DIABETIC RETINOPATHY DETECTION

Universität Stuttgart

Ram Sabarish Obla Amar Bapu (st169693) and Swetha Lakshmana Murthy (st169481)

Institute of Signal Processing and System Theory, University of Stuttgart, Germany

1. Introduction

- At the outset, the detection of Diabetic Retinopathy(DR) in its earlier stages required labor, background, and time. For the supervision of fundus images employing trained experts was necessary. Therefore, manually screening the DR elevated inconsistency among experts who conducted the medical diagnosis. Gradually resulting in developing a deep learning model for enhancing the screening process of DR.
- The proposed deep learning model here employs a Convolutional Neural Network(CNN) architecture for binary classification tasks to predict whether the patient has referable or non-referable DR.

2. Data and Features

- The Indian Diabetic Retinopathy Image Dataset(IDRID) is the dataset used for the model. A total of **516** fundus images are available here. It is categorized into **413** (80%) training images and the rest **103** (20%) for testing. The ground truths of the images are available in the form of a CSV file.

3. Data Preprocessing

- **Input Pipeline Creation:** The original IDRID dataset contains severity grades from 0-4 of DR. An emphasis is laid upon binary image classification. Hence, the labels are converted to **NRDR (Labels - 0,1)** and **RDR(Label - 2,3,4)**.
- The training dataset is split into training and validation. We observe that the dataset is imbalanced. Hence, a **random oversampling** method is applied to the training dataset. This is crucial for obtaining a precise model with an equal number of samples belonging to all classes.
- Before mapping the dataset,the images are cropped and normalized to a pixel size in the range(0,1) and finally resized to a resolution of **256 X 256** pixels without distortion.
- **Data Augmentation:** To bolster the diversity of the training dataset and reduce over-fitting,the images are **rotated, flipped, shifted and zoomed**. Subsequently mapping and prefetching is carried out to optimize the performance.

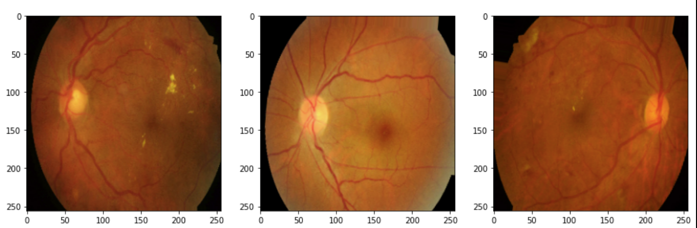

Figure 1: *Augmented Images*

4. Model Architecture

- The model consists of a cascade of convolutional layers with ReLu as activation function followed by max-pooling layers correspondingly.
- The dense layers, in the end, have the ReLu as activation function and softmax activation function respectively. To avoid over-fitting dropout layers are added.
- **Training:** The **batch size** used here is **32**. The hyper-parameters of the models are optimized using the hyper-parameter tuning approach. The model is trained for about **150 epochs**. An **adam optimizer** and the sparse categorical cross-entropy loss function is used.

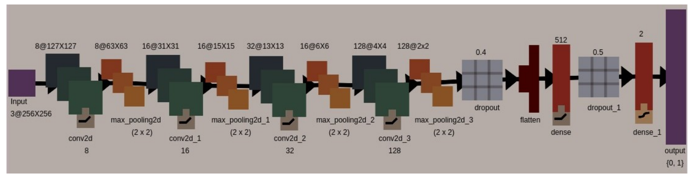

Figure 2: *Model Architecture*

5. Evaluation and Metrics

- The model is trained and evaluated with the test dataset. The obtained test accuracy is **72.81%**

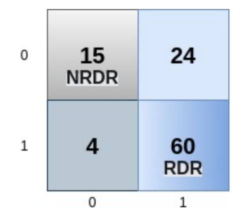

Figure 3: *Metric : Confusion Matrix*

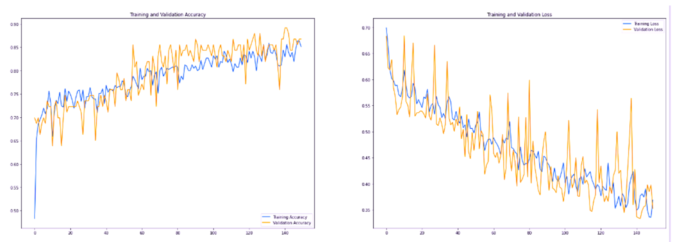

Figure 4: *Train and Validation Accuracy and Loss*

6. Deep Visualization Methods

- The **Grad-CAM** localizes image regions with relevance and visualizes gradients of the final conv2D layer important for classification.
- The **Guided Back-propagation** visualizes gradients with respect to the image where it nullifies the gradients associated with negative value. Eventually aiming to capture pixels detected by neurons.

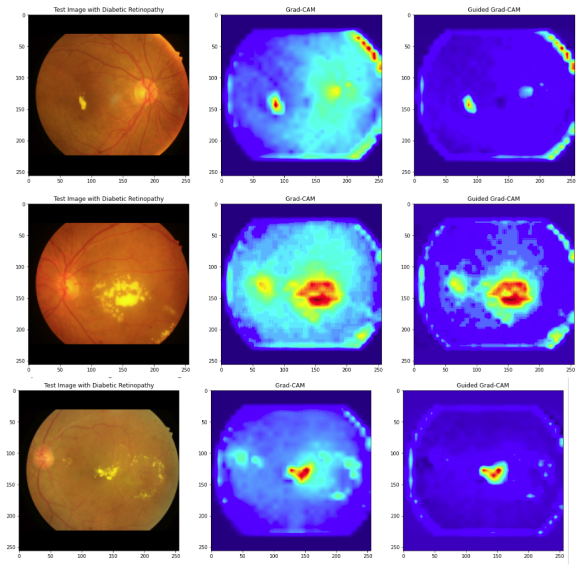

Figure 5: *Deep Visualization*

7. Conclusion

- Deploying a model for predictive DR progression enables early detection of patients with the highest risk of vision loss. This further enhances the timely referral to a retinal specialist and aims at preventing irreversible vision impairments.

Deep Learning Lab 2021, February 7th, Stuttgart, Germany