

# Smart Training of Alex Net using Fluorescein Angiography Fundus Images Implementing Selective Data Sampling

Shubham Chaudhary, Vipul Sharma, Akshay Khandelwal

Department of Electronics and Communication,  
SRM Institute of Science and Technology, Kattankulathur,  
Chennai, Tamil Nadu, India - 603203

**Abstract**—This research paper pertains to the concept of preliminary detection of Posterior Capsular Opacification, Glaucoma and Genetic Diseases in the human eye using Alex Net, which is one of the earliest CNN found in the MATLAB. CNN are trained with specific layers using the given dataset taking blood vessels as the parameter for detecting pre-abnormalities from the images. The specific layers are used for reducing the time for training and detection. The layers have been chosen based on the different parameters like filters, rate of training, initial rate, statistical parameters and mathematical modelling. After this the system is trained for 100 to 1000 epochs in which the test image is chosen as input and is compared with disease affected and normal blood vessels. It is found that for first training the system had a detection with an accuracy of 50 to 60% and with subsequent training it had an efficiency of 80-90%

**Keywords:** *Posterior Capsular Opacification, Convolutional Neural Network, ALEXNET, Regression-statistical model used in CNN, Sampling, fundus.*

## 1 INTRODUCTION

Computer vision has been used a lot in medical image processing but the novelty in our system is that we have used one of the oldest CNN keeping in mind the computational timing and the no. of epochs used for the training. However, training of these CNNs is time consuming and it requires a lot of computational power, this makes it a challenging process. Also, the pre-existing model of conventional neural networks has some inaccuracies when the no. of images in the dataset is high. In this paper we propose a unique method to improve the overall efficiency of classic CNNs, using the Alex Net CNN, this exists uniquely in MATLAB. Alex Net is an existing CNN which aid us in analysing medical conditions directing to more accurate results. Alex Net is a pre-trained CNN with 25 layers. Here we intend to use some specific layers on the basis of statistical model calculations instead of all layers, this drastically reduces the computational time and improves the decision making. The two predefined layers in Alex Net are image input layer and the classifier layer or the output layer. The input dataset is segmented blood vessels from disease diagnosed fundus images. The training data is sampled based on the classification obtained by the status quo of the CNNs. For each sample weights are assigned and the informative samples are included in the next iteration of the CNN. Linear regression methods and probabilistic sampling are performed on the training dataset. Training is performed between 15 to 60 epochs and gauging is done to determine

whether the person has preliminary glaucoma and other diseases.

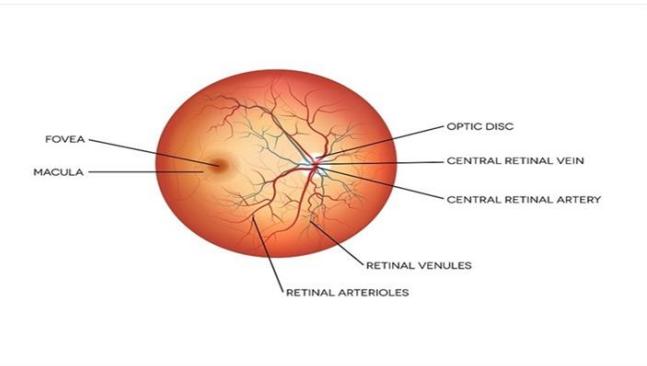


Fig. 1 . Essential Parts of a Human Eye

The essential parts of the human eye are Optical Disk, Macula, hameroge, Blood Vessels, Exudates and this the order of detection is optical disk and blood vessels are easily segmented but, exudates and hamerges are difficult to detect because of vast similarity, false detection and difficult segmentation methods which increases the complexity of the process of detection.

## 2 METHODS AND SYSTEM DESIGN

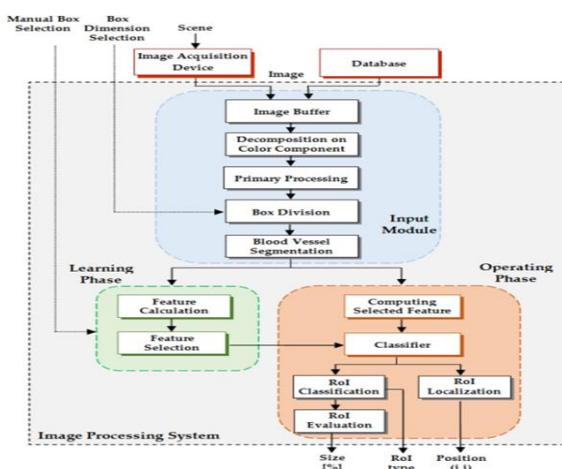


Fig. 2 . Blood Vessel Segmentation Methodology

### 2.1 Database information:

In the following paper we have taken datasets from MESSIDOR, keggle and cataract. In this we have taken the

fundus images in the form of ppm. After this we have used the images for blood vessel segmentation by doing many pre-processing methods which we will discuss now. MESSIDOR has around 1000 disclosed images each having pixels of around mydriatic retinography each having size of 1440\*1000, 2000\*1800 pixels which is used for independent testing.

**Kaggle database:** The database is having a lot of diabetic retinopathy and also having a lot of disclosed 50,000 images from this dataset we have used around 5000 image this is based on specified on the system parameters and this does not affect the system output because for the CNN training it is based on statistical measurements and also on the training of the system.

## 2.2 Pre-processing methods

In the pre-processing methods we have converted the images from rgb to green channel then to gray images. The green color is used because of physiological references. This is done for reducing the matrix calculations since the input images are very big in size. We have PCA(principal component analysis) because after rgb to gray we will perform multiple filtering of images using different filters like average filter and median filtering.

## 2.3 Filtering methods

After converting the image from PCA we use different types of filters like median, gaussian, contrast enhancement using CLAHE algorithm then background noise is removed by using background exclusion. After applying an average filter we then take the difference between the gray and average filter.

## 2.4 Isodata segmentation method(blood vessel segmentation)

The isodata segmentation is based on multispectral image processing where a number of iterations is run based on the threshold which is calculated for each pixel of image. The isodata segmentation is based on the physiological references based on the different channels of the image. In the first stage we are calculating the mean intensity of the image and using that for the threshold of the pixels.

After that the mean above threshold and mean below threshold is calculated because we want a complete noise removal from all the images in datasets.

In the second stage we are calculating the normalisation for the range in [i,1].

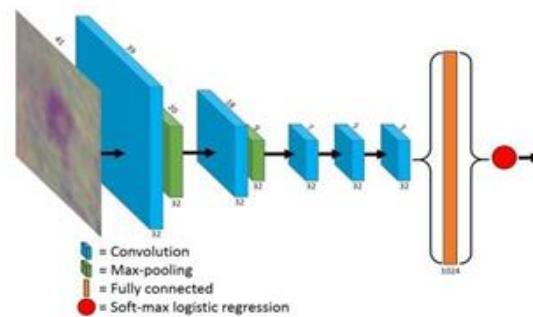
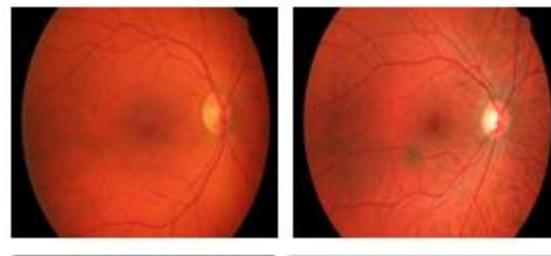
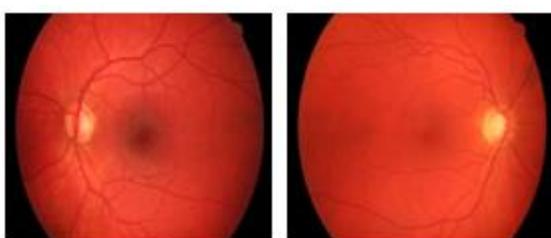


Fig. 2 . AlexNet Layers Used for Training and Detection

Convolution layer has a large number of filters which is used for filtering and it is characterized by the height and width and we have used 2d filtering for this method.

For the 2-d filtering we have not taken more than 20 filters of each having a size of 5x5 dimensions.

The Max-pooling is characterized by the output, indices and size. This gives the maximum value of the output for each layer. It can be both scalar and vector. In this each layer is having size of 2x2 dimensions, stride of size 2x2 and padding vector is included in it.

Rectified Linear Unit layer where the layer is used for pooling and unpooling the contents of the layers. This layer gives us a minima and maxima of each and every layer. We have used the sigmoid function(sgdm) and we have also included the initial learning rate as 0.1-0.2 based on the system specifications and for this we have used epochs between 100-1000.

Fully Connected Layer determines the size of the output layer which multiplies the input by a matrix and then adds a vector.

Softmax layer is used for the classification of the blood vessels and tells if a person has cataract or not.

**Mini Batch-** The total no. of training is large as the dataset is very large hence, we have divided it into small no. of images to reduce the time of training and for determining the output at low no. of epochs.

**Base learning rate-** The learning rate is the no. of times the layers are trained at different no. of epochs. The weights are updated regularly to determine the output fast.

Gradient descent- It is an iterative algorithm that computes the gradient of a function and uses it to update the parameters of the function to find maximum and minimum value of the function.

### 3 RESULTS

From the blood vessel segmentation algorithm we derive that after applying the different filters like Average filter and Gray filter, it is obtained that the threshold value is constant which reduces the number of iterations and this is true for all filter values.

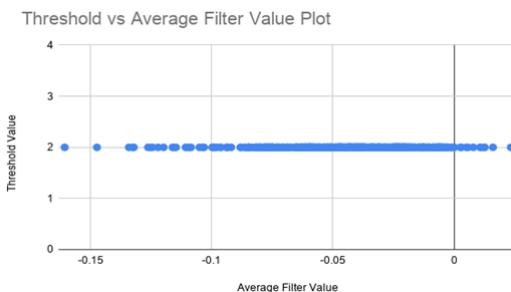


Fig. 3 . Constant threshold with varying Average Filter Value

AlexNet has 22 layers but for our system, only five layers are being used with varying parameters such as regression based models where we use statistical parameters. This reduces the computational time by a considerable amount and reduces the processing power requirements as opposed to using GPUs.

Layer	Operation	Input size	Details
1	Image Input Layer	605X700X3	3x3x3, Kernel = 31
2	Convolution2dLayer	5x20	2x2, stride = 2
3	ReluLayer	20x20	3x3x32, Kernel = 31
4	MaxPooling2dLayer	18x18	2x2, stride = 2
5	Convolution2dLayer	5X20	3x3x32, Kernel = 31
6	ReluLayer	7x7	3x3x32, Kernel = 31
7	MaxPooling2dLayer	5x5	3x3x32, Kernel = 31
8	Soft-max	3x3	1020 nodes
9	Classification Layer	1024x1	2 classes

Fig. 4. Learning and Training Rates of various AlexNet layers

```

AlternateFileSystemRoots: {}
ReadSize: 1
ReadFcn: @readDatastoreImage

layers =
1 '' Image Input
2 '' Convolution
3 '' ReLU
4 '' Max Pooling
5 '' Convolution
6 '' ReLU
7 '' Max Pooling
8 '' Fully Connected
9 '' Softmax
10 '' Classification Output
Training on single CPU.
Initializing image normalization.
=====
| Epoch | Iteration | Time Elapsed | Mini-batch | Mini-batch | Base Learning |
|       |           | (hh:mm:ss) | Accuracy | Loss | Rate |
=====
| 1 | 1 | 00:00:13 | 33.33% | 0.8636 | 0.0100 |

```

Fig. 5. Training Accuracy with 15 Epochs with learning rate of 0.01

```

10 '' Classification Output - crossentropyex
Training on single CPU.
Initializing image normalization.
=====
| Epoch | Iteration | Time Elapsed | Mini-batch | Mini-batch | Base Learning |
|       |           | (hh:mm:ss) | Accuracy | Loss | Rate |
=====
| 1 | 1 | 00:00:07 | 50.00% | 0.7557 | 0.0100 |
| 15 | 15 | 00:01:37 | 50.00% | 0.6932 | 0.0100 |
=====

convnet =
SeriesNetwork with properties:
  Layers: [10×1 nnnet.cnn.layer.Layer]

```

Fig. 6. Classification Results for 15 epochs with learning rate of 0.01

From the given figures (Fig. 5 and Fig. 6), it can be observed that the accuracy increases with the number of iterations and it is possible to get the output with base learning rate of 0.01 as opposed to 0.1. This reflects, the capability of the system that the number of epochs used for training does not affect the number of kernels denoted by k and the novelty in the proposed method is the reduced number of epochs between 15 - 60. It is possible to get the output and the system is capable enough to classify the image as healthy or unhealthy.

### 4. CONCLUSION

After evaluating the blood vessel segmentation of fundus images of various subjects and using it for training the AlexNet we came across the fact that if we use it for different system where the processors are of different generation and this system is compatible with all kinds of all units of epochs and the lowest epochs is 15 and maximum number of epochs is high as possible. The laptop used for this is having i3 generation and 2GB RAM with 1 TB SSD. The training of alexnet in limited number of epochs and with very high efficiency makes the system very unique for detection of all kind of eye diseases including glaucoma, diabetic retinopathy, genetic and genial diseases.

### 5. REFERENCES

- [1] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," 2015
- [2] J Y. Guo, G. Wu, L. A. Commander, S. Szary, V. Jewells, W. Lin, and D. Shent, "Segmenting hippocampus from infant brains by sparse patch matching with deep-learned features," 2016

- [3] M. J. J. P. van Grinsven, Y. T. E. Lechanteur, J. P. H. van de Ven, B. van Ginneken, C. B. Hoyng, T. Theelen Auto-’ matic drusen quantification and risk assessment of age-related macular degeneration on color fundus images,” 2013
- [4] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” 2015
- [5] Li, J.Q.;Welchowski, T.; Schmid, M.; Letow, J.;Wolpers, A.C.; Holz, F.G.; Finger, R.P. Retinal Diseases in Europe Bonn, Germany, 2017.
- [6] Li, J.Q.;Welchowski, T.; Schmid, M “Advantages and examples of resampling for CAD evaluation,” in IEEE International Symposium on Biomedical Imaging, Germany 2017
- [7] J. N. Cheung, P. Mitchell, and T. Y. Wong, “Diabetic retinopathy,” 2010.
- [8] G. B. Kande, T. S. Savithri, and P. V. Subbaiah, “Automatic detection of microaneurysms and hemorrhages in digital fundus images,” Journal of Digital Imaging, 2010
- [9] H. Seyedarabi, and A. Javadzadeh, “A comparative study on preprocessing techniques in diabetic retinopathy retinal images: illumination correction and contrast enhancement, 2015
- [10] S. Deepa “Certain investigation of the retinal hemorrhage detection in fundus images,” International Journal of Electronics and Communication Engineering 2015
- [11] R. K. Ghosh, and P. Gupta, “An efficient parallel algorithm for random sampling,” 1989
- [12] M. Gambardella, and J. Schmidhuber, “Mitosis detection in breast cancer histology images with deep neural networks. 2013
- [13] “Kaggle diabetic retinopathy detection competition report,” University of Warwick 2015
- [14] R. Mann, N. Karsse-’ meijer, and B. Platel, “Automated localization of breast cancer in DCEMRI,” 2015
- [15] A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” Journal of Machine Learning Research, 2014