Classifying Pulmonary Embolism cases in Chest CT scans using VGG16 and XGBOOST

A Major Project Report

Submitted to the Faculty

of

St. Joseph's College (Autonomous)

Bangalore University

by

Reshma Dua

Ronald Wallace

Tashi Chotso

in partial fulfilment of

the requirements for the degree of

Master of Science

in

Big Data Analytics

April 2022

Classifying Pulmonary Embolism cases in Chest CT scans using VGG16 and XGBOOST

A Major Project Report

Submitted to the Faculty

of

St Joseph's College (Autonomous)

Bangalore University

by

Reshma Dua

Ronald Wallace

Tashi Chotso

in partial fulfilment of

the requirements for the degree of

Master of Science

in

Big Data Analytics

April 2022

We would like to extend our sincere gratitude to the almighty, our parents and our project guides.

ST. JOSEPH'S COLLEGE (AUTONOMOUS)

36, Lalbagh Road, Bangalore, 560027

DEPARTMENT OF ADVANCED COMPUTING

CERTIFICATE

This is to certify that the Master of Science report entitled, "**Classifying Pulmonary Embolism cases in Chest CT Scans using VGG16 and XGBOOST**" is a bona fide work, done by *Reshma Dua, Ronald Wallace* and *Tashi Chotso* bearing register numbers 20BDA05, 20BDA10, 20BDA01 respectively in the 3rd semester during the year 2021-2022 in the partial fulfilment of the requirement for the award of Master of Science in Big Data Analytics from St. Joseph's College (Autonomous).

Dr Jayati Bhadra
Head of Department
Department of Advanced Computing
St. Joseph's College (Autonomous)

DECLARATION OF THE CANDIDATE

We hereby declare that this work entitled, "**Classifying Pulmonary Embolism cases in Chest CT Scans using VGG16 and XGBOOST**", has been originally carried out by *Reshma Dua, Ronald Wallace and Tashi Chotso* under the guidance of Francis Densil Raj V, Assistant Professor, Department of Advanced Computing, St. Joseph's College (Autonomous). This work has not been submitted elsewhere for the award of any other degree or diploma certificate.

Reshma Dua – 20BDA05
Ronald Wallace – 20BDA10
Tashi Chotso – 20BDA01

April 2022

M.Sc. Big Data Analytics Candidates

TABLE OF CONTENTS

LIST OF FIGURES

1. ABSTRACT

Pulmonary Embolism, often referred to as PE is a condition in which a blood clot becomes trapped in a pulmonary artery and prevents blood flow to the lungs. If left ignored, this might be life-threatening and, in most circumstances, fatal. Since the identification of whether a scan contains an embolus or not is a cumbersome process, we propose an approach using VGG16 and XGBOOST to classify whether an image contains an embolus or not. The dataset used has been downloaded from Kaggle and Segregated into two classes, namely 'PE' (The images that contain embolus) and 'No PE' (The images without any embolus in the lungs). Each directory contains over 1000 images. The methodology employed in this paper using VGG16 to extract features and XGBOOST to further classify images rendered an accuracy of 97.59% and a sensitivity of 97.00% with 5 misclassifications.

Keywords— Pulmonary Embolism, VGG16, XGBOOST, Mask RCNN

2. INTRODUCTION

Pulmonary Embolism (PE) is a condition in which a blood clot forms in the lung, when a clot from another region of the body (usually the leg or arm) travels through the bloodstream and lodges in the lung's blood arteries. This reduces blood flow to the lungs, lowers oxygen levels in the lungs, and raises pulmonary artery blood pressure. A thrombus is a blood clot that forms in a vein and stays there. An embolus occurs when a blood clot breaks free from the vein's wall and travels to another portion of your body. Deep vein thrombosis (DVT), a disorder in which blood clots form in veins deep within the body, is the most common cause of PEs. In today's practice, computed tomography pulmonary angiography (CTPA) is the major means of diagnosing PE, and it can successfully reduce the mortality rate[1].

A CT scan called a 'CTPA' looks for blood clots in the lungs *(also known as pulmonary embolism or PE).* The blood veins that travel from the heart to the lungs are photographed on a CT pulmonary angiography. A dye will be injected into a vein in the patient's arm that leads to their pulmonary arteries during the test. On scan images, this dye makes the arteries seem bright and white. The doctor will next be able to determine whether there are any blockages or blood clots.

Manually analysing a CTPA volume, on the other hand, requires a radiologist to meticulously trace each pulmonary artery over 300-500 slices for any suspected PEs, which is time-consuming and error-prone due to inexperience and eye strain. Medical image classification is one of the most pressing issues in image recognition, with the goal of categorising medical images into different groups to aid doctors in disease diagnosis and study.

3. OBJECTIVE

To build a model using Deep Learning and Machine Learning techniques to classify CT scan images into whether they contain embolus or not.

4. LITERATURE REVIEW

1. A novel method for pulmonary embolism detection in CTA images[2013]

This paper proposes a new computer-aided detection (CAD) – based method to detect pulmonary embolism (PE) in computed tomography angiography images (CTAI). Since lung vessel segmentation is the main objective to provide high sensitivity in PE detection, this method performs accurate lung vessel segmentation. To concatenate clogged vessels due to PEs, the starting region of PEs and some reference points (RPs) are determined. These RPs are detected according to the fixed anatomical structures. After the lung vessel tree is segmented, the region, intensity, and size of PEs are used to distinguish them. Data sets that have heart disease or abnormal tissues because of lung disease except PEis used to achieve the goal. According to the results, 428 of 450 PEs, labeled by the radiologists from 33 patients, have been detected. The sensitivity of the developed system is 95.1% at 14.4 false positive per data set (FP/ds). With this performance, the proposed CAD system is found quite useful to use as a second reader by the radiologists. [1]

Methodology Employed:

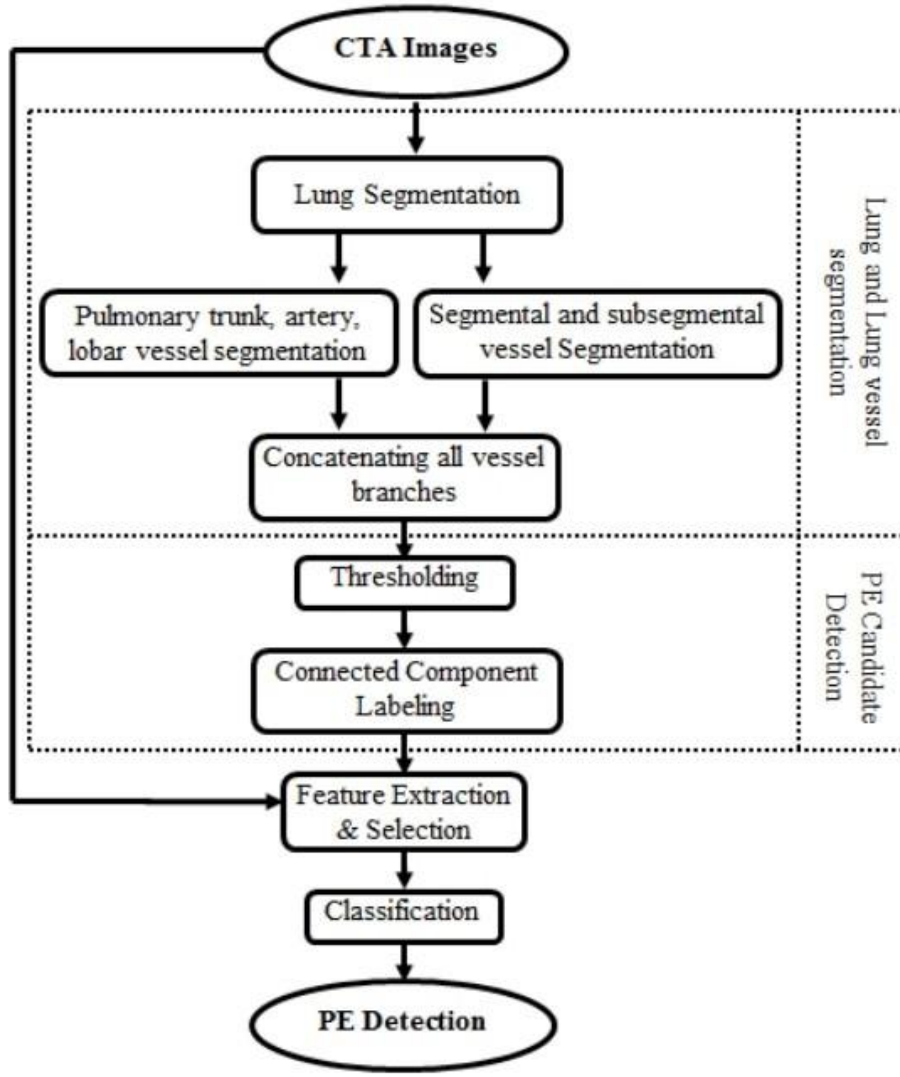

Figure 1:Flow chart of the proposed method

2. **Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks[2015]**

Computer-aided detection (CAD) can play a major role in diagnosing pulmonary embolism (PE) at CT pulmonary angiography (CTPA). However, despite their demonstrated utility, to achieve a clinically acceptable sensitivity, existing PE CAD systems generate a high number of false positives, imposing extra burdens on radiologists to adjudicate these superfluous CAD findings. In this study, the feasibility of convolutional neural networks (CNNs) as an effective mechanism for eliminating false positives is investigated. A critical issue in successfully utilizing CNNs for detecting an object in 3D images is to develop a “right” image representation for the object. Toward this end, a vessel aligned multi-planar image representation of emboli is developed. The image representation offers three advantages: (1)

efficiency and compactness—concisely summarizing the 3D contextual information around an embolus in only 2 image channels, (2) consistency—automatically aligning the embolus in the 2-channel images according to the orientation of the affected vessel, and (3) expandability—naturally supporting data augmentation for training CNNs. The author’s of the paper have evaluated their CAD approach using 121 CTPA datasets with a total of 326 emboli, achieving a sensitivity of 83% at 2 false positives per volume. This performance is superior to the best performing CAD system in the literature, which achieves a sensitivity of 71% at the same level of false positives. The system is further evaluated using the entire 20 CTPA test datasets from the PE challenge. The proposed system outperforms the winning system from the challenge at 0mm localization error but is outperformed by it at 2mm and 5mm localization errors. The performance at 0mm localization error is more important than those at 2mm and 5mm localization errors.. [2]

Proposed methodology: Vessel-Aligned Multi-planar Image Representation.

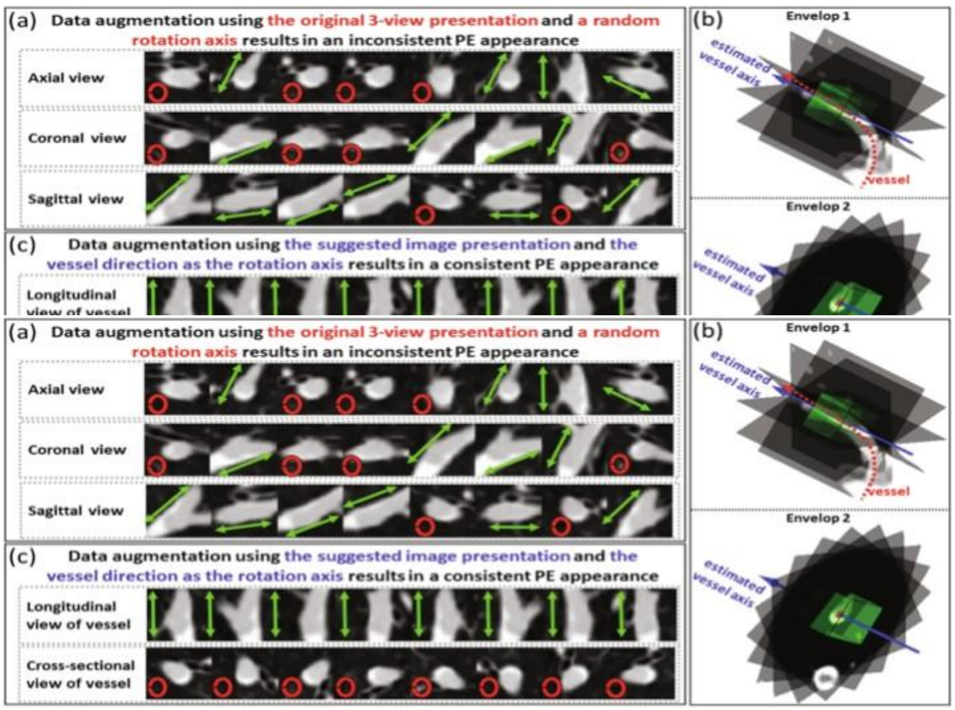

Figure 2:Data augmentation

3. **A Two-Stage Convolutional Neural Network for Pulmonary Embolism Detection From CTPA Images[2019]**

This paper presents a two-stage convolutional neural network (CNN) for automated detection of pulmonary embolisms (PEs) on CT pulmonary angiography (CTPA) images. The first stage utilizes a novel 3D candidate proposal network that detects a set of cubes containing suspected PEs from the entire 3D CTPA volume. In the second stage, each candidate cube is transformed to be aligned to the direction of the affected vessel and the cross-sections of the vessel-aligned cubes are input to a 2D classification network for false positive elimination. The approach is evaluated using both the test dataset from the PE challenge and the dataset consisting of 129 CTPA data with a total of 269 embolisms collected by the authors themselves. The experimental results demonstrate that the method used achieves a sensitivity of 75.4% at two false positives per scan at 0 mm localization error, which is superior to the winning system in the literature (i.e., sensitivity of 60.8% at the same level of false positives and localization error). On the dataset collected by the authors, the method employed achieves sensitivities of 76.3%, 78.9%, and 84.2% at two false positives per scan at 0, 2, and 5 mm localization error, respectively.[3]

For PE detection, Yang et al combined a 3D fully convolutional neural network (*FCN*) to detect candidate regions where the embolus is likely to be present, followed by a second stage that consists of a 2D cross-section of the vessel-aligned cubes and a ResNet-18 model for classifying and extract vessel-aligned 3D candidate cubes and eliminate the false positives.

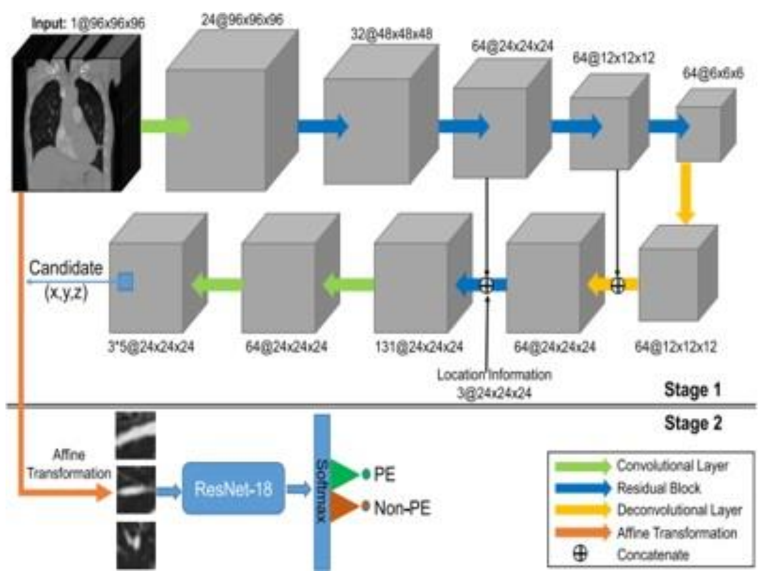

Figure 3: Framework of two stage PE detection

The figure above depicts the approach taken by Yang et al. The first stage is a 3D FCN that uses an encoder-decoder network in which the encoder enciphers hierarchical feature maps and the decoder up-samples the feature maps. This happens with the help of max pooling layers and residual blocks that are the middle layers of the FCN. Since PEs found in some unique regions, the apparent location information is also added to the derived feature map. An FCN feature map that consists of 64 channels is directly concatenated with the 3-channel location map. In the candidate proposal subnet, anchor cubes are used to identify candidate proposals from the concatenated 3D feature map, allowing for more accurate identification of variable size. The anchor cubes, in particular, are multiscale 3D windows that are predefined and centered at each scale. The second stage uses a classifier to remove as many false positives as feasible while maintaining a high sensitivity. Since all possible Embolus' could vary significantly in their appearance on all three cross-sections, vessel-aligned 2.5D is used to limit the apparent differences of PEs in the three cross-section slices. Each potential proposal is aligned to the orientation of the afflicted vessel using image representation.

4. **Deep convolutional neural network based medical image classification for disease diagnosis[2019].**

Medical image classification plays an essential role in clinical treatment and teaching tasks. However, the traditional method has reached its ceiling on performance. Moreover, by using them, much time and effort needs to be spent on extracting and selecting classification features. The deep neural network is an emerging machine learning method that has proven its potential for different classification tasks. Notably, the convolutional neural network dominates with the best results on varying image classification tasks. However, medical image datasets are hard to collect because it needs a lot of professional expertise to label them. Therefore, this paper researches how to apply the convolutional neural network (CNN) based algorithm on a chest X- ray dataset to classify pneumonia. Three techniques are evaluated through experiments. These are linear support vector machine classifiers with local rotation and orientation free features, transfer learning on two convolutional neural network models: Visual Geometry .

VGG16 and InceptionV3, and a capsule network training from scratch. Data augmentation is a data preprocessing method applied to all three methods. The results of the experiments show that data augmentation generally is an effective way for all three algorithms to improve performance. Also, Transfer learning is a more useful classification method on a small dataset compared to a support vector machine with oriented fast and rotated binary (ORB) robust independent elementary features and capsule network. In transfer learning, retraining specific features on a new target dataset is essential to improve performance. And, the second important factor is a proper network complexity that matches the scale of the dataset.

The comparison table of VGG16 and InceptionV3is shown below.

5. Convolutional Networks for Biomedical Image Segmentation[2015]

There is large consent that successful training of deep networks requires many thousand annotated training samples. In this paper, the authors present a network and training strategy that relies on the strong use of data augmentation to use the available annotated samples more efficiently. The architecture consists of a contracting path to capture context and a symmetric expanding path that enables precise localization. It is shown that such a network can be trained end-to-end from very few images and outperforms the prior best method (a sliding-window convolutional network) on the ISBI challenge for segmentation of neuronal structures in electron microscopic stacks. Using the same network trained on transmitted light microscopy images (phase contrast and DIC) the authors won the ISBI cell tracking challenge 2015 in these categories by a large margin. Moreover, the network is fast. Segmentation of a 512x512 image takes less than a second on a recent GPU.

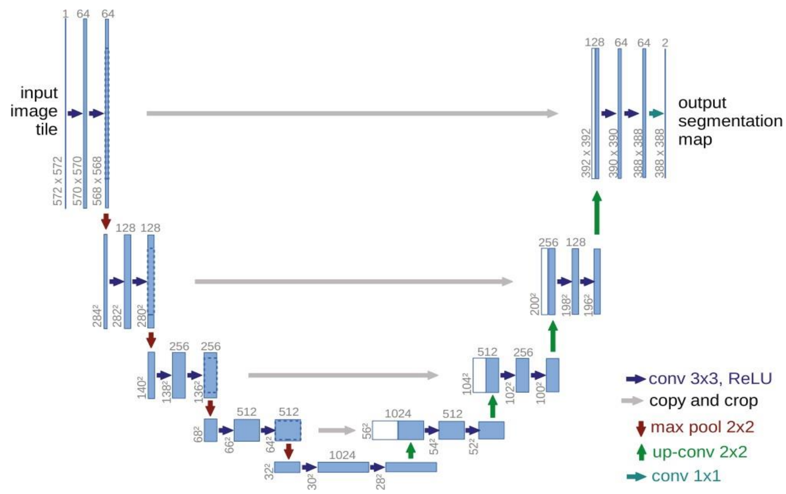

Figure 4: U-net architecture

U-net architecture (example for 32x32 pixels in the lowest resolution) is depicted above. Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

6. Deep Residual Learning for Image Recognition

Deeper neural networks are more difficult to train. A residual learning framework is presented to ease the training of networks that are substantially deeper than those used previously. The layers explicitly reformulated as learning residual functions with reference to the layer inputs, instead of learning unreferenced functions. A comprehensive empirical evidence showing that these residual networks are easier to optimize, and can gain accuracy from considerably increased depth is provided. On the ImageNet dataset the authors evaluate residual nets with a depth of up to 152 layers---8x deeper than VGG nets but still having lower complexity. An ensemble of these residual nets achieves 3.57% error on the ImageNet test set. This result won 1st place on the ILSVRC 2015 classification task. Analysis on CIFAR-10 with 100 and 1000 layers is also presented.The depth of representations is of central importance for many visual recognition tasks. Solely due to the extremely deep representations, 28% relative improvement on the COCO object detection dataset is obtained.

7. Accurate Pulmonary Nodule Detection in Computed Tomography Images Using Deep Convolutional Neural Networks.

Early detection of pulmonary cancer is the most promising way to enhance a patient's chance for survival. Accurate pulmonary nodule detection in computed tomography (CT) images is a crucial step in diagnosing pulmonary cancer. In this paper, inspired by the successful use of deep convolutional neural networks (DCNNs) in natural image recognition, we propose a novel pulmonary nodule detection approach based on DCNNs. We first introduce a deconvolutional structure to the Faster Region-based Convolutional Neural Network (Faster R-CNN) for candidate detection on axial slices. Then, a three-dimensional DCNN is presented for the subsequent false positive reduction. Experimental results of the LUng Nodule Analysis 2016 (LUNA16) Challenge demonstrate the superior detection performance of the proposed approach on nodule detection(average FROC-score of 0.891, ranking the 1st place over all submitted results).

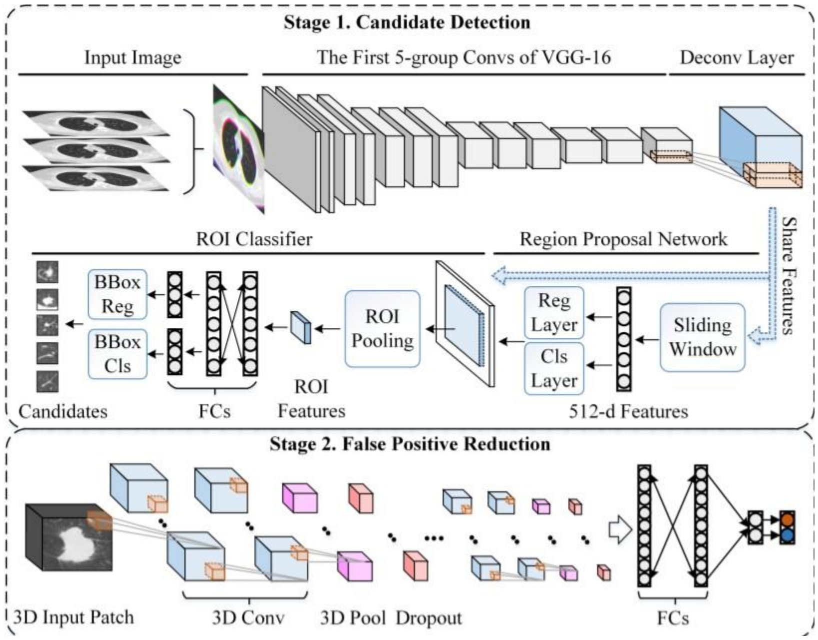

Figure 5: Architecture of the proposed CAD system

8. An investigation of XGBoost-based algorithm for breast cancer classification.

Breast cancer is one of the leading cancers affecting women around the world. The Computer-Aided Diagnosis (CAD) system is a powerful tool to assist pathologists during the process of diagnosing cancer, which effectively identifies the presence of cancerous cells. A standard CAD system includes processes of pre-processing, feature extraction, feature selection and classification. In this paper, we propose an enhanced breast cancer classification technique called Deep Learning and eXtreme Gradient Boosting (DLXGB) on histopathology breast cancer images using the BreaKHis dataset. This method first applies data augmentation and stain normalization for pre-processing, then pre-trained DenseNet201 will automatically learn features within an image and combine with a powerful gradient boosting classifier. The proposed classification technique is designed to classify breast cancer histology images into binary benign and malignant, and additionally one of eight non-overlapping/overlapping categories: i.e., Adenosis (A), Fibroadenoma (F), Phyllodes Tumour (PT), And Tubular Adenoma (TA) Ductal Carcinoma (DC), Lobular Carcinoma (LC), Mucinous Carcinoma (MC), And Papillary Carcinoma (PC). With DLXGB, we have obtained an accuracy of 97% for both binary and multi-classification, improving the existing work done by researchers using the BreaKHis dataset. The results indicated that this method could produce a powerful prediction for breast cancer image classification. The overview is given below.

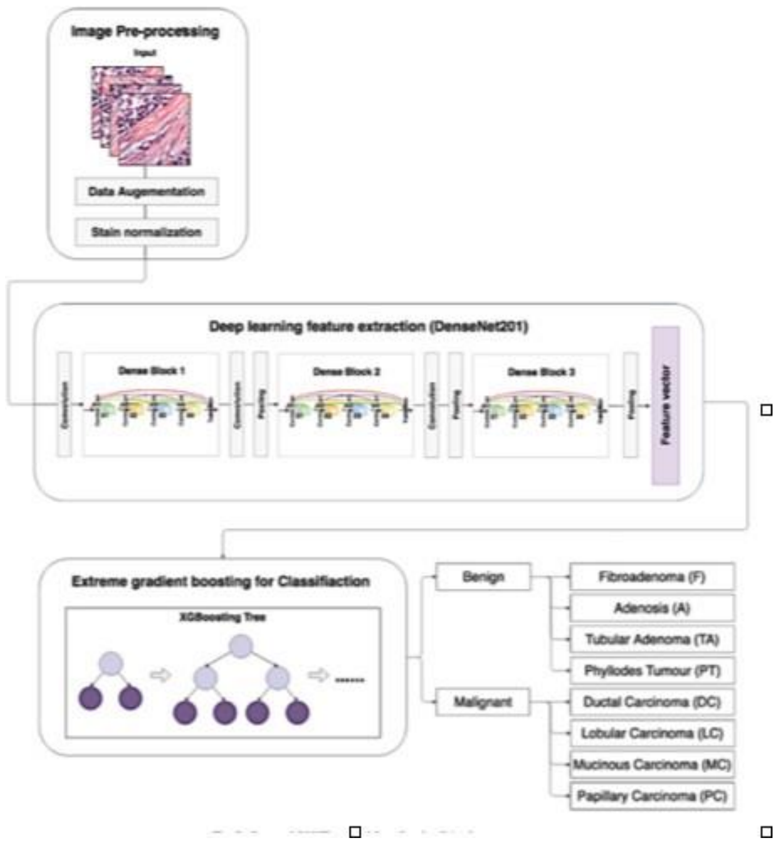

Figure 6: Proposed DLXGB mythology for classification breast tumors

9. Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images.

Traditionally, data mining algorithms and machine learning algorithms are engineered to approach the problems in isolation. These algorithms are employed to train the model in separation on a specific feature space and same distribution. Depending on the business case, a model is trained by applying a machine learning algorithm for a specific task. A widespread assumption in the field of machine learning is that training data and test data must have identical feature spaces with the underlying distribution. On the contrary, in real world this assumption may not hold and thus models need to be rebuilt from the scratch if features and distribution changes. It is an arduous process to collect related training data and rebuild the models. In such cases, Transferring of Knowledge or transfer learning from disparate domains would be desirable. Transfer learning is a method of reusing a pre-trained model knowledge for another task. Transfer learning can be used for classification, regression and clustering problems. This paper uses one of the pre-trained models – VGG - 16 with Deep Convolutional Neural Network to classify images.

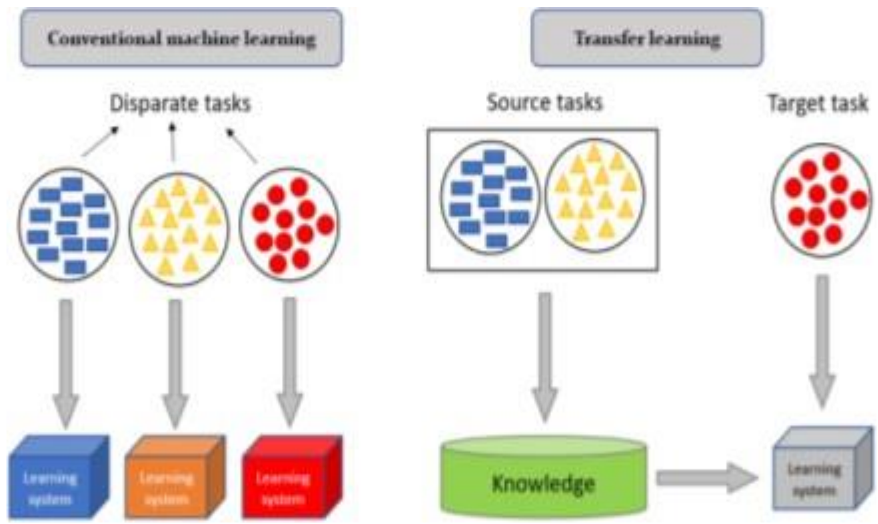

Figure 7:Comparative diagram of Learning Processes between Conventional Machine Learning and Transfer Learning

10. Synthetic Medical Images Using F&BGAN for Improved Lung Nodules Classification by Multi-Scale VGG16.

Lung cancer is one of the highest causes of cancer-related death in both men and women. Therefore, various diagnostic methods for lung nodules classification have been proposed to implement the early detection. Due to the limited amount and diversity of samples, these methods encounter some bottlenecks. In this paper, we intend to develop a method to enlarge the dataset and enhance the performance of pulmonary nodules classification. We propose a data augmentation method based on generative adversarial network (GAN), called Forward and Backward GAN (F&BGAN), which can generate high-quality synthetic medical images. F&BGAN has two stages, Forward GAN (FGAN) generates diverse images, and Backward GAN (BGAN) is used to improve the quality of images. Besides, a hierarchical learning framework, multi-scale VGG16 (M-VGG16) network, is proposed to extract discriminative features from alternating stacked layers. The methodology was evaluated on the Lung Image

Database Consortium and Image Database Resource Initiative (LIDC-IDRI) dataset, with the best accuracy of 95.24%, sensitivity of 98.67%, specificity of 92.47% and area under ROC curve (AUROC) of 0.980. Experimental results demonstrate the feasibility of F&BGAN in generating medical images and the effectiveness of M-VGG16 in classifying malignant and benign nodules

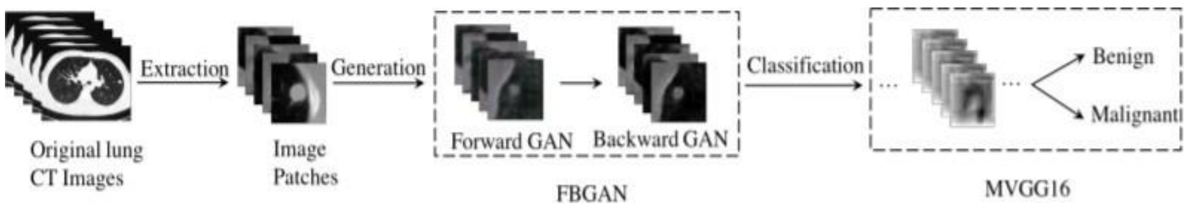

Figure 8:Overview of the whole structure. It is composed of a F&BGAN to generate diverse lung images and a multi-scale VGG16 network for classification

11. The Image Classification Method with CNN-XGBoost Model Based on Adaptive Particle Swarm Optimization

CNN is particularly effective in extracting spatial features. However, the single-layer classifier constructed by activation function in CNN is easily interfered by image noise, resulting in reduced classification accuracy. To solve the problem, the advanced ensemble model XGBoost is used to overcome the deficiency of a single classifier to classify image features. To further distinguish the extracted image features, a CNN-XGBoost image classification model optimized by APSO is proposed, where APSO optimizes the hyper-parameters on the overall architecture to promote the fusion of the two-stage model. The model is mainly composed of two parts: feature extractor CNN, which is used to automatically extract spatial features from images; feature classifier XGBoost is applied to classify features extracted after convolution. In the process of parameter optimization, to overcome the shortcoming that traditional PSO algorithm easily falls into a local optimal, the improved APSO guide the particles to search for optimization in space by two different strategies, which improves the diversity of particle population and prevents the algorithm from becoming trapped in local optima. The results on the image set show that the proposed model gets better results in image classification. Moreover, the APSO-XGBoost model performs well on the credit data, which indicates that the model has a good ability of credit scoring.

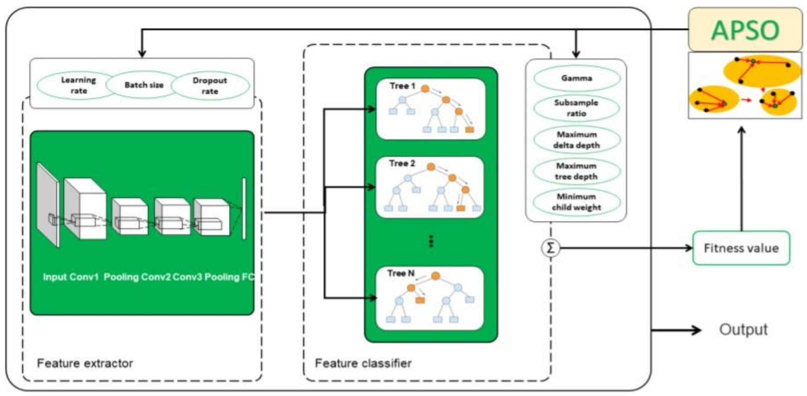

Figure 9:CNN-XGBoost based on APSO image classification model

12. A Novel Image Classification Method with CNN-XGBoost Model

Image classification problem is one of most important research directions in image processing and has become the focus of research in many years due to its diversity and complexity of image information. In view of the existing image classification models' failure to fully utilize the information of images, this paper proposes a novel image classification method of combining the Convolutional Neural Network (CNN) and eXtreme Gradient Boosting (XGBoost), which are two outstanding classifiers. The presented CNNXGBoost model provides more precise output by integrating CNN as a trainable feature extractor to automatically obtain features from input and XGBoost as a recognizer in the top level of the network to produce results. Experiments are implemented on the well-known MNIST and CIFAR-10 databases. The results prove that the new method performs better compared with other methods on the same databases, which verify the effectiveness of the proposed method in image classification problems.

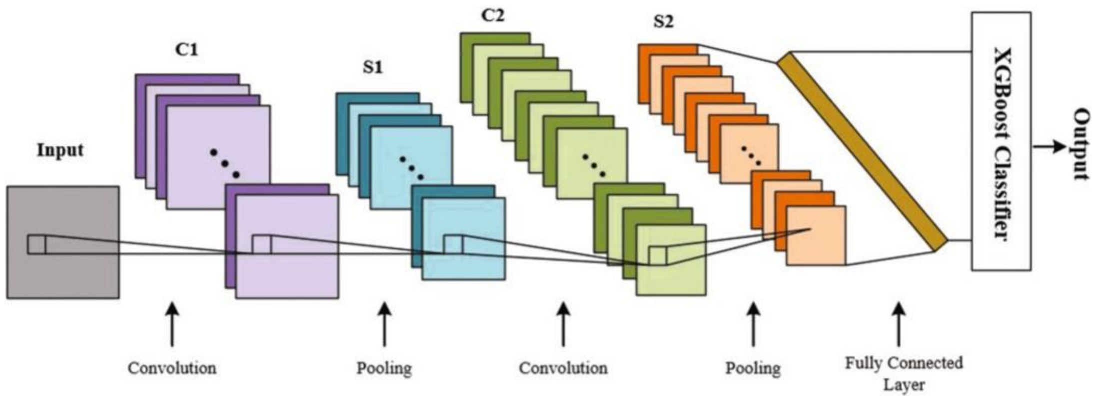

Figure 10:The Specific structure of CNN-XGBoost model for image classification

13. Medical Image Classification with Convolutional Neural Network

Image Classification is the process of classifying or predicting the class of a specific object in an image. This can mainly divided into two different categories : pixel based classification and object-based classification

Medical Image Classification is a key technique of computer-aided Diagnosis(CAD) systems.It is the primary domain, in which deep neural networks play the most important role of medical image analysis. The image classification accepts the given input images and produces output classification for identifying whether the disease is present or not.

Image patch classification is an important task in many different medical imaging applications. In this work, A customized Convolutional Neural Networks (CNN) with shallow convolution layer is being used to classify lung image patches with interstitial lung disease (ILD). While many feature descriptors have been proposed over the past years, they can be quite complicated and domain-specific. This proposed CNN framework can automatically and efficiently learn the intrinsic image features from lung image patches that are most suitable for the classification purpose. The same architecture can be generalized to perform other medical image or texture classification tasks.

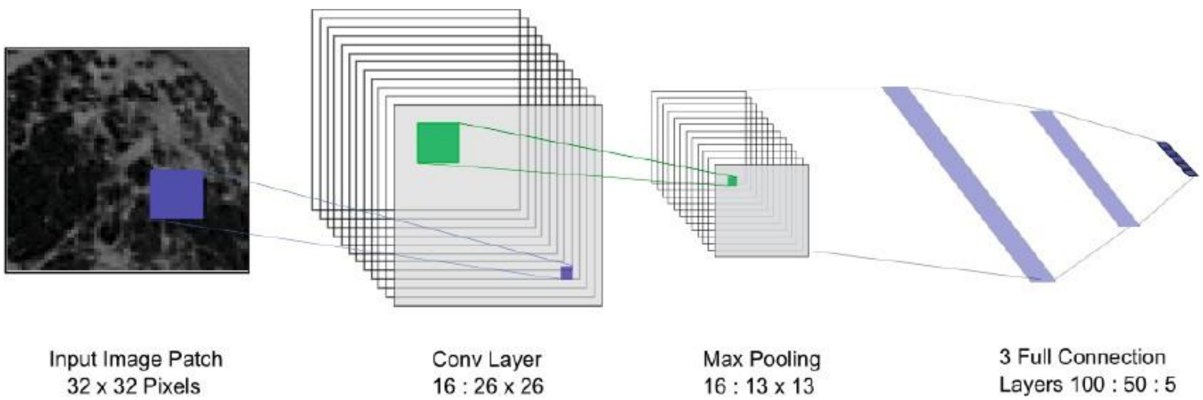

Figure 11: Network Architecture

14. Medical Image Data Classification Using Deep Learning Based Hybrid Model with CNN and Encoder

The Healthcare sector is one of the prime and different from other trades. Society expects high priority and highest level of services and care irrespective of money. Presently the medical field suffers from accurate diagnosis of diseases and it creates a huge loss to society. The prime factor for this is due to the nature of medical data, it is a combination of all varieties of data. Medical image analysis is a key method of Computer-Aided Diagnosis (CAD) frameworks. Customary strategies depend predominantly on the shape, shading, and additionally surface highlights just as their mixes, a large portion of which are issue explicit and have demonstrated to be integral in medical images, which prompts a framework that does not have the capacity to make portrayals of significant level issue area ideas and that has poor model speculation capacity. In this paper we are attempting a medical image data classification technique using hybrid deep learning technique based on Convolutional Neural Network (CNN) and encodes. What's more, we assess the proposed approach on two benchmark clinical picture datasets: HIS2828 and ISIC2017. The proposed algorithm is applied on the considered 2 datasets for performing data classification using deep learning based CNN and encoders. The proposed model is compared with the traditional methods and the results show that proposed model classification accuracy is better than the existing models.

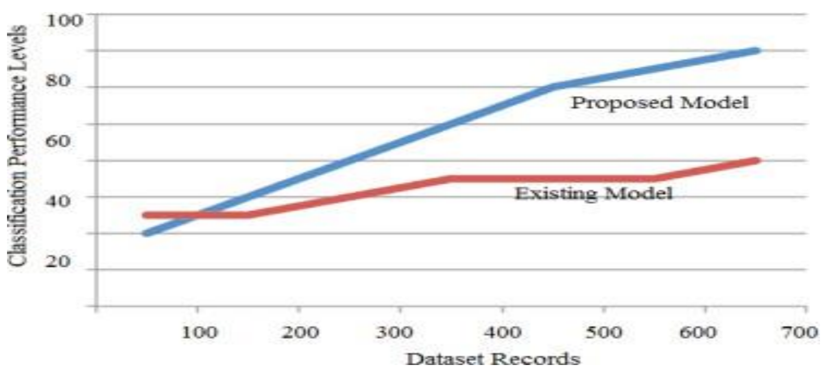

Figure 12: Hybrid Model performance vs Existing model

15. Deep learning for pulmonary embolism detection on computed tomography pulmonary angiogram

Computed tomographic pulmonary angiography (CTPA) is the gold standard for pulmonary embolism (PE) diagnosis. However, this diagnosis is susceptible to misdiagnosis. In this study, we aimed to perform a systematic review of current literature applying deep learning for the diagnosis of PE on CTPA. MEDLINE/PUBMED were searched for studies that reported on the accuracy of deep learning algorithms for PE on CTPA. The risk of bias was evaluated using the QUADAS-2 tool. Pooled sensitivity and specificity were calculated. Summary receiver operating characteristic curves were plotted. Seven studies met our inclusion criteria. A total of 36,847 CTPA studies were analyzed. All studies were retrospective. Five studies provided enough data to calculate summary estimates. The pooled sensitivity and specificity for PE detection were 0.88 (95% CI 0.803–0.927) and 0.86 (95% CI 0.756–0.924), respectively. Most studies had a high risk of bias. Our study suggests that deep learning models can detect PE on CTPA with satisfactory sensitivity and an acceptable number of false positive cases. Yet, these are only preliminary retrospective works, indicating the need for future research to determine the clinical impact of automated PE detection on patient care. Deep learning models are gradually being implemented in hospital systems, and it is important to understand the strengths and limitations of these algorithms.

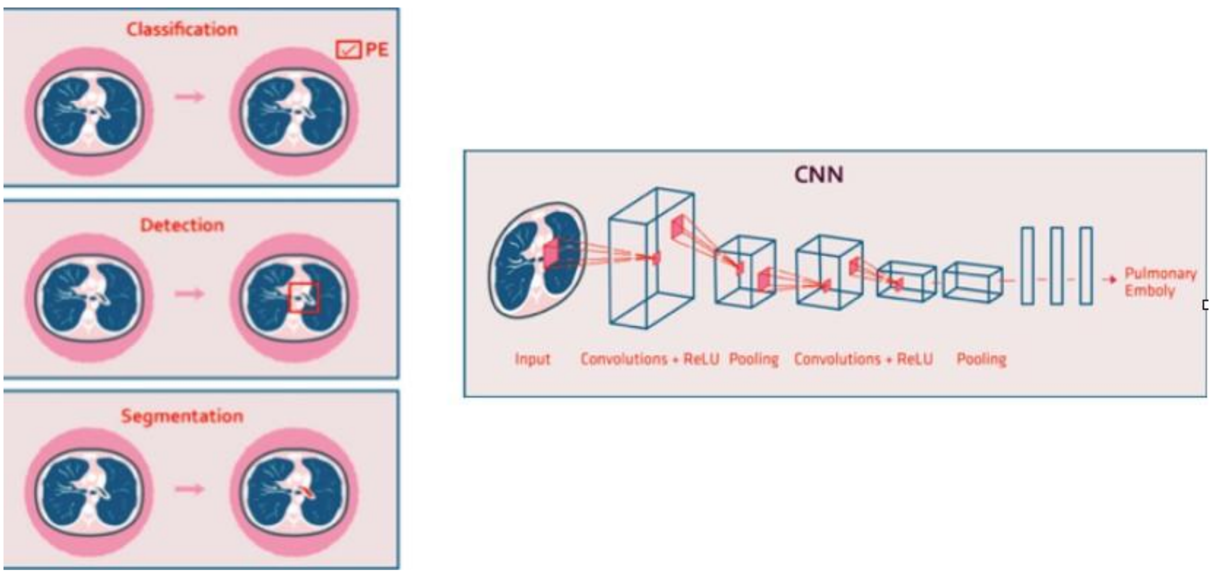

Figure 13: (a) Main computer vision task: classification, detection and segmentation (b) Architecture of proposed system.

Vijaykumar. T explains the reliability of capsule neural networks in comparison to the Convolutional neural networks that despite their hype are prone to performance degradation due to the process of reduction of dimensions for acquiring spatial invariance. The working of CapsNets are emphasised in which the author states that the Capsule neural networks take care of the performance degradation problem seen in convolutional neural networks by dividing the total images into sub – parts and hierarchical relating them and represents the picture with even better resolution than the CNN, the pooling layer in the CNN ignores the location relationship in the features leading to the degradation in the performance ,but the capsule shows effective feature extraction thus enhancing the classification accuracy [10].

The paper “Study of variants of extreme learning machine (ELM) brands and its performance measure on classification algorithm” described the variant of EML with various machine leaning algorithm and compared its accuracy and execution time. Nowadays, deep learning methods are widely used in many fields. Deep learning integrates feed-forward neural networks to train and test the model with high accuracy, but the feed-forward network functioning with slow computation time and increased gain. The ELM algorithm can overcome drawback of FFNN being redesigned the existing neural network with network components such as hidden nodes, weights and biases [11].

In the paper “Artificial Intelligence Algorithm with SVM classification using Dermascopic Images for Melanoma Diagnosis” an ordinal scale is used for assessing the likelihood of melanoma by clinicians on scale of 1 to 4 where 4 indicates most likely and 1 indicates least likely. All biopsied lesions are estimated for the likelihood of melanoma based and the clinical accuracy of the assessment is compared. It generated sensitivity of 95 % and 100% with using multiple decision thresholds.

The proposed AI algorithm with SVM classification is done on images obtained from digital single-lens reflex(DSLR) cameras with specificity and sensitivity 70% and 100% respectively [12].

The paper on “Analysis of Convolutional Neural Network based Image Classification Techniques” proposes a method to classify fruits by detecting the most important features of the images by applying filters or feature detectors to the input image in order to generate the feature maps or the activation maps by using the activation function. Fig.8. represents the overview of the proposed system.

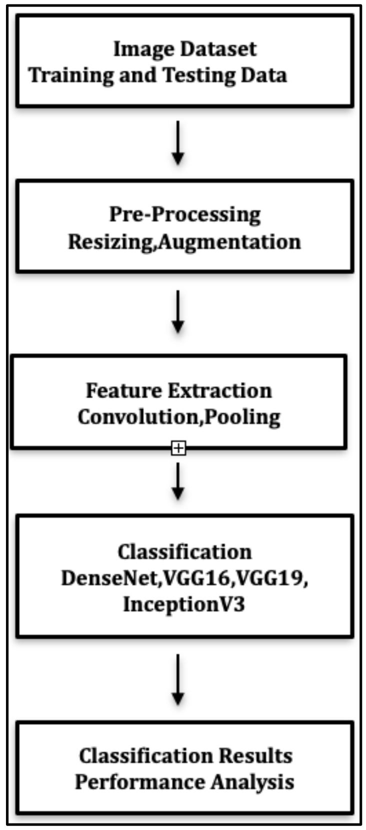

Figure 14: Overview of Proposed System

The proposed deep learning-based image classification system suggests DenseNet based model to classify the images effectively than the other classifiers by obtaining a training and testing accuracy of 99.25% and 100% respectively [13].

CNN is particularly effective in extracting spatial features. However, the single-layer classifier constructed by activation function in CNN is easily interfered by image noise, resulting in reduced classification accuracy. To solve the problem, the advanced ensemble model XGBOOST is used to overcome the deficiency of a single classifier to classify image features. To further distinguish the extracted image features, a CNN-XGBOOST image classification model optimized by APSO is proposed, where APSO optimizes the hyper-parameters on the overall architecture to promote the fusion of the two-stage model. The model is mainly composed of two parts: feature extractor CNN, which is used to automatically extract spatial features from images;

feature classifier XGBOOST is applied to classify features extracted after convolution. In the process of parameter optimization, to overcome the shortcoming that traditional PSO algorithm easily falls into a local optimal, the improved APSO guide the particles to search for optimization in space by two different strategies, which improves the diversity of particle population and prevents the algorithm from becoming trapped in local optima. The results on the image set show that the proposed model gets better results in image classification. Moreover, the APSO-XGBOOST model performs well on the credit data, which indicates that the model has a good ability of credit scoring [14].

5. COMPARATIVE STUDY

Below is a comparative study of a few reference papers from the literature review based on the methodology employed and the results obtained.

Paper	Method ology	Evaluatio n Metric
A Two-Stage Convolutional Neural Network for Pulmonary Embolism Detection From CTPA Images	Resnet-18	Sensitivity (75.4%)
Computer-aided pulmonary embolism detection using a novelvessel-aligned multi-planar image representation and convolutional neural networks	CNN	Sensitivity (83%)
U-net: Convolutional networks for biomedicalimage segmentation,	U-net	accuracy (77.5)

Artificial Intelligence Algorithm with SVM Classification using Dermascopic Images forMelanoma Diagnosis	AI with SVM	sensitivity (100%) speciRicity (70%)

Accurate Pulmonary Nodule Detection in Computed Tomography Images Using Deep Convolutional Neural Networks	CAD system	FROC-score (89%)
U-net: Convolutional networks for biomedical image segmentation,	RPN and FasterR-CNN	accuracy(70.4%)
A novel method for pulmonary embolism detection in CTA images	CAD	sensitivity (95.1%)
Analysis of ConvolutionalNeural Network based Image Classification Techniques	DenseNet	training accuracy: (99.25%) testing accuracy: (100%)
Deep convolutional neuralnetwork based medical image classifcation for disease diagnosis	Capsnet	74%

6. PROPOSED METHOD

The methodology proposed aims to classify Pulmonary Embolism CTPA scans correctly into PE and No PE with minimal misclassifications. The flowchart below illustrates the procedure step by step.

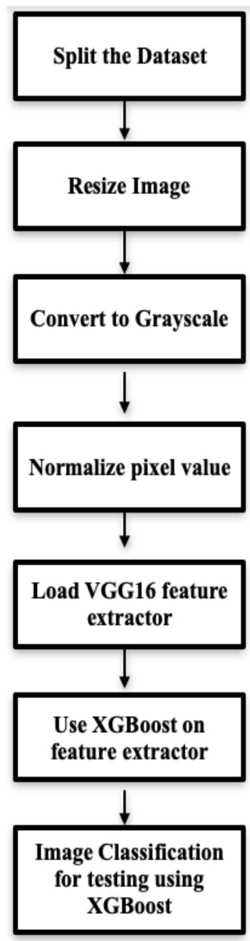

Figure15: Flowchart of Proposed Methodology

Visual Geometric Group-16(VGG16) is a Convolutional Neural Network (CNN) model which is used a lot in medical image classification and detection. It consists of 16 layers which have tunable parameters out of which 13 layers are convolutional layers with non- linear activation function(ReLu) and 3 fully connected layers with the same non-linear activation function. In addition to these 16 convolutional layers there are Max Pool layers for reducing the dimensions of the feature map and dense layers with the Softmax activation function that renders classified output.

Every convolutional layer has a receptive field of 3x3 and a stride of 1. The VGG16 takes in images in the size of 224x224. It uses row and column padding to maintain spatial resolution after convolution. The Max Pool window size is observed to 2x2 with a stride of 2. Not all convolutional layers are followed by max-pool layer. The first two fully connected layer contain 4096 channels each and the final fully connected layer contains 1000 channels. The last layer is a softmax layer with 1000 channels, one for each class of images in the dataset. VGG16 is also a go-to-model when one has very few images to train on.

We have made use of the model to extract features from the images and the input layer which is a convolutional layer takes in a fixed image size of 256x256 instead of 224x224.

Dense layers are dropped as a classification task is taken care of by the XGBOOST model that succeeds VGG16, only convolutional layers are used to extract features and these features are then flattened and fed into the XGBOOST for classification.

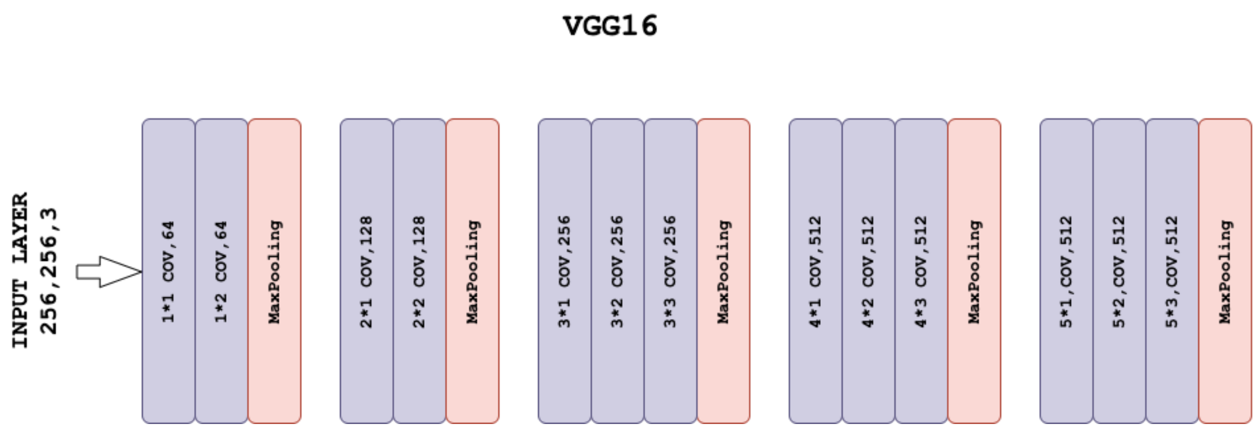

Figure 16:VGG-16 Architecture map for feature extraction

The features of shape (8,8,512) are extracted from the VGG16 model and flattened to 32768. These features are used to train the XGBOOST model which works on the principle of extreme gradient boosting using Gradient Boosted decision trees. All of the independent variables are given weights, which are fed into the decision tree, which predicts outcomes. The weight of variables that the tree predicted incorrectly is increased, and these variables are then fed into the second decision tree.

These individual classifiers/predictors are then combined to form a more powerful and precise model. After the model is trained on the extracted features the same process of feature extraction and reshaping is applied to the test set in order to get the model performance.

The process of how XGBOOST model arrives at a better hypothesis using outputs from number of weak leaners is illustrated in the image below –

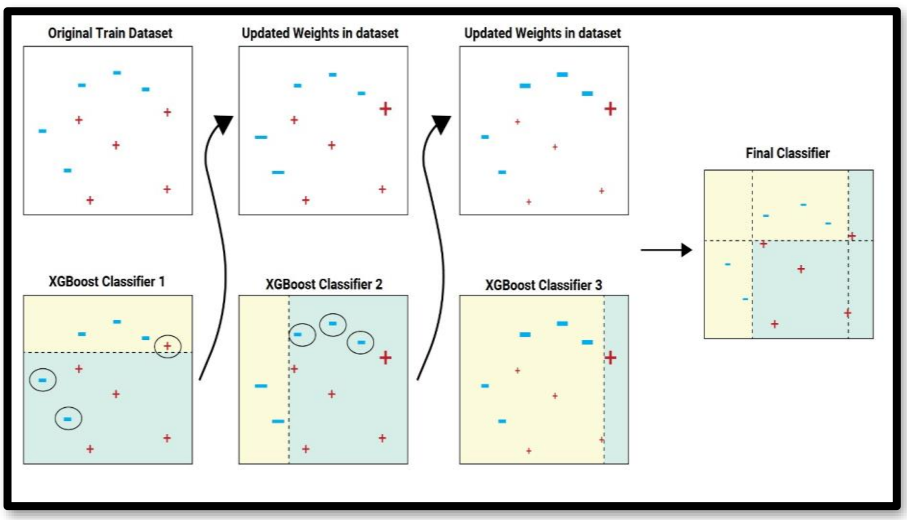

Figure 17: Working of XGBOOST algorithm

6.1. About the Dataset

The RSNA STR Pulmonary Embolism Dataset

The Radiological Society of North America(RSNA) has partnered up with the Society of Thoracic Radiology (STR) to increase the use of machine learning in the diagnosis of PE [15]. We have evaluated our method on the RSNA STR PE Detection Dataset available on Kaggle for academic research and education.

File Description

“train.csv contains UIDs and labels which is the metadata for all images.”

“Dataset contains two subdirectories for positive and negative PE cases.”

Data Fields

“Study Instance UID is the unique ID for each study in the dataset.”

“Series Instance UID is the unique ID for each series within the study.”

“SOP Instance UID is the unique ID for each image within the study.”

“pe_present_on_image indicates the image-level, notes whether any form of PE is present on the image.”

The images were downloaded from Kaggle in .jpg format and the train.csv file was used to segregate the images into two classes, namely ‘PE’ (The images that contain embolism) and ‘No PE’ (The images without any embolus in the lungs) based on “pe_present_on_image” column.

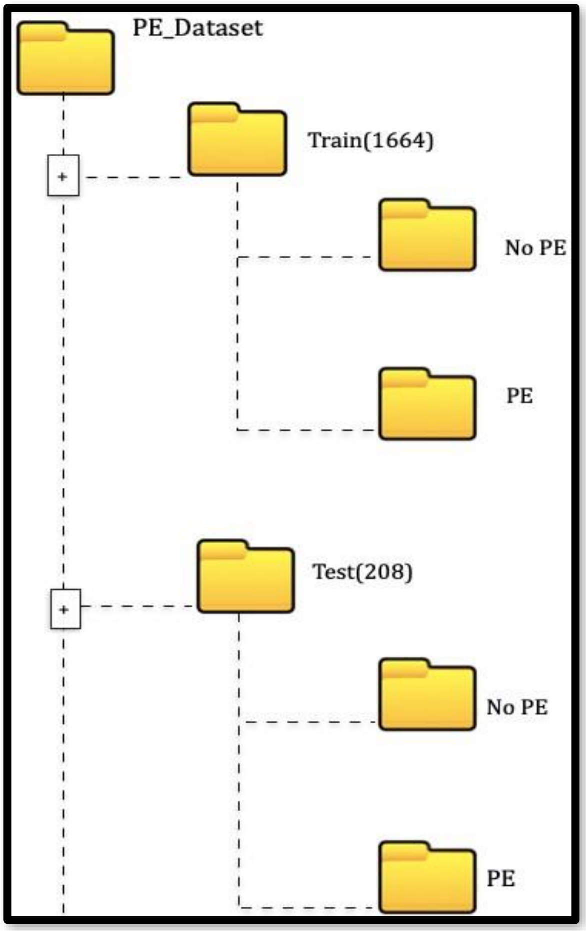

Figure 18:Image split in different directories.

6.2 CREATING TRAIN AND TEST SETS

After the segregation of images into two different classes we split the data into train and test data such that there are 832 images in the train set and 104 images in the test set. This is done so that our model will have enough data with variety to learn from and render better reliability and performance.

6.3 MODEL BUILDING

Image interpretation by computer programmes has been an important and active topic in the machine learning discipline, as well as in application and research specific investigations, due to the rapid growth of digital image collecting and storage technology. Medical image classification is a challenging task in Deep Learning which tries to classify medical images into distinct categories to aid doctors and radiologists in diagnosing disease or doing additional study.

The classification of medical images can be broken down into two parts. The first step is to take a photograph and extract information from it. The characteristics are then used to classify the image collection and generate models in the second stage. Doctors/radiologists formerly used their professional experience to extract features from medical images in order to classify them into distinct classes, which was a tedious, monotonous, and time-consuming task.The approach to classify the medical image is time consuming if done manually. Medical image classification application study has a lot of merit, based on previous research. The efforts of the scholars have resulted in a huge number of published works in this field.

However, we are still unable to complete this mission effectively at this time. If we can complete the classification process well, the data will aid medical doctors in diagnosing disorders that require further investigation.

As a result, figuring out how to complete this duty properly is crucial, and the topic of our paper

6. 4 MODEL EVALUATION METRICS

The classification report includes various metrics used for evaluation of our model such as –

Precision which gives us the percentage of correctly classified images of the in both the classes, Recall or the true positive rate explains the ability of our model to classify whether a particular image belongs to a certain class, the f1 score, a classification metric that combines the values of both precision and recall to give an overall measure of model performance. The accuracy is seldom used as a classification metric, but in our case, since the number of images in either of the classes are equal we can also calculate the accuracy that determines the percentage of correctly classified images of the model. Specificity or true negative rate is a metric that evaluates a model's ability to predict if an image does not belong to a certain class. The misclassification rate depicts the number of misclassified images with respect to the total number of images in that particular class. The formulas for the above metrics are given as follows –

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives}$$

$$f1 - score = \frac{2*Precision*Recall}{Precision + Recall}$$

$$Specificity = \frac{TrueNegatives}{TrueNegatives + FalsePostitves}$$

$$Accuracy = \frac{TruePositives + TrueNegatives}{TruePositives + FalseNegatives + TrueNegatives + FalsePositives}$$

An ROC is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters:

- True Positive Rate
- False Positive Rate

7. RESULTS

Fig.13. represents the confusion matrix which describes the performance of our model on test data. The proposed methodology has performed well with very few misclassifications- 5 out of 104 unseen test data points.

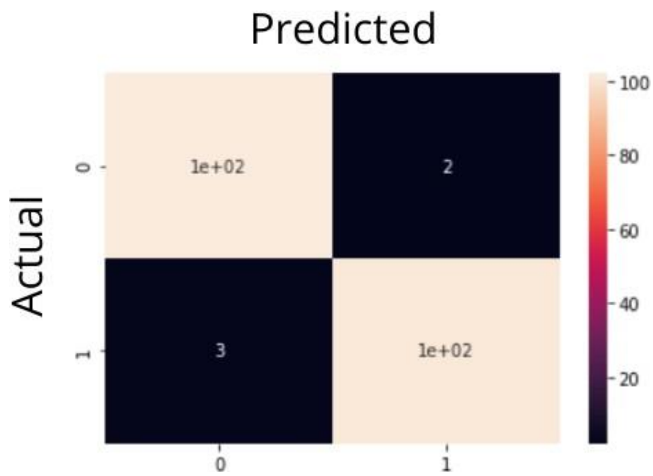

Figure 19: Confusion Matrix

Fig.14. represents the classification report that is used to assess the quality of predictions made by the algorithm.

	precision	recall	f1-score	support
No PE	0.97	0.98	0.98	104
PE	0.98	0.97	0.98	104
accuracy			0.98	208
macro avg	0.98	0.98	0.98	208
weighted avg	0.98	0.98	0.98	208

Figure 20: Classification Report

Our model achieves an accuracy of 97.59% and an overall sensitivity of 97.00%, i,e; The model can successfully predict if a given image does not belong to a particular class 97 times out of 100. For The "PE" class the sensitivity is observed to be 97.00% and for the "No PE" class it is seen to be 98.00% respectively. The misclassifications are also very low at about 5 misclassifications out of 104 unseen images which leaves us with a misclassification error rate of 4.80%. As for the Precision, Recall and f1-score, Our model achieves a precision of 98.00% on the "PE" class and 97.00% of the "NO PE" and a recall of 97.00% on the "PE" class and 98.00% on the "NO PE" class respectively. f1-score of our model is observed to be 98.00% for both the classes.

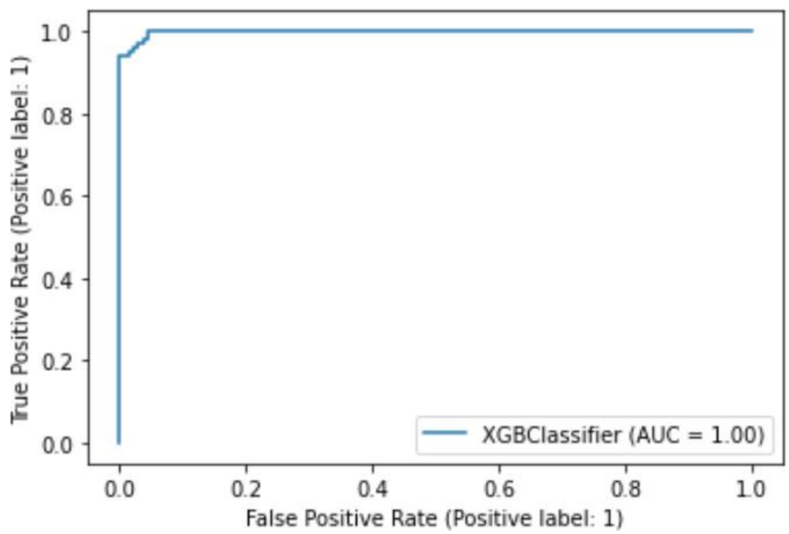

Figure 21. Receiver Operator Characteristic (ROC) curve

Classifiers that give curves closer to the top-left corner indicate a better performance and hence our proposed classifier (XGBOOST built on top of VGG16) has proven to perform better.

8. CONCLUSION

This paper represents a VGG16 model in combination with boosting technique (XGBOOST) to successfully classify the images into two classes.

Prior to selecting the model stated above, different models and techniques were experimented on, but they did not prove to be effective due to the lack of data needed to feed the model. ResNet50 was one of the architectures that was used, rendering a validation accuracy of 71.98%.

Since there was a constraint on the number of images available for building our model an alternative and a much more effective way was employed in which a VGG16 model was used for feature extraction, which is observed to perform well on medical images in combination with Boosting techniques (XGBOOST). This technique on the other hand, was observed to perform much better than all other techniques with an accuracy of 97.59% and a sensitivity of 97.00% with 5 misclassifications. Alternate approach is also mentioned in the literature review section, which describes a two-stage approach to classify and detect 3D images

9. FUTURE WORK

Now that we have successfully constructed a model to classify our image to whether it has an embolus or not, our next task is to optimize this objective by coming up with methods that would serve the purpose of localising and segmenting the embolism without compromising on the efficiency of our model.

A model that uses an instance segmentation method to segment the clot from an image can be used. Though the images acquired are not annotated and the approximate coordinates are unknown, VGG image Annotator software can be used to segment the area where the clot is likely to be present. The annotated images are then used to train models that use the principle of object detection and image segmentation. Mask RCNN would be a better model to start from and proceed further as it is shown to perform well with medical data. This method comprises of two stages –

STAGE-1: ANNOTATION OF IMAGE

The segment where the clot is likely to be present in the image is annotated using VGG image Annotator software. This is done for about 30-40 images. These many images would suffice for this approach of classifying the images and most models that use instance segmentation approach tend to perform better with the number of images stated above.

Figure 22: Segmented Pulmonary Embolus

Fig.16. is annotated using the annotation software – VGG image annotator, the segmented part is where the embolus is observed to be present in the lungs of the patient

STAGE-2: MODEL BUILDING AND OPTIMISATION

The next step in the process is to train a model like Mask RCNN on the annotated images and use it to segment embolus from the CTPA images and optimise it.

10. REFERENCES

1. X. Yang, “A Two-Stage Convolutional Neural Network for Pulmonary Embolism Detection from CTPA Images,” IEEE Access, vol.7, (2019).
2. “Pulmonary Embolism imaging and outcomes,” AJR American Journal of Roentgenology, (2012).
3. O. O. S. .. a. A. F. B. H. Özkan, “A novel method for pulmonary embolism detection in CTA images,” Comput. Methods Programs Biomed, Vols. vol. 113, no. 3, pp. 757 766, (2014).
4. S. J. S. D. Yadav, “Convolutional neural network based medical image classification for disease diagnosis,” Vols. J Big Data 6, 113 , (2019).
5. M. B. G. a. J. L. N. Tajbakhsh, “Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks,” Springer, (2015).
6. P. F. a. T. B. O. Ronneberger, “U-net: Convolutional networks for biomedical image segmentation,” vol. pp. 234–241, (2015).
7. X. Z. S. R. a. J. S. K. He, “Deep residual learning for image recognition,” IEEE, vol. pp. 770–778., (2016).
8. M. W. Y. Xin, “Research on image classification model based on deep convolution neural network,” (2019).
9. A. Bashar, “Survey on evolving deep learning neural network architectures,” Journal of Artificial Intelligence 1, vol. no. 02, pp. 73-82, (2019.
10. T. Vijayakumar, “Comparative study of capsule neural network in various applications,” Journal of Artificial intelligence 1, vol. no. 01, pp. 19-27, (2019).

11. J. S. Manoharan, "Study of Variants of Extreme Learning Machine (ELM) Brands and its Performance Measure on Classification Algorithm," Journal of Soft Computing Paradigm (JSCP)3, vol. no. 02, pp. 83-95, (2021).
12. V. Balasubramaniam, "Artificial Intelligence Algorithm with SVM Classification using Dermascopic Images for Melanoma Diagnosis," Journal of Artificial Intelligence and Capsule Networks 3, vol. no. 1, pp. 34-42, (2021).
13. M. Tripathi, "Analysis of Convolutional Neural Network based Image Classification Techniques," Journal of Innovative Image Processing (JIIP) 3, vol. no. 02 , pp. 100-117, (2021).
14. S. J. S. Yadav, " Deep convolutional neural network based medical image classification for disease diagnosis," Vols. J Big Data 6, 113, (2019).
15. "Radiology: Artificial Intelligence," The RSNA Pulmonary Embolism CT Dataset, (2021). [Online]. Available: https://pubs.rsna.org/doi/full/10.1148/ryai.2021200254.

11. CONFERENCE CERTIFICATES

Springer

ICICV

FRANCIS XAVIER ENGINEERING COLLEGE

AN AUTONOMOUS INSTITUTION

Certificate of Presentation

This certificate is awarded to

Reshma Dua

for successfully presenting the paper entitled

Classifying Pulmonary Embolism cases in Chest CT scans using VGG16 and XGBOOST

at the

4th International Conference on

Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2022)

organized by Francis Xavier Engineering College, Tirunelveli, India

on the 10th & 11th of February 2022.

Session Chair

Conference Chair

Dr. G. Rajakumar

Principal

Dr. V. Velmurugan

Springer

ICICV

FRANCIS XAVIER ENGINEERING COLLEGE
AN AUTONOMOUS INSTITUTION

Certificate of Presentation

This certificate is awarded to

Ronald Wallace G

for successfully presenting the paper entitled

Classifying Pulmonary Embolism cases in Chest CT scans using VGG16 and XGBOOST

at the
4th International Conference on
Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2022)
organized by Francis Xavier Engineering College, Tirunelveli, India
on the 10th & 11th of February 2022.

Session Chair	Conference Chair Dr. G. Rajakumar	Principal Dr. V. Velmurugan

Springer

ICICV

FRANCIS XAVIER ENGINEERING COLLEGE
AN AUTONOMOUS INSTITUTION

Certificate of Presentation

This certificate is awarded to

Tashi Chotso

for successfully presenting the paper entitled

Classifying Pulmonary Embolism cases in Chest CT scans using VGG16 and XGBOOST

at the
4th International Conference on
Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2022)
organized by Francis Xavier Engineering College, Tirunelveli, India
on the 10th & 11th of February 2022.

Session Chair

Conference Chair
Dr. G. Rajakumar

Principal
Dr. V. Velmurugan