Ensemble Model for X-ray Image Classification

MSc Research Project

Arun Bennur
Student ID: X17139813

School of Computing
National College of Ireland

Supervisor: Anu Sahani

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name :	Arun Bennur
Student ID:	X17139813
Programme:	Data Analytics
Year:	2019
Module:	MSc Research Project
Lecturer:	Anu Sahni
Submission Due Date	12-08-2019
Project Title	Ensemble Model for X-ray Image Classification
World Count	4512

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ………

Date: ………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Ensemble Model for X-ray Image Classification

Arun Bennur
X17139813

Abstract

Medical image Classification is powered by Deep learning solutions. The Research work proposes a novel methodology for the medical image classification and aims to introduce a reliable model for medical diagnoses. The research work is carried out on Publicly available Dataset chest X-ray images dataset, classifying images into Pneumonia and Normal. The approach uses the state of art image classifiers MobileNet, Inception-V3 and Xception and harness the power of transfer learning and Data Augmentation in image classification problem and finally the ensemble model is produced, based upon the average of predictions of all the three image classifiers. The accuracy obtained from this method far exceeds the result from the previous studies. All the models used, are evaluated and comparative study of all the image classifiers will be carried out. The best image classifier is identified. The limitations and Future scope of this research methodology is presented.

Keywords: Medical Imaging, CNN, Transfer learning, Ensemble, Data Augmentation

1 Introduction

Medical imaging market was approximately 34 billion in 2018 and expected to grow 48.6 billion by 2025. There is a huge investment by governments in developing countries and the advancement in technology which fuel the growth further. The widespread adoption of electronic medical records contributes to the exponential generation of biomedical data in size, dimension and complexity(Livieris *et al.*, 2018). The medical imaging uses number of technologies like MRI, X-ray CT-scan, etc., to diagnose the disease. Deep learning and Artificial intelligence can help radiologists to achieve diagnostic excellency. CNN-motivated deep learning algorithms have become the standard choice for medical image classifications state-of-the-art CNN-based classification deep learning solutions has the potential just not to aid human experts in image diagnosis but also has the capability to replace human experts. Replacing the practising radiologists depends on the reliability of the model. With the advancement of new image classification algorithms and the techniques, the accuracy at which the image can be classified is increasing. The search for the better architecture for medical image classification is never ending. It is also noted that the data available in the local clinic is limited. There is always a need to develop a sustainable model which performs excellently at low data environment and is reliable. There is a need to develop an app or portable device which can diagnose the disease by just scanning the picture with an accuracy of expert radiologist the research work aims to use the state of art architectures in the image classification problem. The experiment make use of pretrained convolutional Neural network like MobileNet, Inception Xception and leveraging the power of Transfer learning and Data augmentation for pneumonia classification. In this research work novel approach of ensemble of all the above classifiers Mentioned is carried out. The motivation of the research work comes from the availability of the labelled dataset, Chest X-ray images. The Dataset is labelled into two categories pneumonia/Normal and is organized into test train and validation. The dataset is graded by expert physician and evaluated by third expert. In this research work However classification of pneumonia from X-ray images is a challenging task due to less availability of data for training the model.

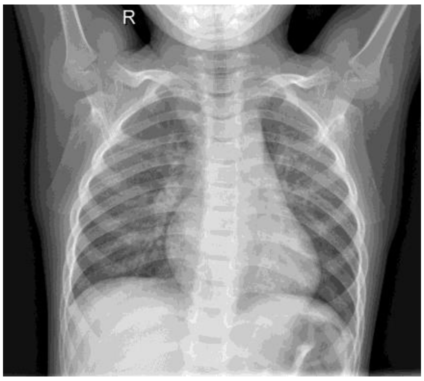


Fig 1 : X-ray Normal

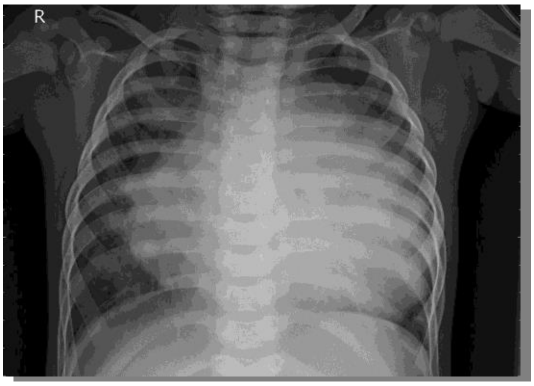


Fig 2: X-ray Bacterial Pneumonia

The report study is organised in to following sections. Section 1 introduces the need and gives you the overview of the experiment. Section 2 covers the related works done in this domain. Section 3 gives the overall methodology followed in this experiment. Section 4 gives the design specifications of the architectures used. Section 5 discusses the implementation of the project. Section 6 evaluations of all the methods are discussed and ends with conclusion discussing the direction of the scope and future work of the research.

2 Related Work

An outstanding amount of work is already done in the field of image classification. The related work section is grouped into 2 major sections. The first sections explore the studies conducted in the architectures used which has led in the selection of classifiers and parameters used in conducting the research. The second section focuses on the previous studies on the subject of pneumonia classification from x-ray images.

2.1 Literature review of image Classifiers

In 2017, The author (Xia, Xu and Nan, 2017) conducted a research work which aimed at classifying images of flowers using inception v-3 model of tensor flow platform for further improving the accuracy he used transfer learning technique. He used two datasets for the experiment. Oxford-17, Contained 17 species of flowers and each species containing 80 flower images. The flowers in the dataset had a similarity with the different species and had a difference between the same species. Oxford 102, the dataset contained 102 species with each species containing a range from 40-258. The results of the experiment achieved higher training accuracy up to 100% with the validation accuracy of 99% for Oxford-17 Flower Dataset. The same model performed with an accuracy of 100% training accuracy and 95% validation accuracy. Overall accuracy on both the dataset being 95% and 94% respectively

The author (Avery *et al.*, 2014) interprets the inception model in CNN. He states that the inception module is an intermediate step between the regular convolution and the depth wise convolution, and he proposes the new state of art architecture and refer it as Xception. The model Xception is just an extension of the module where inception module has been replaced by depthwise separable convolutions. He compares the classification performance of the two models using the dataset ImageNet dataset. The model Xception significantly scores better than the model Inception (Inception v3) on a larger dataset containing 350 million images with 17,000 classes. He concludes stating that the performance increase in Xception is due to the effective utilization of model parameters.

The author (Gavai *et al.*, 2018)carried out the experiment of classifying flowers using MobileNets on TensorFlow platform. The dataset used in this experiment is Oxford102. The images are used multiple times during training. Training accuracy reached during the research work between the value 85-99%. The Training precision is high compared to the validation accuracy resulting in overfitting of the model. The results of the experiment is compared with other classifiers namely VGG16, Squeezenet, AlexNet. The accuracy outperforms every other model mentioned above and the only model which stands close with the accuracy is VGG 16. The advantage of using MobileNets over VGG 16 is that the model is 32 times less compute intensive. The research work emphasis on width and resolution multiplier over a rational amount of correctness to shrink size and latency to increase efficiency.

The author (Kieffer *et al.*, 2018) carried out a research work on classification of medical images dataset based on a feature vector. The feature has been extracted from the deepest layer of pre-trained CNN. The author evaluated the impact of transfer learning with very less samples, and the performance of pre-trained deep features versus CNNs. The dataset, kimia path24 consists of 27044 histopathology training patches in 24 tissue texture classes. The results of the experiment show the pre-trained networks performs better than the networks which are trained from scratch. The experiment also sheds a light on additional tuning of VGG16 did not result any significant performance, but the inception model's accuracy was improved after fine tuning.

The Research work shows the importance of Data Augmentation in image classification problem. The research work was conducted on a very small set of ImageNet dataset in the classification of cats and dogs by an author (Perez and Wang, 2017). For this experiment 500 images for cats and dogs is used Among them 400 is taken for training and the rest 100 is used in validation set. The Adam optimizer is used with a learning rate of 0.001 and the model is trained for 40 epochs. Various Augmentation methods and its effectiveness are tabulated. Augmentation in the above method contributes in achieving better accuracy. Author concludes stating data augmentation helps avoid over fitting during training and regularization problems upfront when used with neural networks.

The author(Esteva *et al.,* 2017) has conducted a research work of classifying skin cancer with deep neural network. The research is assumed to be challenging due to the reason of very less variability in the appearance of skin lesions. Author trains CNN from a scratch using only pixels and disease labels as inputs. The dataset used contains 129,450 clinical images- which are categorized into 2032 different diseases. The author claims the results are par with dermatologist level of diagnoses. Google’s Inception v3 is used and pretrained to 93.33% top -five accuracy on the 1000 object classes of the 2014 ImageNet Challenge. The sensitivity and specificity is found to be AUC of 0.96 and 0.94 respectively.

The author (Ortiz *et al.*, 2016) conducts a research study for the early diagnosis of Alzheimer’s disease. The study focuses on classification of brain regions by Automated Anatomical Labelling. For this experiment the images of grey matter pertaining to specific brain regions has been split into 3D patches about AAL atlas. Each patch is used to train different deep belief networks. The final predictions are made by ensemble of different deep belief networks by voting scheme. In this experiment four voting schemes and two deep learning-based structures are compared. The above method is also efficient in classifying Mild Cognitive Impairment (MCI) subjects. The experiment classifies AD images with accuracy up to 90% and forming AUC of 0.95 for NC/AD classification.

The author (Gargeya and Leng, 2017)designed a new algorithm based on deep learning for automatic detection of diabetic retina. The research work is carried out on fundus images obtained from EyePACS dataset. The validation for the above research work is also carried out using dataset MESSIDOR and E-Optha Dataset. Customized deep convolutional neural network uses convolutional parameters to learn iteratively filters which transforms images into ranked feature maps. The layer learns discriminative features with varying spatial levels. The precison-recall tradeoff of the method Is evaluated by plotting receiver operating curve. The method achieves 0.97 AUC with a 98% and 94% specificity and sensitivity respectively with the local dataset. Testing the method with other two dataset MESSIDOR and E-Optha) resulted in 0.94 and 0.95 AUC score. The author claims the model is found to be reliable for opthalmoglogists.

The author (Shin *et al.,* 2016) conducted a research work and sheds light on understudied factors of implementing deep CNN to aid computer aided diagnoses problems. Authors compares different CNN architectures which contains 5 thousand and 160 million parameters with varying number of layers. Experiment is conducted on diagnoses of Thoraco-abdominal lymph node by using transfer learning. The research work achieves the sensitivity of 85% at 3 false positive per patient. The cross -validation classification predicts axial CT slices with ILD categories. The author claims the insights gained from the research work can be used to design high performance computer aided diagnoses for medical imaging tasks.

The Author ((Khalifa *et al.*, 2017)) presents a galaxy classification base on its features Elliptical spiral and irregular. The convolution network author uses has 8 layers with 96 filters and two fully connected networks at the end. The model is trained on over 1350 images. The testing accuracy is remarkable with 97.272%

2.2 Literature review of X-ray image classification models

The author (A.A Saraiva, et.al; 2019) conducted a research work on image classification of chest X-ray images. He used Convolutional layer in the first stage which extracts certain characteristics of images using several filters followed by the pooling layer which is used to reduce the spatial size the both layers are linked with Fully Connected layers (FCN). FCN converts the weights received from previous layer to new resource vector after the weight is multiplied to form a computation matrix forming a dense matrix vector multiplication. Model is evaluated using K-fold cross validation. The performance of the model exceeds in terms of accuracy of 95.30% in the tests against the work of (Kermany *et al.*, 2018).

The author (Kermany *et al.*, 2018) addressed the problem of interpretability and reliability of medical imaging algorithms. He used the power of deep-learning frameworks and screened on retinal diseases patients. The power of transfer learning is applied on the dataset of optical coherence tomography images. His work forms the basis of utilizing Transfer learning for medical image diagnostics. The model was able to distinguish images of choroidal neovascularization and diabetic macular edema with accuracy of 96.6% and with a sensitivity of 97.8%. Occlusion test was performed on 491 images to identify the most significant areas predicted by neural network. Deep learning algorithms identified the region of interest and predicted with accuracy of 94.7%. The author also conducted an experiment on Chest X-ray Pneumonia dataset. The model performed exceptionally well classifying images with an accuracy of 92.3% despite the challenge of presence of large number of irrelevant features.

In the year 2017 author (Rajpurkar *et al.*, 2017) claim to develop an algorithm which can classify pneumonia images from chest X-rays with an accuracy exceeding practising radiologists. The Algorithm is named as ChexNet. The Algorithm is trained on dataset of 100000 frontal view Chest X-ray images with 14 diseases. The Algorithm is based on the Convolutional Neural Network with 121 layers. The algorithm also has the capability to produce the heatmaps for visualization indicating the disease using activation mappings. To generate the feature, map the image is processed into fully trained network where feature maps are captured from the output of final convolutional layer. The algorithm CheXnet classifies Cardiomegaly with an accuracy of 92.48% and Pneumonia with 76%. It is found that the model has a margin of >0.05 AUROC compared to other state of art models.

In the year 2018 Author (Baltruschat *et al.*, 2019) aims to explore the capability of architecture ResNet-50 by applying on chest X-ray classification. He also considers training the model from the scratch by using ChestX-ray14 dataset. The advantage of availability of high spatial resolution of X-ray data is leveraged with the combination of non-image data of patients by using ResNet-50. The different approaches for the classification of pathology is compared by using ROC statistics. For the network which is trained from the scratch it achieves the AUC of 0.730 after fine tuning it increases to 0.819. author also states that the AUC increases after associating with non-image features. The ResNet-50 architecture achieves 99.3% and 99.1% is sensitivity and specificity respectively.

The author (Vianna, 2018) conducts an experiment on x-rays chest images dataset and makes a comparative study of the powerful architectures AlexNet, SqueezeNet, ResNet and inception in classifying images into normal and pneumonia. He uses transfer learning to train the models with limited data and using data augmentation techniques for avoiding over fitting problems in the models. He credits the data augmentation technique helped to avoid overfitting problems by not allowing training accuracy to reach 100%. The model ResNet is has the better accuracy of 96.3% when compared to other models. The author has not mentioned the other metrices like specificity, sensitivity to determine the realiablity of the model.

3 Research Methodology

The above research methodology is based upon the Cross-industry process for data mining (CRISP-DM). The research work deals with the image classification of pneumonia/normal from chest X-ray images. The below figure indicates the methodology followed in research work

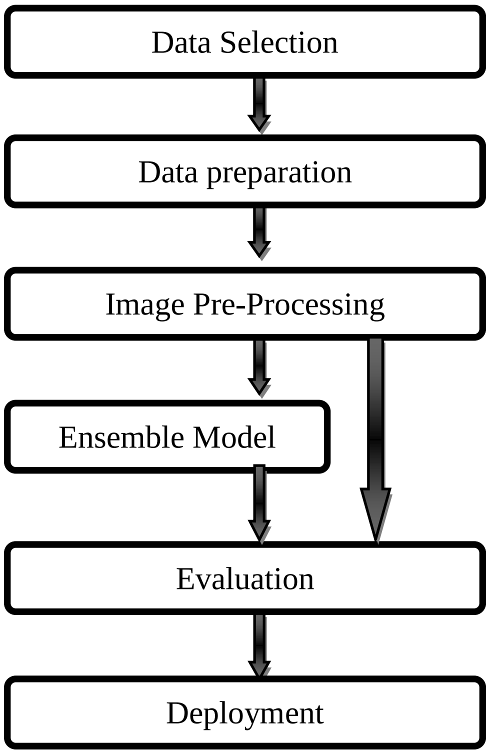
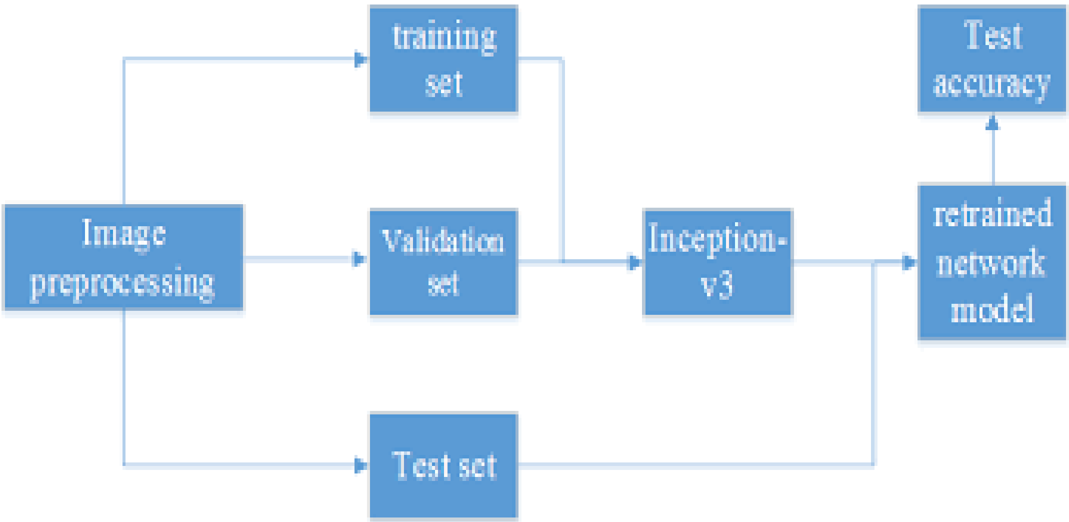


Figure 3: The steps involved in research Methodology

The acquired data consists of labelled 5863 X-ray images categorised into Pneumonia/Normal. The dataset is split for Training Testing and validating the model. Since, the dataset required to train the model is limited, 88% of the images is kept for training the model. 10% of the images are used for Testing and the small portion of dataset is used for validation. The next stage involves the data preparation. In this stage the dataset images which are available in JPEG format are not available in uniform resolution. This serves as an advantage since the data is augmented further. Data Augmentation process involves Resizing, And Squashing of the images which could have taken many lines of code if it were available in the uniform resolution. Data Augmentation helps avoiding the overfitting of the model.

Figure 4: Steps involved in Image Pre-processing



The augmented data is processed with pre-trained CNN architecture with MobileNet, Inception-v3 and Xception. Fine tuning is done for all the above architectures. Each architecture is added with an extra layer with two neurons connected to the last output layer and trained with the X-ray image dataset. Each Model is evaluated for training, testing and validation accuracy. The weights of each models are saved. The saved weights are used as input for ensemble model. The above three models are then ensembled and evaluated.

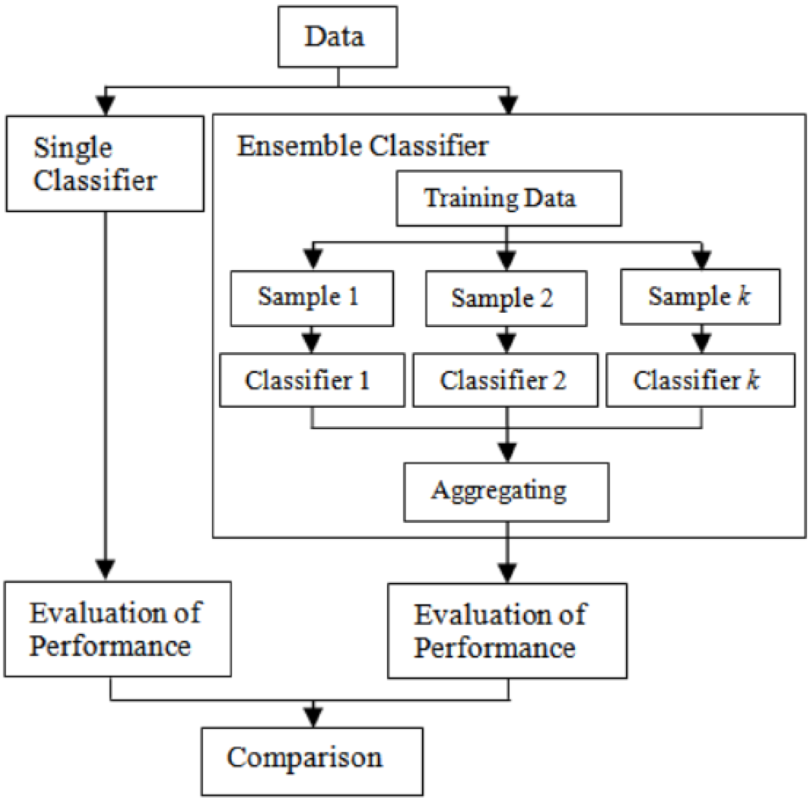


Figure 5: Workflow of single vs Ensemble Classifier (*Utami, et al., 2014)*

The method employed for ensemble is based on simple linear probabilities obtained from each model. For evaluation, Confusion matrix and Receiver Operated Characteristics (ROC) which explains the selectivity of the image classifiers are created. the metrices Specificity, Sensitivity, Precision The negative prediction values for each model are recorded and compared. The Comparison for the each metrices of the model is later discussed in the evaluation section.

4 Design Specification

4.1 MobileNet

One of the pretrained Convolution network is used for the experiment is MobileNet. MobileNet, being light weight require low maintenance and performs quite well with high speed. The single filter to each of the input channel is applied by depth wise convolution and then 1 x 1 convolution is applied by the pointwise convolution to put together the outputs. The filtering and combining of the inputs into a new set of outputs happens in single step. Then the outputs ae split further into two layers by a standard convolution. The split output of a separate layer for filtering and combining is formed. This factorization is responsible for the drastic reduction in computation and the size of the framework. The MobileNet totally has 30 layers with convolution layer with stride 2, depthwise layer, pointwise layer that doubles the number of channels etc. The last layer of the Mobilenet is discarded and is replaced by two denser layers. We have two neurons connected to the final layer with softmax activation.

4.2 Inception V3

The Inception-v3 Convolutional Neural Network is used as the second classifier for the research work. The Inception V3 has 48 layers and is trained with more than a million images from the dataset. Inception V3 module has pooling layers and convolutional filters and SoftMax activation function. Two Dense layers are added at the end.The finetuning parameters are same as the previous architecture MobileNet.

4.3 Xception

The Xception architecture is just an extension of inception. Where the depthwise convolution layers are replaced in the inception module. The replaced layers have the following operations 1. Spatial convolutional, being n x n convolution and 2. Pointwise convolution which is actually 1 x 1 convolution which is used to change the dimension 3. Batch Normalization. Xception architecture also has dense layers as output layer
With Global Average Pooling and SoftMax as the activation function.

4.4 Ensemble

The weights of each image classifiers are saved and are ensembled. In this method, the average of predictions of all the three image classifiers are taken to make the final prediction

5 Implementation

Research work is carried out using python language using Keras library which uses Tensorflow as the backend. The data is transformed from 3D to 4D tensor with shape 1,120,120,3. The augmented data is processed through the pretrained image classifiers MobileNet, Inception and Xception. Final layer of all the three image classifiers were fine tuned adding two dense layers with Global Average Pooling and softmax as the active function. Adam optimizer with learning rate of 0.001 is used which has the fastest convergence and the lower loss during the training. The images are processed in the batch of 8 and the loss function being categorical crossentropy. All image classifiers are trained for 4 epochs. The minimum time taken to train is MobileNet which trains with accuracy of 95.58% with validation accuracy of 95%. The Inception model trains with an accuracy of 85% with the validation accuracy being 90.04% and the validation loss did not improve from 0.19519. The best among the three being Xception with training accuracy of 88.08% and the validation accuracy of 90.80%, for evaluating the model the confusion matrix is created for each architecture. Then all the three architectures are ensembled with average prediction method. The accuracy is tested with a sample size of 200 for each model. The best test accuracy achieved is 97% with Ensemble among all 4 models. The performance comparison among all the 4 models with respect to the metrices is done in the Evaluation section.

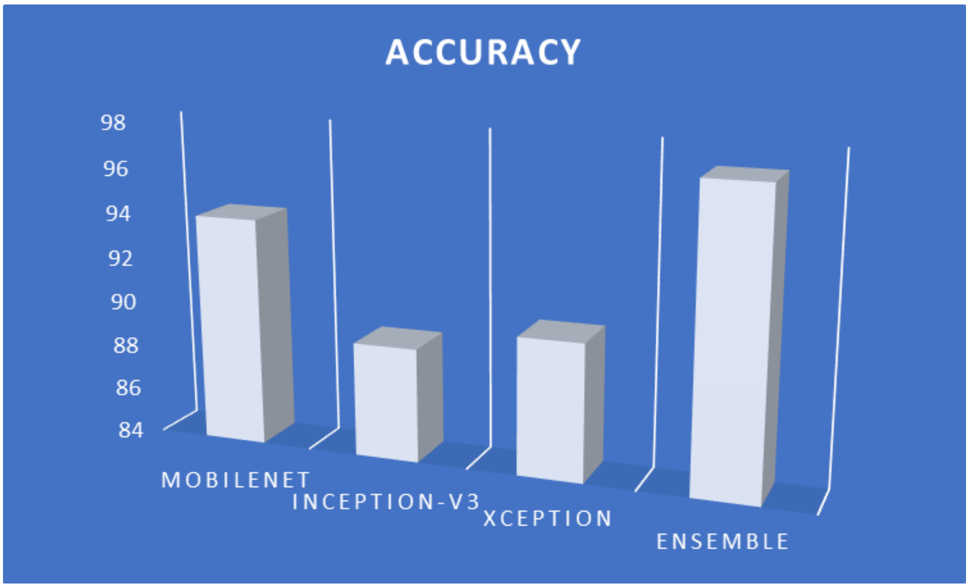
6 Evaluation

Performance Evaluation

In this research work the Dataset was split into 88:10 for training to test ratio. The test dataset is balanced with 234 images being normal 390 images having pneumonia. The accuracy is one of the important parameters in evaluating performance. Since the classification problem is in regard with medical field, not only the accuracy but Sensitivity and Specificity also determines the reliability of the model. The ideal model classifies true positive and true negative efficiently. Here in this section all the metrices like sensitivity specificity precision and negative prediction values are compared.

6.1 Accuracy

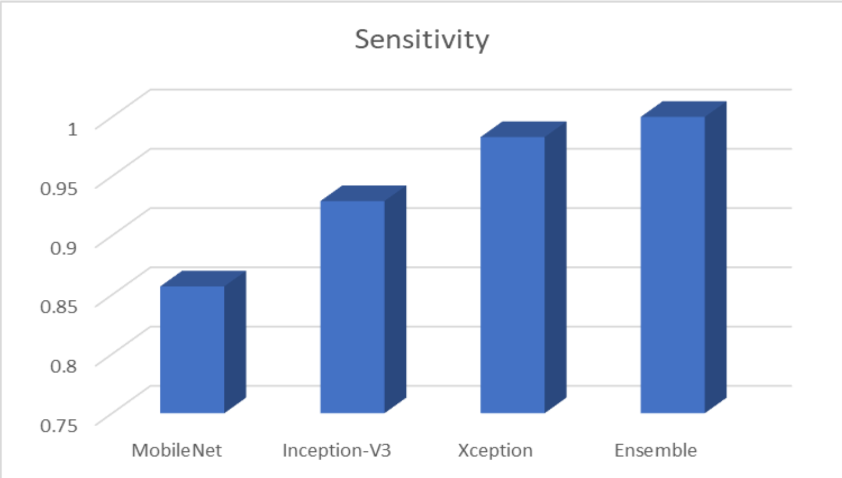
Model	MobileNet	Inception - v3	Xception	Ensemble
Accuracy in %	94	82.14	96.	97.5



Although all the four models performed good. From the above chart it is obvious that the Xception classifier performed well in terms of Accuracy and Ensemble method performed slightly lesser in.

6.2 Sensitivity or Recall or True positive rate

Model	MobileNet	Inception-v3	Xception	Ensemble
Sensitivity	0.857	0.929	0.983	1



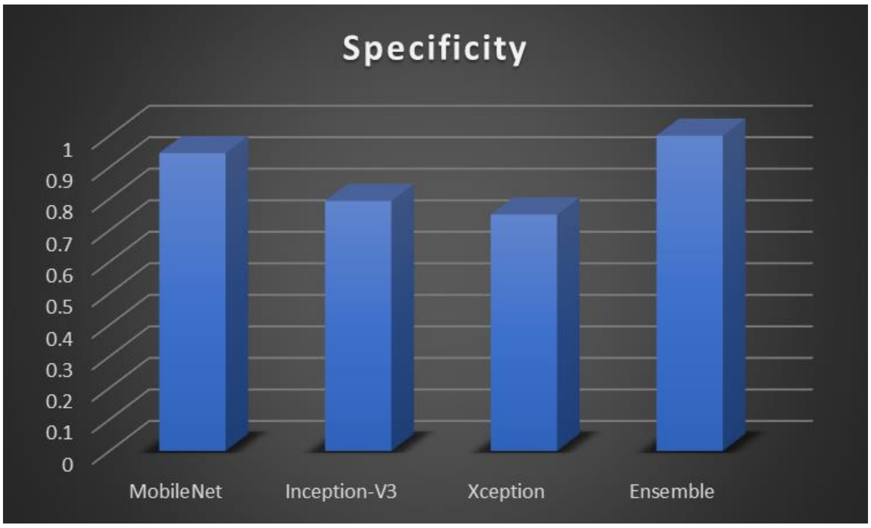
Sensitivity in this research work denotes what proportion of images that had pneumonia being classified as pneumonic. Sensitivity is given by the equation

Sensitivity = TP/(TP + FN)

6.3 Specificity/Selectivity

Specificity -= TN/(TN+FP)

Model	MobileNet	Inception-v3	Xception	Ensemble
Sensitivity	0.945	0.793	0.75	1

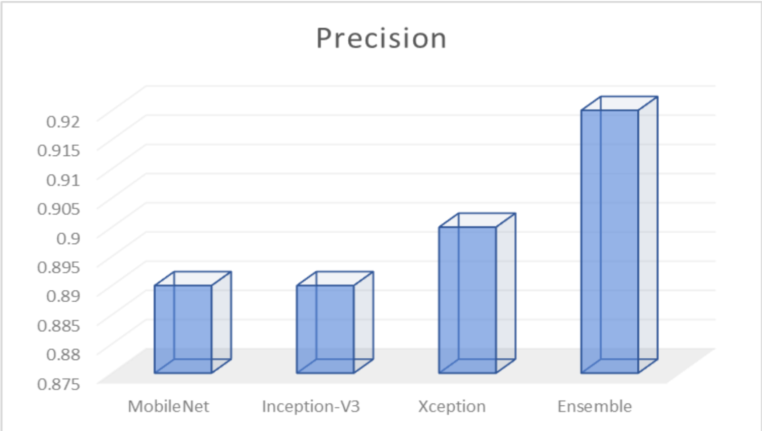


In this research work the specificity is the measure of capability of model that tells us the proportion of images are classified as normal, were predicted by the model as normal Specificity is exactly opposite of Recall. Ensemble model has the specificity of 1

6.4 Precision

Model	MobileNet	Inception-v3	Xception	Ensemble
Sensitivity	0.89	0.89	0,9	0.92

Precision = TP/(TP+FP)



Precision is the measure that tells us what proportion of images that we classified as having pneumonia, actually had Pneumonia. Ensemble model has precision value of 0.92 followed by Xception with 0.905.

From all the above metrices it is evident that the Ensemble architecture performs best in classifying images with higher accuracy and has the value of specificity and sensitivity greater than other models. Ensemble performs well when it comes to precision with precison 0.92

6.5 AUC-ROC curve

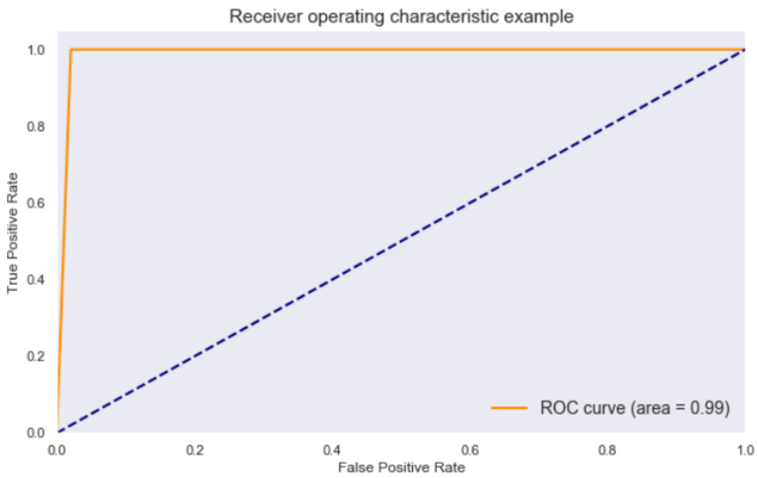


Fig: AUC-ROC for Ensemble Model

From the above AUC-ROC (Area Under the Reciever Operating Characteristics) Curve for the Ensemble model. It is noted that Capability of model for distinguishing between classes is high with area under curve being close to 1. It indicates the Ensemble model is better in classifying the images between pneumonia and Normal.

7 Conclusion and Future Work

The capability and the reliability of the ensemble model for classification of X-Ray images is studied. It is proven that the ensemble model outperforms all the other models in the experiment and from the previous studies. The Model is reliable, and it decreases the type 2 error compared to other models illustrated in the research. The model is highly efficient with an accuracy of 97% and with the precision 0.92. This model is slower compared to a single image classifier.

The future work involves the study of performance of the model with other medical image dataset with random data splits of training and testing data. The research work carried out forms a foundation for further research. There is a huge scope of conducting a research on bringing the right ensemble method and appropriate classifiers together to make a reliable and robust model.

Acknowledgements.

I express my sincere gratitude to my Mentor Prof. Anu Sahani for her unrelenting support and encouragement. Thanks again for her immense support, motivation and guidance on my every phase of my thesis.

Adding to that, I would like to express my gratefulness to my sister, Kiran Bennur and her husband (Bhava) Naveen Sunag. And at last I finally thankful to my mother Smt. K.C. Jayakotramma. Her blessings have helped me to reach this far.

References

Avery, K. *et al.* (2014) *Xception Net, SAE International Journal of Materials and Manufacturing*. doi: 10.4271/2014-01-0975.

Baltruschat, I. M. *et al.* (2019) 'Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification', *Scientific Reports*. Springer US, 9(1), pp. 1–10. doi: 10.1038/s41598-019-42294-8.

Esteva, A. *et al.* (2017) 'Dermatologist-level classification of skin cancer with deep neural networks', *Nature*. Nature Publishing Group, 542(7639), pp. 115–118. doi: 10.1038/nature21056.

Gargeya, R. and Leng, T. (2017) 'Automated Identification of Diabetic Retinopathy Using Deep Learning', *Ophthalmology*. American Academy of Ophthalmology, 124(7), pp. 962–969. doi: 10.1016/j.ophtha.2017.02.008.

Gavai, N. R. *et al.* (2018) 'MobileNets for flower classification using TensorFlow', *2017 International Conference on Big Data, IoT and Data Science, BID 2017*, 2018-Janua, pp. 154–158. doi: 10.1109/BID.2017.8336590.

Kermany, D. S. *et al.* (2018) *Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning, Cell*. Elsevier Inc. doi: 10.1016/j.cell.2018.02.010.

Khalifa, N. E. M. *et al.* (2017) 'Deep Galaxy: Classification of Galaxies based on Deep Convolutional Neural Networks'. Available at: http://arxiv.org/abs/1709.02245.

Kieffer, B. *et al.* (2018) 'Convolutional neural networks for histopathology image classification: Training vs. Using pre-trained networks', *Proceedings of the 7th International Conference on Image Processing Theory, Tools and Applications, IPTA 2017*, 2018-Janua, pp. 1–6. doi: 10.1109/IPTA.2017.8310149.

Livieris, I. *et al.* (2018) 'An Ensemble SSL Algorithm for Efficient Chest X-Ray Image Classification', *Journal of Imaging*, 4(7), p. 95. doi: 10.3390/jimaging4070095.

Ortiz, A. *et al.* (2016) 'Ensembles of Deep Learning Architectures for the Early Diagnosis of the Alzheimer's Disease', *International Journal of Neural Systems*, 26(07), p. 1650025. doi: 10.1142/s0129065716500258.

Perez, L. and Wang, J. (2017) 'The Effectiveness of Data Augmentation in Image Classification using Deep Learning'. Available at: http://arxiv.org/abs/1712.04621.

Rajpurkar, P. *et al.* (2017) 'CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning', pp. 3–9. Available at: http://arxiv.org/abs/1711.05225.
Shin, H. C. *et al.* (2016) 'Deep Convolutional Neural Networks for Computer-Aided Detection: CNN

Architectures, Dataset Characteristics and Transfer Learning', *IEEE Transactions on Medical Imaging*, 35(5), pp. 1285–1298. doi: 10.1109/TMI.2016.2528162.

Vianna, V. P. (2018) 'Study and development of a Computer-Aided Diagnosis system for classification of chest x-ray images using convolutional neural networks pre-trained for ImageNet and data augmentation', pp. 1–7. Available at: http://arxiv.org/abs/1806.00839.

Xia, X., Xu, C. and Nan, B. (2017) 'Inception-v3 for flower classification', *2017 2nd International Conference on Image, Vision and Computing, ICIVC 2017*, pp. 783–787. doi: 10.1109/ICIVC.2017.7984661.