**Title of research:**

Multi Class Fundus Image classification for Diabetic Retinopathy identification

**Objectives:**

1. To identify and resolve existing barriers in fundus image analysis for diabetic retinopathy classification
2. Automation and speedup of existing models and optimization of deployment pipelines such that wider range of doctors can utilize the technologically advanced product

**Introduction:**

Diabetes is one of the common disease in the world today and a lot of people are affected with it. Diabetes in not just related to glycogen conversion to glucose but if somebody has diabetes from long time it can affect their eyes as well. There is where we have picked up the idea for this project. Eye disease caused due to diabetes is often termed as diabetic retinopathy, and surprisingly it’s one of the highest frequently occurring eye disease when it comes to the retina based eye diseases. Scary it may sound, but the fact is that diabetic retinopathy is a very silent disease, which augments the blood vessels in the eye which might even result in abnormal vessels forming in the eyes, and it’s one of the disastrous side effect is that it can lead to eye blindness in the person who is suffering from diabetes from a very long time. It is very much recommended that people suffering from diabetes should go for doctor check regularly for its timely diagnosis, only disappointment here is that traditional diabetic retinopathy diagnosis is carried out manually where some doctors take digital images of our fundus through fundus camera and search for some hand crafted features in the images, and thus it’s diagnosis depends on the doctor carrying that out, and results of the diagnosis may vary according to his expertise levels and years of experience and number of cases he has worked upon until now. Even after all this shortcomings, the traditional solutions are very expensive as a very experienced doctor is required for examining with high accuracy, it takes a lot of time for the doctors to give results on one single cases. This makes it very difficult for the poor to afford this solution and early detection of disease becomes almost impossible for them.

**Related work:**

In the past some years, considerable research has been done in regard to the Image Processing based diabetic retinopathy classification and people have put appreciable efforts to solve this particular problem. Some of the traditional image processing methods that were used extensively are Hough transforms, difference of Gaussian methods, different filters like Gabor filters, various intensity variation methods etc. After feature extraction methods mentioned above, traditional object classification methods were used like k-means clustering, Support Vector Method based classification etc. In spite of those hard works even in today’s solution there exists a lot of shortcomings. In a lot of automated Diabetic Retinopathy screening devices hand crafted features has been used which are susceptible to many environmental condition like lighting, blurriness, motion artefacts due to small motion in eye while capturing images, electronic noises due to bad sensors. Even if we try to capture images in ideal conditions, a lot of devices only perform two class diabetic retinopathy classification which kind of don’t addresses the problem fully and is not suitable for this whole diabetic retinopathy framework and it’s requirements. In the recent few people have tried to tackle the problems face by traditional image processing methods by trying to utilize deep neural networks. Beginners usually rely on standard deep learning architectures like Alex NET and Google Nets to build their models. And based on the experiments these models outperforms the two stage processed by a good margin and is more robust to minor changes in the parameters. Even the trend in a recent Kaggle competition is that most of the people relied on CNN based architectures to achieve the submitted results they achieved. All the above mentioned convolutional nets requires complex neural networks and medical practitioners faces a lot of difficulty to understand the inferences. A convolutional neural network which can also help the doctors to understand the region responsible for classification into a certain category is also very much required.

**Proposed Plan:**

**Dataset:**

After exploring a lot of DR Datasets, Kaggle – EYE Pacs dataset looks to be the best to train and test the developed model. It has lost of variation and sufficient number of images representing 5 classes. Large number of images can give us the much required confidence that the developed model can be generalized very nicely. Distribution of data is as follows and displayed in the below histogram:

Figure 1 Represents downloaded data distribution

**Data Preprocessing:**

**Data Imbalance:**

In medical image classification data imbalance is a major problem faced by all data scientists around the world. In the above histogram you can clearly see that data distribution is highly skewed thus a variety of treatments will have to be tried to negate data imbalances. Some of the data curation methods will be described below

**Data Augmentation:**

Due to rarity of certain disease occurrence it is very natural to find less number of images belonging to that disease. In recent time people have employed a lot of data augmentation techniques to mitigate this problem. In Data augmentation we try to generate more copies of the already existing images through some mathematical and image processing techniques. These methods can be taken as analogous to oversampling and under sampling techniques. Sometime we can have a lot of normal images, in that case to make the number of data points same we under sample from normal or healthy images and over sample from affected images

1. Variance Of Laplacian  
   Laplacian of image gives edges present in an image. Blurry images have less edges so variance of edges will be lesser than as compared to sharp images. Thus a well-focused image is expected to have a high variation in grey levels. When we have abundance of images in class zero which are basically normal images , we can use VOL to discard bad images such as which don’t have good quality edges which actually would not help our CNN to understand properly

1. Histogram Equalization  
   when the images has been captured under uneven lighting condition, it happens sometimes that some pixels of images are not evenly lit. In that case we have to take away light from the pixels where intensity is more and distribute it among those pixels where intensity is less
2. Log enhancement  
   As sometimes after capturing an image we can see that dynamic regions of pixels remain very high, due to which a lot of features in images are not clearly visible as the pixel values remain very low, but the pixels values whose intensity is high should not be disturbed as well. In that case we have to apply some transformation which can increase the lower level pixel values and does not affect much to the higher level values. To solve this we add one to every pixel values and take log transformation
3. Sigmoid Correction  
   as the name suggest a continuous non-linear sigmoid function will be used to adjust and correct the bad images. It is a point process operator and controls the appearance of the image. It tries to adjust shadows and highlights present in the image in a good way such that all the features of the images are clearly visible.

**Model Development:**

In this thesis research we have planned to study various state of the art deep learning architecture, old and new, and to propose a new state of the art architecture which can cater to the shortcomings displayed by existing models

1. VGG Net  
   Zissermen and Simonyam first proposed VGG Net, which is a Convolutional Neural Network. The model was able to achieve staggering accuracy of 92.5 % in the prestigious imagenet challenge back then. Naturally it is one of the most favoured architecture when it comes to feature extraction. Authors have been kind enough to open source weight files of VGG Net, thus it can be used a backbone in many problem statements.  
      
   

Figure 2 VGG Net architecture

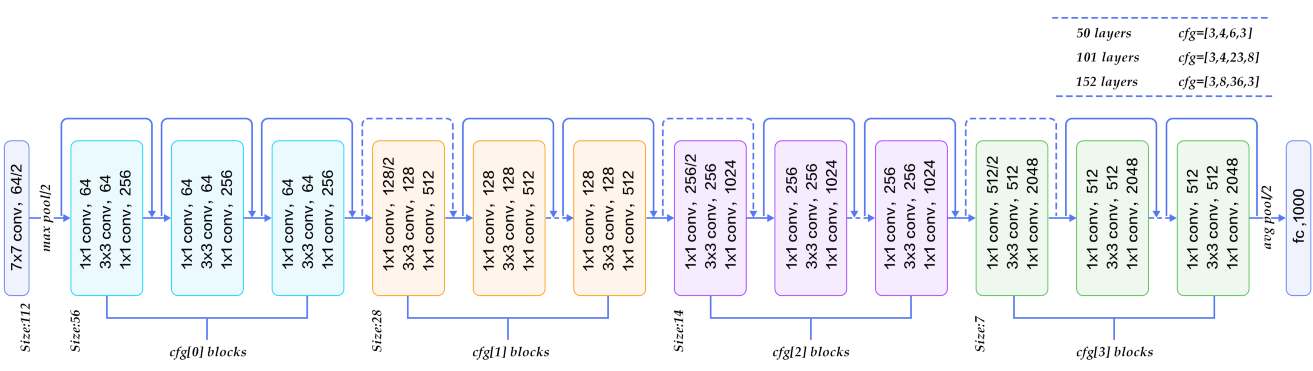
1. ResNET  
   Residual Neural Network also known as ResNET was introduced by Kaiming He et al. It consists of skip connections and heavy batch normalization. Such skip connections are called gated units or gated recurrent units resembling strongly to RNNs. ResNET consists of heavy 152 convolutional layers while still less complexity than VGGNet.  
     
   

Figure 3 ResNET architecture

1. Dense NET  
   In DenseNET all the layers pass its own and its preceding layer information to its succeeding layer. Thus each layers will have the information of the gradients which its previous layers have as shown in figure 4

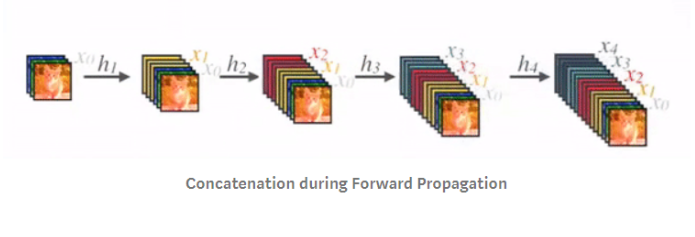


Figure 4 DenseNET architecture

**Transfer Learning:**

While using state of the artaadeep learning architectures, transfer learning will be tried out. For that, the last fully connected layer would be removed, thenaaa transfer learningaascenario would be followed by treating theaaremaining network componentsaaas a fixed featureaaextractor for theaanew dataset. The transfer learning retainsaainitial pre-trained modelaaweights and uses final network layer to extract image features.

**Results Evaluation:**

Since this is a classification based problem so Confusion Matrix will be used to generate the results. As the dataset is not fully balanced so instead of accuracy, precision, recall and f-score will be calculated for comparing different architectures.

**Please outline the proposed sample group, including any specific criteria:**

As the data used for this project is open source, a series of guidelines and methodologies were followed by the data aggregating organization. While referring to the report, it was mentioned that while doing collection of data variance of data being collected was given highest priority, in which people irrespective of their state, caste, color, race or ethnicity were requested to contribute for a noble project which will benefit the society. It was made sure that we will have equal representation of males and females, as well as children from all age group.

**Describe how the proposed sample group will be formulated:**

Data collection agency which made the data available informed that a call for representatives was sent out to targeted groups and communities in the areas of representation identified, and through the use of cascading methods.

**Indicate clearly what the involvement of the sample group will be in the research process, How their consent will be obtained, potential risks to them, Anonymity of data being collected:**

The main purpose of the sample group is to help us in collecting data which will help in the development of better healthcare devices and technologies which can be used for the betterment of the society. The result of this exercise was to develop a cutting edge technology which can be further sent to healthcare doctors for their review. Data that will be collected from customer will not have their name, address or any information that promotes racism. In cases of images being collected, all the meta-data from images will be removed to find any backpropagation way to locate the Participants. All the hardware devices use for collecting data would be properly sanitized and softwares would go through several antivirus scan so as to avoid any form of data leakage from our database. Consent of participants was taken on a paper, duly signed, by data collecting organization while collecting data

**Indicate any potential risks to people using the product how you propose to minimize these:**

The developed product will be a part of academic curriculum and unless it has cleared clinical trials, it should only be used a reference and not as a actual result. People should not use the product without getting validated from a certified health official. Results displayed by the products are meant to assist healthcare professionals and not replace them.

**Ownership of data collected**

Organization that collected data and made it open source is the full owner of data and the data has been provided to us mainly for academic research purpose.

**References:**

[1] HerbertaaBay, AndreasaaEss, Tinne Tuytelaars, and Luc Van Gool. 2008. Speeded-UpaaRobust Features (SURF). aaComput. Vis. ImageaaUnderst. 110, 3 (June 2008), 346-359. DOI=http://dx.doi.org/10.1016/j.cviu.2007.09.014

[2] Shah AR, aaGardner TW. Diabeticaaretinopathy: researchaato clinical practice. ClinaaDiabetes Endocrinol. 2017;3:9. aaaaPublished 2017 Oct 19. aadoi:10.1186/s40842-017-0047-y

[3] Gardner TW, aaChew EY. Future opportunitiesaain diabetic retinopathyaaresearch. Curr OpinaaEndocrinol DiabetesaaObes. 2016;23(2):91–96. aadoi:10.1097/MED.0000000000000238

[4] LeeaaR, WongaaTY, SabanayagamaaC. A two-stageaafilter foraaremoving salt-and-pepperaanoise usingaanoise detector basedaaon characteristic differenceaaparameter and adaptiveaadirectional meanaafilter. Eye Vis (Lond). 2015;2:17. Published 2015 Sep 30. doi:10.1186/s40662-015-0026-2

[5] Ankita Gupta, Rita Chhikara, Diabetic Retinopathy: aaPresent andaaPast, Procedia Computer Science, Volume 132, 2018, Pages 1432-1440, ISSN 1877-0509

[6] S. izza Rufaida and M. I. Fanany, “Residualaaconvolutional neuralaanetwork for diabeticaaretinopathy,” in Proceedings ofaaIEEE InternationalaaConference on AdvancedaaComputer Science andaaInformation Systems (ICACSIS). IEEE, 2017, pp. 367–374.

[7] D. Zhang et al., “Review of Videoaaand Image DefoggingaaAlgorithms and RelatedaaStudies on Image Restorationaaand Enhancement,” inaaProceedings of IEEE SmartWorld, UbiquitousaaIntelligence and Computing, Advancedaaand Trusted Computed, aaScalable Computing andaaCommunications, Cloudaaand Big Data Computing, aaInternet of People andaaSmart CityaaInnovation (SmartWorlda/SCALCOM/UICa/ATC/ CBDCom/IOP/SCI). IEEE, 2017.

[8] M. Alban and T. Gilligan, aa “EmpiricalaaEvaluation of RectifiedaaActivations inaaConvolutional Network.” Report ofaaStandford Education, 2016.

[9] Z. Wang et al., “A ComprehensiveaaStudy of Alzheimer'saaDisease ClassificationaaUsing Convolutional Neural Networks,” inaaProceeding of theaaInternational Conferenceaaon Medical ImageaaComputing and Computer-AssistedaaIntervention (MICCAI). Springer, aa2017, pp. 267–275

[10] Xavier Glorot, aaYoshua Bengioaa; Proceedings ofaathe ThirteenthaaInternationalaaConference on Artificial Intelligenceaaand Statistics, PMLR 9:249-256, 2010.