pixels in our sample dataset

4 Experimental Setup

Diabetic Retinopathy has been provided with high-resolution images broadly divided into five main categories:

```
0- No DR
1 - Mild
2 - Moderate
3 - Severe
4 - Proliferate
```

The dataset is organized into TRAIN and TEST folders, each with five sub-folders corresponding to the disease categories. Image paths from each folder are stored in a list and shuffled into the train, validation, and test sets using train test split 70 percent training, 20 percent validation, and 10 percent testing. Pytorch's Dataset is then applied to this list, which reshapes and applies transformations to ensure uniformity. This pre-processed dataset is passed into a DataLoader with a batch size of 16 and split into three such DataLoader objects. Once this setup is complete, the images can be trained using various state-of-the-art models.

5 Observations

The data is processed, normalized, and then tested on 3 state-of-art architectures for image classifications. The architectures are VGG16, Resnet50, and Inception netV3 (all pre-trained on image-net weights), and 4 epochs were run for them. The Observations are as follows in fig 4:

Architecture	Loss	Accuracy	Precision	F1
VGG16	Training:1.604 Validation:1.608 Testing:1.580	Training:0.300 Validation:0.296 Testing:0.322	Training:0.2372 Validation:0.2464 Testing:0.2309	Training:0.115 Validation:0.111 Testing:0.108
Resnet50	Training:1.435 Validation:1.379 Testing:1.41	Training:0.45 Validation:0.508 Testing:0.458	Training:0.39 Validation:0.408 Testing:0.385	Training:0.317 Validation:0.332 Testing:0.291
InceptionNetV3	Training:1.40 Validation:1.411 Testing:1.350	Training:0.482 Validation:0.466 Testing:0.564	Training:0.431 Validation:0.387 Testing:0.505	Training:0.382 Validation:0.342 Testing:0.463

Figure 4: Observation table

Based on the results presented in the table, it is evident that Resnet50 and InceptionnetV3 achieve the highest performance scores. However, due to the limited availability of computational resources, the models were trained for only four epochs. It is worth noting that optimizing these models further and achieving even higher scores with additional resources is possible.

After passing the testing data through the model, the confusion matrix is as follows in fig 5: Upon analyzing the matrix, it is evident that the model exhibits a high degree of accuracy in detecting individuals with no diabetic retinopathy ('No DR'). However, it shows some limitations in accurately identifying the severity level of the condition, particularly in distinguishing between Moderate and Severe types. One potential solution is to augment the dataset by including more Severe and Moderate samples. This would allow the model to understand better the patterns and characteristics of these types of DR and improve its ability to classify them correctly. The inception net model's underlying architecture serves as the foundation for the analysis, and the data is visualized through TSNE plots to

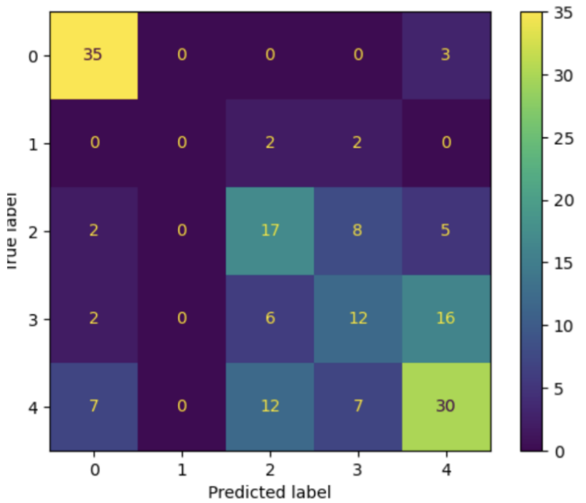

Figure 5: Confusion Matrix

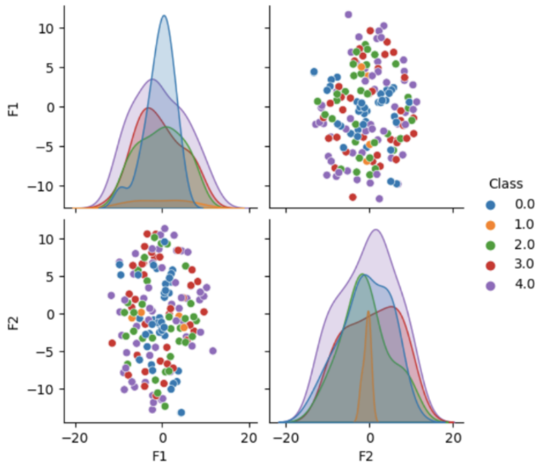

Figure 6: TSNE plot

reduce the dimensions and assess their separability. The plots demonstrate that the instances belonging to class 4 are dispersed near the borders of the distribution, while those from class 0 are situated towards the center. Conversely, the other categories exhibit no discernible trends or distinguishing features that set them apart.

6 Future Work

The future work in image classification for diabetic retinopathy using attention and ensemble learning techniques holds great promise. Attention-based approaches can help to identify the most critical regions of an image for diagnosis, while ensemble learning can combine the strengths of multiple models to improve classification accuracy. During future work, we will develop more sophisticated attention mechanisms and explore new ways to connect various models. Additionally, more extensive datasets and diverse patient populations will be essential to validate these techniques' effectiveness and develop models that can generalize to new data. Ultimately, the use of attention and ensemble learning in diabetic retinopathy classification can significantly improve the accuracy and efficiency of diagnosis, leading to better patient outcomes.

7 References

Dataset for the problem- https://www.kaggle.com/c/diabetic-retinopathy-detection

Diabetic retinopathy detection through deep learning techniques- https://www.sciencedirect.com/science/article/pii/S2352914820302069

Early detection of diabetic retinopathy based on deep learning and ultra-wide-field fundus images- https://www.nature.com/articles/s41598-021-81539-3

Diabetic Retinopathy Detection via Deep Convolutional Networks for Discriminative Localization and Visual Explanation- https://paperswithcode.com/paper/diabetic-retinopathy-detection-via-deep