Methodology

# Pre-processing:

The dataset consists of several images and will hence need to be pre-processed appropriately in order to make the data suitable for usage in Machine Learning Techniques. Several techniques to pre-process the images exist, however, existing literature was surveyed wherein the following methodologies were proposed.

Pre-processing is required to ensure that the dataset is consistent and displays only relevant features. This step is necessary to simplify the workload of the following processes. Next, the images are segmented to differentiate between the normal and abnormal substances. Several features can be obtained from the images, some of them are: Exudate Number, Exudate Area, Micro aneurysms etc. Of the three colour channels in the image (Red, Green, and Blue) the contrast between the blood vessels, exudates and haemorrhages is best seen in the **Green channel** and this channel’s neither under- illuminated nor over-saturated like the other two. **Contrast Enhancement** to further enhance the features of the image, **Cropping and Resizing,** since the original images vary widely in size and some images were chopped at the top and bottom, they had to be standardized. It is observed that **Exudate Area** is the best feature out of all the features which can primarily be used for diabetic detection, followed by blood vessels and other features.

# Machine Learning:

Pre-processed images will then be subjected to several Machine Learning Techniques, in order to generate an optimal model. Due to the data being in the form of images, CNNs will be of importance in generating the model, whose layer by layer breakdown will be understood during the implementation. Multiclass Classifier is generated which will classify the data into one of the following classes: 0 – No DR, 1 – Mild, 2 – Moderate, 3 – Severe, 4 – Proliferative DR.

# Web Application:

The trained model will be then made accessible to the end user via a web application, wherein they will be asked to input their Left and Right Fundus images and will then be provided the result in the form of one of the above mentioned classes.

**Note**: The proposed methodology is pure speculation due to the very early stages of implementation and surveying. Further refinements will be made as and when needed and the methodology is subject to change accordingly.

