# Super Learner Model to Detect Abnormalities - OCT and blood smear imaging case studies

**Dipam Paul[1,2], Alankrita Tewari[1], Imon Banerjee[2,3]**
[1]Department of Electronics Engineering, KIIT University, Bhubaneswar, India
[2]Department of Biomedical Informatics, Emory School of Medicine, Atlanta, GA, USA
[3]Department of Radiology, Emory School of Medicine, Atlanta, GA, USA

**Abstract**

We propose a Super-Learner framework by combining Deep Ensemble Network and feature maps to automate the differential diagnosis from medical imaging datasets. The Super Learner achieved state-of-the-art performance on both imaging datasets and mitigated the chances of over-fitting.

## Introduction

Since the past two decades, there have been numerous studies on combining and blending Machine Learning models to achieve optimal performance. On the contrary, there have been hardly any investigations conducted to effectively ensemble Deep Neural Networks. The authors report that ensembling Neural Networks for medical diagnosis can remarkably help to overcome single network limitation by decreasing variance and bias and increasing accuracy, and escalate the application of automated models in clinical practice. Moreover, expert systems for medical diagnosis essentially can tolerate false positive (which could be reviewed by the experts later) but on the cost of minimal false negatives.

Ensemble learning is a well-recognized method in machine learning that effectively obtains exemplary results by combining multiple classifiers. Early work by Hansen and Salamon (Hansen and Salamon 1990) proves that the efficacy of a neural network model can be refined, by ensembling multiple neural networks. Since this technique behaves reasonably well, it is considered to be an effectual approach for the fashioning of sophisticated image classification systems and thus has already found applications in diversified areas such as facial recognition, optical character or pattern recognition, scientific image analysis, medical diagnosis systems, signal classification, etc.

In the present study, we have developed an ensemble model *'Super Learner'* that combines various Deep Convolutional Neural Networks (DNN) based on their compatibility. The technique was further elucidated using two distinct case studies that address two medical problem statements. Firstly, the technique was used to construct a model named *OCTx* to classify four retinal disorders from Optical Coherence Tomography (OCT) images. Secondly, it was used for *MalariaX* which diagnoses parasitized Malaria by performing a binary classification of thin blood smear images that were classified as either 'Parasitized' (infected) or 'Uninfected'. In both cases, the Super Learner yielded state-of-the-art accuracy, increased model reliability and mitigated the chances of over-fitting for the test cases.



## Methodology

Super Learner architecture comprises two primary segments. As depicted in Figure 1, the first layer utilizes the pre-trained architectures and the second layer combines the prediction of individual architectures by pulling the feature maps via a non-linear ensemble network. Thus potentially maximising the optimisation reach one could have in this dataset. It is also important to take note that there exists no mathematical relationship between the number of DNNs to be used on a particular dataset and which ones to be used - therefore, it is purely empirical and based on the ablation studies conducted.

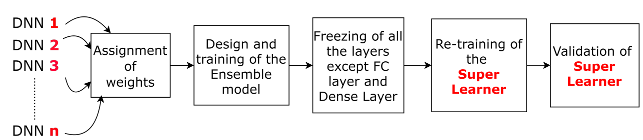


Figure 1: Super Learner ensemble approach.

### Datasets - OCT images and blood smear imaging

The dataset for OCTx has been gathered from the open-source Mendeley repository (Kermany et al. 2018). The images of the dataset have been extracted from random subjects, all gathered by professionals. The dataset is organized into four diagnosis categories, namely Normal, CNV, DME, and DRUSEN. There are 84,484 OCT images and the total distribution of images are - Train (83,484 images), Test (968 images), and Validation (32 images) while the dataset for MalariaX (Malaria Cellular Images dataset (Rajaraman et al. 2018)) has been originally collected by the National Library of Medicine, Bethesda, Maryland. It contains blood smear slide images of healthy and malaria-affected subjects. A collection of Giesma-stained blood was done from 150 subjects, infected with P. falciparum and 50 healthy patients.

The segmented cells are categorized into 'Parasitized' and 'Uninfected' classes. The authors report that this dataset served as a benchmark and comprises of 27,558 images in total - 80% of which were used for Training and the remaining for Testing our Super Learner.

Given the uneven distribution in the classes of the images in both the datasets, in the train set, we re-balanced the minority class using over-sampling techniques. Parameters like rotation_range, zoom_range, horizontal_flip, vertical_flip, fill_mode, etc are randomly modified therefore adding more images to training data and increasing the model accuracy. Then the images were resized up to a fixed dimension of 256 x 256 and 224 x 224 for the two datasets respectively.

## Model architecture

The *Super Learner* is an ensemble approach, constructed using n-Deep Neural Network architectures - (1) Custom CNN architecture (5 hidden layers [conv2d, Pooling, Dropout], and a series of FC layers, followed by an output layer) and (2) Pre-trained well-known architectures, such as DenseNet, VGG16, and InceptionV3, ResNet50. For each case-studies, the specific set of DNN architectures were selected from a set of open-source pre-trained models based on their performance on gathering trainable access points from the image data, when exposed to high-quality images. In addition, they were also selected after evaluating standard performance metrics.

The breakdown of the architecture and process flow of Super Learner is depicted in Fig. 1. We categorize all the results of the first level of classification and collate the predicted outputs obtained from the individual models on all of the test images. Then, we assign non-linear weights based on the constraining of how well the models performed. Unlike a stacking ensemble architecture, these non-linear weights in our proposed model are assigned following an averaging mechanism to ensure adaptive weightage assignment between the classes - irrespective of the dataset used to evaluate. Thereafter, we design an ensemble model after this aforesaid transformation of the data and design and train the Neural Network layers by initially freezing all the layers except the Fully Connected (FC) layer and the Dense Layer. Feature maps (see Fig. 2) were obtained for gaining insights about the edge detection performance of our model and then we feed it to the same and train the ensemble model again after attempting to further fine-tune the hyper-parameters involved which would govern the final classification to take place.

## Results

Table I and II delineates the individual model as well as the Super learner performance on the respective test datasets - 968 OCT images and 3,858 blood smear images. As observed, the Super learner outperformed all the individual models for both multinomial (OCT) and binary (blood smear images) with a statistically significant margin. For the multinomial OCT image classification study, our Super Learner was able to classify 954 images against their correct classes out of 968 images in the test set of the OCT dataset and the accuracy was obtained to be 98.55% and the mis-classification rate was correspondingly 1.45% among the four categories as highlighted in Fig. 3 (a).

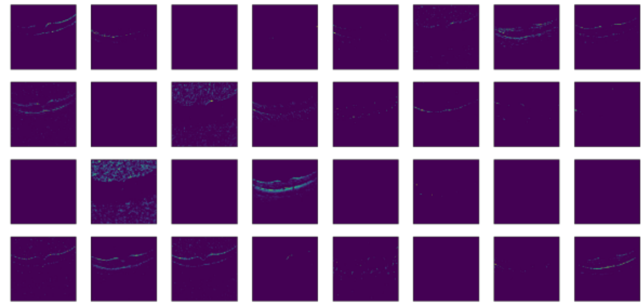

Figure 2: Obtained feature maps: OCT images.

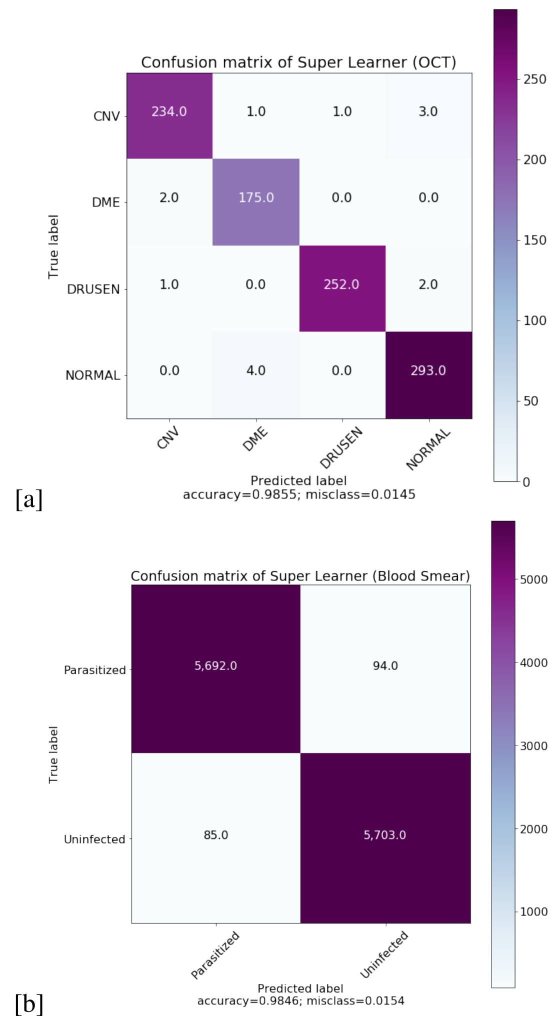


Figure 3: (a) Confusion Matrix - Super Learner (OCT) (b) Confusion Matrix - Super Learner (Blood Smear)

While for the binary Malaria Cellular Images classification, our Super Learner was able to predict 3,799 images

| Metrics Used | Custom CNN | DenseNet | VGG16 | InceptionV3 | **Super Learner** |
|---|---|---|---|---|---|
| Accuracy | 0.940 | 0.963 | 0.958 | 0.955 | **0.985** |
| Precision | 0.924 | 0.938 | 0.944 | 0.965 | **0.968** |
| Recall | 0.901 | 0.941 | 0.952 | 0.969 | **0.973** |
| F1-Score | 0.912 | 0.939 | 0.948 | 0.967 | **0.970** |

Table 1: Performance Evaluation of the Custom and Individual Pre-Trained Architectures on OCT images

| Metrics Used | DenseNet | Custom CNN | ResNet50 | **Super Learner** |
|---|---|---|---|---|
| Accuracy | 0.962 | 0.975 | 0.973 | **0.985** |
| Precision | 0.954 | 0.952 | 0.978 | **0.984** |
| Recall | 0.965 | 0.989 | 0.964 | **0.985** |
| F1-Score | 0.953 | 0.976 | 0.977 | **0.986** |

Table 2: Performance Evaluation of the Custom and Individual Pre-Trained Architectures on the blood smear images

correctly, out of 3,858 test images. The test accuracy is calculated as 98.46%, with an error rate of 1.53% as is shown in Fig. 3 (b). We present the confusion matrix for both tests in Fig 3.

## Conclusion

In this paper, we demonstrate a novel framework based on ensembling multiple Deep Neural Networks which were selected based on their individual performances on the targeted case-study. We were successfully able to access the performance of the framework on two distinct imaging case-studies with a state of the art accuracy. We further welcome research in the domain of utilisation of Deep Ensemble Networks used in this paper in the detection of diseases and disorders from medical images and thereafter achieving better-calibrated results on future image datasets as well.

## References


Hansen, L. K.; and Salamon, P. 1990. Neural network ensembles. *IEEE transactions on pattern analysis and machine intelligence* 12(10): 993–1001.

Kermany, D. S.; Goldbaum, M.; Cai, W.; Valentim, C. C.; Liang, H.; Baxter, S. L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. 2018. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 172(5): 1122–1131.

Rajaraman, S.; Antani, S. K.; Poostchi, M.; Silamut, K.; Hossain, M. A.; Maude, R. J.; Jaeger, S.; and Thoma, G. R. 2018. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *PeerJ* 6: e4568.