**ImageNet Classification with Deep Convolutional Neural Networks**

Author has trained largest convolutional neural network to classify 1.2 million images of high resolution with 60 million parameters on the Imagenet subset and gained the best result. Network contains new and unusual features which reduces the training time and improved the performance using supervised learning. Architecture of proposed system contains five convolutional layers and three fully connected layers and feeded output to the softmax function. Softmax output layer is used for probability distribution of 1000 classes. Relu Nonlinearity activation is connected to all convolutional network as well as fully connected layer. Researchers have demonstrated that Relu activation trains faster than the tanh activation in Deep Neural network. Relu has the property that they don’t need normalization of input to prevent from saturation. Effectiveness of Relu’s property has been tested on the CIFR10 dataset. With normalization model has achieved 11% error rate and without normalization model has achieved 13% error rate. To make training faster the learning rate were chosen independently for each network. To train the data two GPU’s has been used because a single GTX 580 GPU has only 3 giga byte memory. These Gpu’s having ability to read from another GPU and write on another GPU. Resultant model is same as the “columner” CNN which is employed by Cirenan but, only difference is that columns of model are dependent. This strategy has reduced the error rates by 1.7% and 1.2% of top-1 and top-5 errors.

Pooling layers in CNNs sum up the yields of neighboring gatherings of neurons in a similar kernel map. Generally, the areas summed up by contiguous pooling units don't cover. This system contains overlap pooling with s=2 and z=3. This strategy has also reduced top-1 error rate by 0.4% and top-5 error rate by 0.3% as compared to no overlapping pooling. Proposed system has prevented the issue of overfitting using data augmentation and dropout method. Two forms of data augmentation has been employed, in first method transformed images has been generated using python code on CPU and in second method intensities of RGB channels of training images has been altered. Performed principle component analysis on RGB channels of imagenet dataset for data augmentation. This strategy has reduced top-1 error rate by 1%. Also, to reduce overfitting dropout has been applied to first fully connected layers. This procedure has reduced complex co-adjustments of neurons, since a neuron can't depend on the nearness of specific different neurons. It is, subsequently, forced to learn increasingly robust highlights that are valuable related to a wide range of random subsets of the different neurons. Modelhas been trained on stochastic gradient descent with momentum of 0.9,batch size of 128 examples and 0.0005 weight decay. To learn the model small amount of decay is very important which nothing but regularizer for implemented model is. Introduced the weights in each layer from a zero-mean Gaussian conveyance with standard deviation 0.01. also, introduced the neuron biases in the second, fourth, and fifth convolutional layers, just as in the fully connected hidden layers, with the consistent 1. Learning rate used for throughout the model is same.

It is remarkable, that implemented system's performance debases if a single convolutional layer is expelled. For instance, removal of any of the hidden layers brings lost about 2% for networks top-1 performance. So the depth truly significant for accomplishing the outcomes.

**"Going Deeper with Convolutions"**

The Paper proposed a deep CNN architecture for computer vision, alias inception which attains the new state of act for ILSVRC14's classification and detection. CNN had one or fully-connected stacked layers where this design was common in the image classification literature by giving in best results on MNIST, CIFAR to date, and is employed successfully for localization despite the concern of losing accurate spatial information due to max-pooling layers. Recent trends suggested that Increasing the number of layers classifies larger datasets while the overfitting problem was addressed using dropout. The area’s current state object detection is the R-CNN which parts the complete detection issue towards two subproblems: utilizing low-level cues to generate object location proposals, CNN classifiers are used in identification of object grouping at those locations. Multi-box prediction and ensemble approaches are used to enhance both stages.

Evolution in convolutional networks and deep learning has improved our detection and classification potential in the last three years. A result of new ideas, efficient algorithms in accordance with the ongoing pull of embedded computing and mobile, and Improved network architecture along with huge datasets, strong hardware, and larger models have resulted in the advancement while not using any new data sources apart from the classification dataset of ILSVRC 2014 competition where GoogLeNet submission has actually used 12 times lesser parameters than the Krizhevskys architecture which was won.

As the Inception architecture is made to improve the performance of deep learning and convolutional networks (CNN) the most straightforward way would be to increase the network. Which means increasing the depth, number of network layers and width of the network uniformly. This is an easy and safe way to improve performance but there are two major drawbacks. Larger the network means more parameters which means the training model is prone to overfitting. The other drawback is deeper the network higher the computational potential required, which means in order to process the data from this deep network a very powerful computer is required which takes a hit on our budget.

These drawbacks can be solved with a simple solution, which is by introducing sparsity and by replacing the fully connected layers by sparse layers. This is similar to widening a fully connected network by branching it. Even this solution has a drawback which is with the current technology; computation of a non-uniform sparse matrix is very inefficient. Therefore, the “Inception” model proposed by the authors of this paper mainly focuses on this drawback.

Inception architecture's main idea is to consider how a network can be approached by readily dense components and finding the optimal local construction and repeating it spatially is what is needed. A combination of the layer-by-layer constructions with their output filter banks is suggested and additionally, pooling operation is essential for the current convolutional network's success. The cost gets high when there is an addition in pooling sets. This results in another inception idea which is, sensibly reducing the dimension whenever there is an increase in computational requirement which embeddings success based.

Generally, the Inception network is a stacked module with occasional max-looping layers and is beneficial at the higher layers only while keeping traditional convolutional fashion in lower layers. "GoogLeNet" refers to the particular incarnation which used one deeper and wider Inception networks which improved the results marginally and are submitted in ILSVRC 2014 competition. These networks were trained using Disbelief distributed machine learning system with the help of model and data-parallelism and it was found that photometric distortions were used in combating overfitting.

Due to addition of the auxiliary classifiers to connect the intermediate layers, classifiers lower stage discrimination was expected and at the inference time, these classifiers are discarded. Later

, control experiments expressed that these networks have a minor effect only.

Independently, 7 versions which are of alike model were trained and ensemble prediction was performed where only unarranged input image order and sampling procedure differed with a more aggressive cropping approach where the resizing of the image was done to 4 scales and softmax probabilities over multiple crops are averaged to obtain a final prediction.

The detective task of this challenge is to process bounding boxes in the images among 200 classes possible. They count to be correct if they go with class of the ground-truth and the GoogLeNet approach is almost equivalent to the R-CNN with Inception model augmented as the region classifier.

Our approach shows that moving to the sparser architectures is feasible and is a very useful idea for both classification and detection.

**“Deeppose: Human pose estimation via deep neural networks”**

Human pose estimation, which is human joints localization, had been shown interest by the computer vision community. There were so many models made to estimate human pose which had less efficiency. In this paper, DNN has been used to develop a method to estimate human pose with the help of body joints regression problem as DNN had shown very good performance in visual classification and object localization.

Seven layers of generic convolution DNN is employed as it has two advantages. Firstly, each body joints context can be fully captured by DNN, and secondly, previously graphical methods were used which were not very simple so, a simpler sustainable approach was employed. Furthur, cascade was proposed to increase joint localization’s precisión.

Previously, PSs were tractable using distance transform but there were some limitations as it was not depending on image data. So, this research focused more on highlighting the representation power while sustaining tractability. There were other approaches where locality-sensitive hashing was used , semi global classifiers were proposed for expressing arms only, another work which is approaching our research has used CNN with analysis of the neighboring component.

The approach we have taken is, we first encode the body joints locations in a pose vector, then normalize them concerning a box, then the joint is translated and scaled, pose vector is normalized and finally, by the box, we de-facto normalize image.

In this word, we handle the pose estimation problem as regression and the image coordinates are read using normalization transformation, and not only this simple formulation, it shows less complexity and power as it is based on the convolutional DNN approach which has seven layers of simultaneous linear and non-linear transformation. Input is taken in the first layer as image and target value which is outputs is given by the last layer. The outstanding results of DNN architectures and its need to de-sign the pose model have motivated us to use it on localization problems and classification problems.

Linear regression is trained on the top of the last layer to predict the pose vector by depreciating a certain distance between foresight and the real pose vector. Optimization is written over each joint, it is supposed to be noted this can be used even for those joints in the image which are not labeled.

Thus, the estimation of the joint is done examining the full image and it totally depends on context but, there is insufficient capacity in the network due to the fixed input and it is not easy to improve the size of input as there are already many parameters. To come up with a solution for this problem, i.e., A cascade of pose regressors are proposed to train where the cascade estimates the pose initially at the first stage, to foretell a joint location's movement from the last stage to the real location, extra DNN regressors were trained and by this learning of the current pose can be evaluated by each consequent stage where these stages used predicted joint location for focussing on images relevant parts. Image’s sub-images are cropped out and by this, the resolution of images gets higher and this is used in every stage of cascade. Thus, this paper discusses the cascade-based refinement effects.

The datasets used were FLIC, Dataset of Leeds sports where the diameter (shoulder to hip distance) was defined for a pose and these datasets were arranged by imitating the human body in a tree kinematically. PCP measures were used to detect limbs rate but this technique penalizes short limbs so, different PDJ was used in detection criteria as this metric allows us in varying the distance( prediction to ground truth ) threshold.

Future work: novel architectures will be examined which could be extremely tailored potentially towards the problem of localization and in the estimation of a pose.

**Robust Object Recognition with Cortex-like Mechanism**

The understanding between object recognition using neural networks and the biology involved in complex optical and visual scenes captured by the cortex in our brain is what the authors of this paper have worked on. Their goal is to create an object recognition model for the complex scenes captured by our visual cortex which outputs great classification results of the images on the basis of shape and texture. So, to accomplish this difficult task, they have decided to use hierarchical architecture as this has provided significantly great results in flat (single-template) object recognition models like for example recognizing facial features or extracting vehicle features.

Explaining the model in a simple way, it has four layers. It has two types of units: *simple* ‘S’ and *complex* ‘C’. S units along with their inputs are combined to a bell-shaped function to improve selectivity. All the C units together are passed through a MAX (maximum) function which increases the evenness. This model tells us that the features captured from the middle and higher layers must be trained from visual experience so, then authors expand this architecture by telling the model how to extract visual features from the complex images of real-world examples for object recognition. The way that the performance of this Standard Model Features (SMFs) are evaluated in two different techniques: 1) *Shape-based feature extraction,* and 2) *Texture-based feature extraction* these are further categorized. The shape-based feature extraction is again classified into two i.e., Object Recognition with clutter and Object Recognition without clutter.

In object recognition with clutter, the final image for both training and testing datasets is shown at random positions and scales also the image is unsegmented. For, this type of image features, scale-invariant and position-invariant SMFs are used. In object recognition without clutter, the input image which is chosen is extracted into large number of image grids for classification of the output image. In this method the randomness in positions and scales of the image is limited. Texture-based feature extraction mainly focuses on the background image features like trees, sky and roads. For this technique, each and every pixel of the image containing the target element/subject is extracted in order to distinguish between the background texture and the main target object.

On reviewing the results of these two feature based image extractions, the authors have concluded that in shape-based extraction the model is divided into images grids/windows which detects target elements in the image (like, bicycles, cars, pedestrians) which are nottexture elements. In object recognition with clutter, the model uses scale-invariance to detect the target element in the image instead of checking from edge-to-edge of the input image whereas in object recognition without clutter it learns the position of the target element from the training data produced by the model. Then the model wherein each and every pixel is scanned to extract the texture from different background subjects like trees, roads and sky. Various textures were sorted out by checking the boundary where the pixels overlap one another is where one texture is differentiated from another. In addition to this, the authors have also depicted a brilliant visual demonstration of these two feature extraction models and also comparing their performance through a data flow diagram.

Finally, the authors have also mentioned when and where to use these SMF’s. When the images used for training dataset are unsegmented, we use object recognition with clutter as the target image is embedded somewhere in the clutter and also there is randomness in positions and scales. When the images for the training data are normalized and clutter-free we can use object recognition without clutter. Where the position and scales are limited and also clutter is non-existent. When the images lack proper features and when the data is not reliable we can use texture-based feature extraction as it extracts each and every pixel to distinguish between background textures and foreground subject.

**Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials**

The authors of this paper implement a highly efficient technique for multi-level image segmentation and labelling using conditional random fields (CRFs) which improves the performance of image segmentation and accuracy of labelling. There are some very basic CRF models like Unary classifiers where the edges in output image are often smoothened out, which leaves us with lack of information on different textures present in the image. In this paper the authors use a *fully connected* CRF which is capable of performing pair-wise potentials on each and every pair pixels present in the image. But even the fully connected CRF models have a restriction of limited datasets up to few hundred image regions or even less. In addition to this, the model used by the authors connects all the pixel pairs of the image without having any limits or restrictions but the only challenge for this model is the ‘size of the image’ which means, even if the image is lower quality if the size of the image is large the model has to cover and segment more number of pixels which in turn decreases the efficiency.

As this model is a fully connected CRF, the pair-wise edge potentials can be defined by a linear combination of Gaussian kernels. This algorithm mainly focuses on an approximation of mean field to CRF distribution. Which means that the approximation is done in such a way that it optimizes all the message passing steps by updating a single variable in each step which again approximates from the rest of the variables. This is shown in a fully connected CRF using Gaussian filtering. Doing this reduces the complexity of performing message passing of variables from quadratic to linear number. The output gained from this approximation is the number of edges present in this model. In the paper it is shown that this model outperforms many other CRF models using the MSRC-21 dataset including the MCMC inference inn fully connected CRF which took 36 hours to run wherein this model produced the output in 0.2 seconds. This model may look simple but gets weak while labelling nearby pixels of different class as the model detection is a bit poor in classification of different objects. For example, when the model is labelling nearby pixels it cannot tell the difference between “sky” and a “bird” to the same as “sky and “cat”. This issue was solved by the authors by training the model with a symmetric compatibility function which includes interactions with labels.

Along with some improvements and modifications to the fully connected CRF model, the authors of this paper implement many techniques from other models proposed by different authors. They train their model using two datasets which are: MSRC-21 dataset and PASCAL VOC 2010 dataset. The dataset is split into training, validation and test data in 45%, 10% and 45% respectively. But in the case of PASCAL VOC 2010 dataset due to lack of ground truth of test data, the give data is partitioned in the same standard splitting of training, validation and test data as mentioned above. The training time for the MSRC-21 dataset took around 40 minutes but the ground truth values provided by this data was not precise. Which means that, some of the ground truth images were not labelled properly, so in order to tackle this problems the authors have manually labelled the improper images of the dataset where labelling each image took almost 30 mins. This model has average classification accuracy of 29.1% and the accuracy when the model learned the labelling from the compatibility function as discussed above is 30.2%. The time taken by this model was roughly around 2.5 hours.

Finally, the authors of this paper have presented a highly efficient fully connected CRF model and also stated that denser the interconnections of pixels more accurate classification performance.

**CNN-based Segmentation of Medical Imaging Data**

The authors of this paper perform image segmentation using three-dimensional Convolutional Neural Networks (CNN) on medical images of hand and brain MRI. Most of the CNN models are trained over two-dimensional images while here they use three-dimensional CNN model for segmentation of bones of the hand and the nervous system of the brain. They achieve this by constructing a CNN modelsimilar to “U-Net” architecture of Ronneberger et al. by including two modifications. They are mixing multiple segmentation maps together created at different layers and summation of each element of the map from one stage of the network to another.

The huge memory of three-dimensional medical images of hand and brain MRI’s are down sampled instead of dividing the 3D image and then up sampled to maintain the original size of the image. So, when the image is down sampled the network has to process the identical patches nearby in order to classify the nearby pixels. While up sampling the image alongside this feature maps skip connections from one stage of the network to the next stage which means image extraction has twice the number of paths compared to segmentation. When we compare this network design architecture to other previously made architectures like U-Net, 3D U-Net and V-Net, they are all based on patch-based training whereas this architecture is pixel-based training. In patch-based training, a three-dimensional ground truth is created for medical images to train the model which require many image samples and a lot of time. But, sample datasets for medical images are very scarce and difficult to get while all the above mentioned patch-based architectures reply on huge chunks of datasets for extensive training. These sample datasets are used for concatenation but in pixel-based training they are used in summation which helps the overall performance of the network architecture.

When switching from patch-based training to pixel-based the most important point to be noted is the loss function. Class imbalance is a major issue while switching from patch-based to pixel-based so a loss function much similar to dice similarity is used for training. It is a metric which is commonly used for checking thesegmentation maps quality. One limitation for this loss function is that some of the MRI samples does not contain any foreground data which maximizes the loss, so samples with no foreground data produce poor outputs.

The four main labels from the ground truth are (1) ‘metacarpal phalanx’, (2) ‘proximal phalanx’, (3) ‘middle phalanx’, (4) ‘distal phalanx’ in our output they are represented as green, yellow, orange and red colors respectively. Getting these parameters accurate when compared to the ground truth after training the architecture is important. The results obtained are compared on the basis of 3 major factors, (1) “whole”, (2) “core” and (3) “active”. Segmentation of healthy volumetric pixels fromimproper ones is whole, segmentation of core tumor from healthy and improper regions on edema and healthy brain tissue is called core and finally, active is segmentation of healthy class from the remaining regions. When same output is compared between two different experts the accuracy of the above three mentioned parameters are 85%, 75% and 74% respectively. But when the output of one expert is compared with several other experts the outcome is ever better with accuracies of 91%, 86% and 85% respectively.

The results suggest that long skip connections of feature maps which was discussed earlier is an important part of the network removing them would worsen the performance of the network. In conclusion, from the two modifications of U-Net architecture as discussed above summationreplicated better results and when multiple segmentation maps has combined it has sped up the convergence without affecting the final performance and Jaccard similarity index was used as loss function. The issue of huge memory size of the 3D medical images was resolved by down sampling the images when required rather than dividing. We can conclude that despite the lack of samples of proper ground truths of medical image samples, three-dimensional CNN network architectures proved to yield good quality results.

**CheXNet: Radiologist-Level Pnumeonia Detection on Chest X-Ray with Deep Learning**

The purpose of the study was to develop the Pneumonia automated algorithm. This is a 121 layer-convolutional neural net work model, named as CheX Net. The model was trained on largest dataset containing 100000 X-Ray images of front view of chest with 14 different diseases .The biggest challenge for the model was that the result could be misunderstood by other benign diseases. Hence, the performance of the model was compared with 4 practicing academic radiologist and found it was working like expert level radiologists. The model was also extended to detect all 14 diseases. The automated detection of diseases will play a benefitting role in healthcare and diagnosis of diseases as large population will be able to access such imaging specialist.

The model is a binary classification problem. Since it takes it input as a chest X-Ray image And outputs a binary label showing probability of Pneumonia which belongs to {0,1} only . It tells whether Pneumonia is present or not. Optimization was also done using a weighted binary cross entropy loss.

CheX Net is a Dense Network which improves flow of information and gradient though the network. The network was trained using standard parameters and in small batch of 16. The learning rate to train the data kept 0.001,i.e. small to avoid misbehavior of the model. The dataset used here is chest X-Ray 14 released by Wang et al (2017) which contains 112120 Chest Ray images of 30805 unique patients. Images form dataset that have Pneumonia were labeled positive and other images were labeled negative. The dataset was then randomly splitted into trained (70%),validated (10%) and tested (20%) dataset, using which model is trained, validated and the tested respectively . No overlapping patients were taken between splits. Before putting the images into the network model , images form Dataset were downscaled to 224x224 and also normalized using mean and standard deviation Method .For testing dataset, the 420 frontal view of Chest X-Rays were obtained form 4 practicing radiologists at Stanford University with different level of experiences. After training , the performance of trained model and radiologists were compared by taking mean of F1 scores of all radiologists that was produced individually . The difference in F1 scores of CheXNet(0.435) sand average of radiologists (0.387) was significantly higher.

However, the limitations of the comparison included lateral view of X-Ray also for 15%of accurate diagnosis and only frontal view of images were not enough to diagnose in many cases. Also there was no permission to use patient history. Some modifications to the model were done in order to classify all 14 diseases by making 3 changes. The model output was extended to vector t of binary labels rather than 1 binary label for presence or absence of all diseases. The network layer was replaced by 14 dimensional output rather than single output. Also, the probability of presence of each disease class became the final output instead of presence of Pneumonia only. At last, optimization loss function was also modified as per the output. The model was added with Heat maps to visualize the affected area which indicated the disease using Class Activation Mappings (CAMs).

The research has gained lot of attention due to advancements in deep learning and artificial intelligence which has increased the performance in the field of medical imaging tasks, including skin cancer classification, diabetic retinopathy, etc. The large amount of population lack access to radiology diagnostics according to W.H.O. The availability of equipments even not able to help as access to medical imaging facilities and skilled radiologists is limited. The conclusion is that the development of expert level automated algorithm solves this problem and early diagnostic treatment will be possible which leads to reduced death rate and improving healthcare.

**U-NET: Convolutional Networks for Biomedical Image Segmentation**

The author of this study has proposed a new method to train a machine learning algorithm for recognizing and segmenting different biomedical microscopic cells with a help of very small training images or more efficiently with an available data set. In the bio medical the main problem very less and the algorithms that are currently present are very time consuming, this problem is solved by U-Net.

This paper introduces a new network and a training method that depends on data augmentation. First, the segmentation maps are generated for each image which defines the border of the microscopic cells. The architecture consists of a contracting path to capture context and a symmetric expansive expansive path that enables precise localization. In the architecture, the input image is contacted to a very small patch while applying convolution 3 times along the breadth followed by max-pooling along the height then the patch is expanded by up-convolution along the height followed by convolution along the breadth respectively. Small segmentation maps are generated to predict border pixels of a particular segment of a large arbitrary input image and the missing input data of the segment is interpolated by mirroring. The pre –computed segmentation map helps the algorithm to learn the small separation borders between the touching cells, this separation is computed with morphological operations. The training data is a set of 30 images from serial transmission electron microscopy. Data augmentation is very important for training the network for the required invariance’s and robust properties. Random elastic deformations are the main feature to train the segmentation network with a very small number of annotated images.

The performance of U-Net architecture on various biomedical segmentation applications is ground breaking. The resultant Wrapping Error, Rand Error, Pixel Error by U-Net is the lowest compared to other segmentation algorithms. It uses a very less number of annotated images and very less training time of 10-hours with an NVidia Titan GPU of 6GB.

Direct and early detection of Mycobacterium tuberculosis complex and rifampicin resistance from sputum smears

The rise of extensively drug-resistant (XDR) and multidrug-resistant (MDR) tuberclosis (TB), Tb is an widely emerging difficulty especially in India with the maximum number of cases between 2 to

2.5 million as reported in 2010. The detection of Rifampicin (RMP) susceptibility may process to timely pertinent treatment. One of the important factors in fighting TB is to find positive patients who are responsible in community spreading of the disease. Reports conclude that when mutation occurs in codons 509–533 of the 81-base pair (bp) RMP resistance-determining region of the rpoB gene we observe a 96% of RMP resistance. To screen the detection of resistance by RMP there are several techniques implied on the positive slides, to avoid the long waiting times.

Sterile screw-capped packages were used to accumulate the positive sputum specimens which were then stamped with the patient’s identity. Further the mucopurulent sputum was disseminated onto fresh and sterilized glass slides. The samples were then indexed on the basis of guidelines assessed by RNTCP. After a period of 8 weeks of preparation, a positive and negative control slide were taken and processed for ZN staining. One another slide was also used which was pacified by using 5% phenol in ethanol for 5 min to test for AFB positive samples. The arrangement of lysos and mycobacterial cell suspension was assessed. The whole procedure was processed as per the standard followed by the Drug Susceptibility Testing was executed out implementing the proportion approach, as mentioned in the Revised National Tuberculosis Control Programme or RNTCP guidelines.

The research evaluated the practicalibility of recognizing M. tuberculosis complex and RMP susceptibility from 3+, 2+ and 1+ index positive sputum slides. The INNO-LiPA kit for the cytological detection of RMP resistance and M. tuberculosis complex was documented by Wallis to have less sensitivity in laboratory specimen and more efficiency when implemented on new culture accumulated DNA samples, 12 in the current research, the increased performance of 88.6% after 2nd tour of ethanol distillation of DNA to discharge the PCR inhibitors present in the laboratory specimens arouses an advanced beginning for futuredevelopment.

It can be seen from various researchs conducted till now that LiPA generates crucial outcomes with DNA produced from the recent perception. The current research was able to accomplish considerable conclusions, from DNA excerpt to hybridisation in a tenure of 1–2 days. On hybridisation, 31 intensified goods proccesed a positive reactive probe for M. tuberculosis complex. A similarity of 96.8% (6 RMP-susceptible and 24 RMP-resistant) and dissimilarity of remaining 3.2% was achieved. A probable interpretation for the failure to recognize RMP resistance using LiPA kit in a single sample could be the happening of a mutation exterior to the 81-bp segment of the rpoB gene.

Thus it can be concluded that the aforementioned approach will be beneficial for the conveyance of specimens from the tangential health centres, there the only necessity is to process an pacified sputum slide, even without the requirement of transit sputum in cetylpyridinium-chloride, which eventually decrease the danger of both laboratory infection in between the specialists as well as that of sample biohazard amid transit to the remarked laboratory. This technique can be advantageous in the cytological means of M. tuberculosis complex and early prognosis of multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis.

**Automated object and image level classification of TB images using support vector neural network classifier**

Tuberculosis or TB is an extensively emerging problem especially in India with the maximum number of cases between 2 to 2.5 million cases reported in 2010. It is an airborne infection; Tuberculosis mainly affects lungs along with lymphatics, pleura, bones and joints or meninges causing pulmonary TB. The proposed methodology presents diagnosis of TB on the basis of a multi layer neural network architecture activated by SVD algorithm to enhance its generalization potential. The integrity of the proposed study is its rare pipe-line flow including the recognition and categorization at both image and object grade of the sputum smear images. A shape like rod, which presents straight curved or bend shape in the sputum smear images represent bacilli. Therefore it is crucial for the feature extraction and categorization plan. In the proposed research, Fourier descriptors (FD) are depicted like the bacilli which are probably unalterable against scaling, translation and rotation.

Samples were taken for both positive and negative cases, which were spreaded onto the slides which are air dried before examining under the fluorescence microscope. The camera which was used to capture images was attached in monochrome binning mode. Various techniques and calculations were employed to have an optimized and reliable solution including image acquisition, chan-vese model formulation, contour-based feature segmentation; entropy-based feature selection followed by categorization which isrelied on State Vector Machine training for MLPNN.

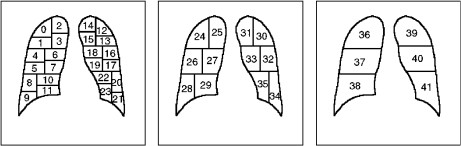
The obtained pre-processed images are regulated for active boundary-based extraction approach to recognize the existence of bacilli and outliers. MLPs trained using Support Vector Neural Network (SVNN) are expected to have improved rationalization potential due to the fact that SVNN is based on the Maximum Margin (MM) principle. The margin is increased which restricts all the training data from the hyper plane as possible thus decreasing the chances of wrong allocation. The sputum smear images considered containing outliers, which can be due to non-specific or poor staining of the smear slides or due to the coinciding bacilli. Uniform morphology is not observed in outliers. Results displays that the extraction model detects the bacilli which maintain their shape Fourier descriptors irrespective of amount specified in the images. Scanty or no bacilli are observed in negative images. The proposed system adopted important Fuzzy entropy measures which are appraised with the use of SVNN. An appreciable accuracy of 91.3% for object segmentation and 92.5% for image level classification was achieved rather than using Back Propagation NN.

From the proposed research it can be concluded that images took using a low cost Fluorescence microscope have been used in both image and object level recognition and categorization, which are documented at MRC/UCT Medical Imaging Research Unit, University of Cape Town, South Africa. The active contour extraction approach is based on level set method and Mumford-Shah technique which are used in enhancing and sharping the image’s features. The results state that the technique sharpens the pattern-based features required for the outline characterization of every entity in these images. For the additional segmentation of the TB objects, the significant contour controlled Fourier descriptor features are abstracted from images. It is noticed that the most related features preferred are relied on vague entropy measures which enhance the categorization accuracy effectively in both image and object level. Also, the chosen multi-layer neural network architecture to execute regular image level and irregular object level categorization is proved to be correct. Hence, because of the higher sensitivity and efficiency, this study can be effectively used for mass screening of pulmonary tuberculosis.

**Automatic Detection of Abnormalities in Chest Radiographs Using Local Texture Analysis**

Tuberculosis or TB is an extensively emerging problem especially in India with the maximum number of cases between 2 to 2.5 million cases reported in 2010. In order to fight against this epidemic, mass chest screening can be done to detect the number of active cases present. It has been observed in different radiographic patterns of TB patients that there are diffused abnormalities in different areas of lungs. Computer-aided diagnosis for mass chest screening for TB identification is the new era that has not been researched before. The proposed methodology displays a standalone approach for texture analysis of lung area produced in chest radiographs. Various state-of-the are approaches are used to interstitial disease. Regions of Interest are segmented which are further processed to extract the texture features based on Fourier Spectrum along with geometric features, pixel profiles, fractal dimension. The approach is based on analysis of texture in local segments in the image, which are extracted autonomously. Multi-layer neural network architecture was designed to accomplish the task.

For this research, 241 normal cases along with 147 abnormal cases along with texture abnormalities form the TB database. The ID database comprises of 100 each normal and abnormal posterior-anterior chest radiographs. Segmentation of different regions within the lung fields was achieved using the Active Shape Model (ASM) technique. To

analyze the minute regions in lung fields, it was divided into smaller segments. Initially it was divided into three parts (sliced horizontally), then each part was divided into two, and further divided again into two sections; making a total of 42 different segments (*as seen in figure*

*1*). A criteria was adapted to test *Fig. 1: Image showing the 42 segments in lung fields.* Whether a segment is abnormal or not: If there is any coincident part between a segment and any abnormal area as drawn by the radiologist, the segment is termed to be abnormal. For the texture analysis module, a multiscale bank was used to obtain features. A texture feature vector has been produced for all segments in the image. These characteristics are segmented from histograms of the reactions of a multiscale filter bank. Every segment is computed individually with a distinct neural network classifier.

A complete abnormality score for the overall image can be calculated by assembling such pieces of evidence. As seen in the figure 2, the black part represents the normal part and the white part shows the abnormal part in the lung field. In order to test the model, experiments were executed on a data base possessing radiographs of chest

with interstitial disease with exceptional results. The results were found to be accurate when experimented on a huge database with images of textural abnormalities, containing many subtle cases. Even though not classes of abnormal texture analysis from this mass TB

Screening were tried in the test, the analysis *Fig. 2: Segmented parts as normal (black) and* Proclaim that this approach can be used by *abnormal (white).*

Radiologists for interpreting mass chest screening images rather doing manually. For the chosen database, results were 97% sensitive at a specificity of 90%.

**KNN Model-Based Approach in Classification**

The author of this study has proposed a new classification algorithm to classify a given data set into different classes. This classification algorithm is an improved version of KNN classification and deals with the drawbacks that arise while using KNN classification. The KNN classification is effective but has many drawbacks, for example, the extremely bad efficiency of the algorithm makes it impossible for data scientists to use it in dynamic web mining for a huge data warehouse and many other applications that require classification of a new data instance dynamically, similarly, the dependency on the value of the variable K makes it very hard to find a K value which can make the algorithm most efficient. These problems are resolved by the KNN Model classification.

This paper introduces a new approach to KNN classification by using models to classify new data instances, to make the algorithm less dependent on the variable K, one of the authors proposed to check various sets of nearest neighbors in place of checking only one set of K nearest neighbor, this will help the new data instance to find the true class for itself by aggregating the support values of various sets of nearest variables. This method will be less dependent on k and the performance of the classification will be very near to the best value of K. The performance of this algorithm for adding a new data instance is O(n²), it is relatively slow but the dependency on the variable K is exceptionally less and comes with a performance as good as the one using the best value of K.

Fusion of Local and Global Detection Systems to Detect Tuberculosis in Chest Radiographs

Tuberculosis or TB is an extensively emerging problem especially in India with the maximum number of cases between 2 to 2.5 million cases reported in 2010. In order to fight against this epidemic, mass chest screening can be done to detect the number of active cases present. The proposed research work presents an automatic TB detection model by merging different subsystems together in digital chest radiographs. The methodology is based on merging detection systems of a lung field, clavicle, and texture and structure analysis all together to enhance the overall performance of model.

The inclusion of the structure based abnormality detection model displays the capability of merging various systems at distinct levels i.e. pixels and image,

Which increase the overall performance of the model. The figure shows the combination of detection systems according to image and pixel level. The segmentation of lung field can be made more versatile so that it outputs accurate segmentations with the presence of pathology. Normal undismayed anatomy already presents important alternative and this situation is even more

when pathology is present. As compare to a system trained on a normal anatomy, the proposed system took much larger amount of data for training.

The serial and parallel configurations of this model were experimented by applying a few probability combination rules. However, when this system is applied in practical scenario, other configurations can be used on the basis of computational and other resource constraints. Instead of using several different individual systems for the detection of disease in images, it is possible to design standalone software to accomplish the task. The generalization potential of such a categorization model decreases when it becomes more intricate (individual system), which then requires greater quantity of data for training to attain better performance. A practical explanation to use several different sub models, such as the clavicle and texture detection system, is that they can be designed parallely by various other researchers.

Inside a middleware for detection system combination such as displayed in this study, other modules can be easily adjusted. In the operation for detection of TB in CXRs, detection of focal lesions and hilar abnormalities could further enhance the performance of the complete system. It was displayed in this research that the consolidation of various detection modules enhances the detection of TB on CXRs. The efficiency of Tuberculosis detection as being calculated using the area under the ROC curve is found to be 67% for the texture abnormality detection system alone. Further, it was also noted that the efficiency was increased upto 86% when three of the systems were integrated together. The optimized result was accomplished using the sum and product rule in a parallel combination of outputs.

Detecting Tuberculosis in Radiographs Using Combined Lung Masks

Tuberculosis or TB is an extensively emerging problem especially in India with the maximum number of cases between 2 to 2.5 million cases reported in 2010. In order to fight against this epidemic, mass chest screening can be done to detect the number of active cases present. The conventional techniques available for screening are time consuming and expensive. The proposed methodology presents an automated system that screens chest x-rays for instances of diseases like TB or other lung problems. For our experiment, two data sets of frontal chest X-rays were used: Japanese Society of Radiological Technology (JSRT) which consists of 247 chest x-rays (154 abnormal x-rays and 93 normal x-rays). Also, we used dataset from Department of Health and Human Services, Montgomery County, Maryland (MC) which consists of a total of 138 x-rays (comprising of 80 normal x-rays and 58 abnormal x-rays). All the images used for dataset are 12-bit gray scale images.

Initially, we provide an input of chest x-ray which is segmented with respect to the lung fields. The task is achieved by using a mixture of a statistical lung model mask, an intensity maskand a Log Gabor mask. The Log Gabor Filter is defined as:

*G(w) = e-(log(w/wo))/ 2log(k/wo)*

Where filter’s center frequency is denoted by *wo*.

The figure shows the energy output of the Log Gabore filtering and our final Log Gabor mask. The resultant mask for lung segmentation is the average of the three masks: the lung model mask, intensity mask and Log Gabor mask.

From the segmented lung field, we then obtain a set

of features for textures, curvature and shapes. We use these extracted features for training a support vector machine (SVM) that differentiates between normal and abnormal x-rays. We experimented our trained model on conventional data from a TB control program and got a AUC of 83.12%. The performance achieved is commensurable with other systems published till now from other such works.

A remark of our proposed method is that to achieve efficient lung segmentation we merged various masks. It is notable that our mask can be applied on lung model trained on a Japanese data set to a very disparate data set of MC. The segmentations were not measured by comparing them pixel-wise with ground-truth segmentation, as done in most other studies so far. We measured the performance of our model on our segmentations directly on the MC data set, which according to us is more realistic.

The future work of our automatic model is to enhance the performance more and to upgrade the recall of our screening system which includes optimizing the lung segmentation and using additional data that will be collected on-site in Kenya. A dynamic lung model alignment can be employed instead of a static one, or even additional segmentation masks can be combined. Furthermore, the extracted lung field can be partitioned into regions and local features be used for classification.

**Combination of texture and shape features to detect pulmonary abnormalities in digital chestx-rays.**

The author of this study has proposed a new classification algorithm to classify a given data set of digital Chest X-Rays into two classes, with Tuberculosis or without Tuberculosis. The authors of the paper, with the help of two radiologist experts, have developed an algorithm for automating the screening of digital Chest C-Rays to diagnose whether the patient has tuberculosis/pneumonia or not. In the real world, it is extremely difficult to diagnose if a patient has TB or not and even the World Health Organization has published that Tuberculosis is the deadliest disease which is highly contagious. As this study will be implemented to solve a real-world problem, it was very important for the data scientists to get a result which will be approximately 100% accurate. When a patient is suffering from TB or Pneumonia due to these pleural abnormalities and pulmonary abnormalities are manifested in the lungs. These deformities can be seen in a Chest X-Ray. For a healthy patient, outlines of the anatomic lung boundary and the air cavity of the lung can be seen overlapped on each other. For a patient suffering from this disease, these outlines do not overlap. This algorithm uses this technique to classify the input Chest X-Ray into Normal or Abnormal.

This paper introduces a new screening system of digital Chest X-Rays which depends on the shape features and texture features of the lungs to diagnose a patient for Tuberculosis. This algorithm has two classifiers, which the algorithm uses in a sequential form, the first classifier is given shape features as input, if the first classifier classifies the input data instance as "Abnormal", then the algorithm is terminated at classifier one itself and the input data instance is labeled "Tuberculosis case". If the first classifier classifies the input data instance as "Normal", the algorithm sequentially goes to the second classifier, now the second classifier is given the texture features of the lung as input and it further classifies the input data instance into "Abnormal" or "Normal". The Architecture of this algorithm uses two types of segmentation, first Lung Anatomy Segmentation, and second Air Cavity Segmentation of the lung. The first segmentation helps the algorithm to find shape features and some of the texture features of the input Chest X-ray, the shape features were very basic geometric characteristics of the lung, for example, bounding box coordinates, centroid coordinates, extent, eccentricity, orientation, size. While the second segmentation Air Cavity Segmentation, help the algorithm to find many texture features of the lung and these texture feature of the lung consists of local binary patterns, the histogram of oriented gradients, gradient magnitude histograms and intensity histograms.

The sequential use of two classifiers helps the algorithm to increase its performance by 16% from its previous performance and as rigid registration is used for the segmentation methodology, it further increases the performance of the algorithm. We can confidently say that the performance of this algorithm is very close to 100%.

**Auto-matic detection of pleural effusion in chest radiographs**

Tuberculosis (TB) is a deadly disease having affected around 9 million people world-wide and causing around 1.5 million deaths in 2013 alone according to World Health Organization (WHO) 2014. However, TB is curable if proper diagnosis is done at an early stage. This becomes a massive problem in countries with low-resources where individuals are screened using Chest Radiography (CXR) initially, due to its wide availability and cost-effectiveness. Subjects with test-positive are made to undergo several, other time-taking and costly examinations for confirmation.

Reading CXR’s for signs of active TB requires medical staffs that are trained well which is limited especially, in countries with high TB rates. Moreover, human-errors start creeping in when number of test-cases increase. This is when Computer-Aided Detection (CAD) systems come into play. CAD systems are powerful enough to make decisions even in absence of medical personnel, thereby making the process extremely simple.

TB primarily affects the parenchyma cells of the lungs and may also cause consolidations and cavitations. Most of the Computer-Aided Detection systems therefore analyse the lung fields for presence of the slightest abnormality in parenchyma cells, however some do focus on manifestations of cavities. Another common form of TB is pleural TB which occurs due to accumulation of extra-pulmonary fluid in the pleural space affecting the hilum and pleura.

Sadly, current CAD systems are not particularly focussed in training to analyse this CP deferment and therefore they are unable to detect the disease if Pulmonary Effusion (PE) is the only abnormality in CXR.

Since, very less research has been done on automated detection of PE, researchers are coming up with different methods to address the issue.

One such method uses CP angle (angle between hemi diaphragm and lateral chest) to identify the presence of PE. If fluid is present, it causes bluntness of the CP angle with respect to normal case and this difference between the two angles can be used as an important feature in PE detection.

* First, two anatomical structures (the Lung Fields and Chest Walls) are detected. These are required to locate the CP point with acute accuracy.
* Next, we go on to detect the CP point. This is done in two steps:

1. Coarse CP point detection
2. Fine CP point detection

* After this, a ROI around CP point is extracted and three feature descriptors namely Angle Measure, Intensity Histogram and Morphology are measured to construct a feature vector.
* Finally, a Supervised Classification methodology is employed to assign the PE severity score.

Based on the following three parameters, the performance of the proposed PE detection model was analysed and compared in terms (AUC):

1. Angle Featurecomputedusing xcp,lung andBlung
2. Angle Feature computed using xcp and Blung.
3. PE Severity score

The proposed PE detection model shows promising results for both the hemi-thoraces. It basically uses the technique of lung segmentation and chest wall contour to detect the CP point. As an overall result, higher accuracy is obtained.

**An introduction to Convolution neural networks**

The research paper gives a basic understanding of the Convolution Neural Network (CNN). The recent work done in the field of CNN is described in a brief and simple way.

In the introductory part of the paper, the authors give an introduction of artificial neural network (ANN) and explain how CNN are different from ANN. The formation of ANN i.e. the layers involved in constructing an ANN and how they work are explained. An Artificial neural network consists of three layers, i.e. input layer, hidden layers and output layer. CNNs are designed to solve image recognition problems and are less complex than ANN .It also explains that if the ANN is made more complex, it could lead to over fitting which in turn degrades the performance of the model.

In the second part, CNN architecture is described briefly.CNN consist of four layers, i.e. input layer, convolution layer, pooling layer and fully connected layer .An activation function called Rectified unit (Relu) is embedded in the layers to provide activation to the output produced by the previous layer. The input neurons of a full image are difficult to handle, so convolution layer applies filters to the whole image and divide the image into small portion. It gets connected to a portion of an image at a time. This is done by taking into consideration the hyper parameters I.e. depth, stride and zero padding. The output given by convolution layer is called an activation map obtained from a potion of image. This activation map is given as input to the pooling layer, which performs some down sampling, i.e. it selects the activation map with the maximum value from all the activation maps generated. The fully connected layer is the final output layer similar to ANN, which operates on the inputs of activation map, selected by pooling layer and convert the inputs to corresponding neurons as output.

In the third part, it is explained that how the basic configuration of CNN can be changed and used to reduce the complexity of the neural network.

The fourth part of the paper concludes the topic saying that a brief understanding of CNN would help in doing further research development in the field.

**Dropout: Preventing Neural Networks from Overfitting**

The author of this study has proposed a new method of data dropout to prevent the machine learning algorithm from overfitting and help the algorithm to have higher generalizability. In complex deep neural networks, the main problem that arises for a data scientist is that, if an algorithm trains itself too much, it ends up training itself on noise data also, now, this prevents the algorithm to find the correct output for general solutions and ends up with a very small scope of generalizability. This reduces the performance of the algorithm exponentially, as the algorithm will be computing every small piece of information the input. An input, even with a slight deformation will be rejected, as the algorithm will be highly sensitive. Dropout resolves this issue by temporarily removing some of the units and their input and output neural nets out of the network. This helps the algorithm to maintain the generalizability and improves the performance of the algorithm. Dropout has a hyperparameter probability, for every unit in the network. While training the dataset, Dropout drops a random unit from the network with a probability (1-p), the retention of units is based on probability p, and this helps the algorithm to have perfect generalizability. The testing phase does not contain any removal of units from the network, which means all the important computing is done in the training phase, which in turn makes the training phase of the algorithm a little slow, but the performance of the algorithm on a general data set increases exponentially.

This paper has introduced the implementation of dropouts in a machine learning algorithm which depends on the way how a data scientist chooses a probability p on a data set of n units. As we know, the probability of retaining a unit is p, the total units become n\*p, if both, the probability, and the data set are too small, then the algorithm will not be able to perform well for the type of data set. Experimentally, it is shown that, if a probability p i is used in an algorithm then, ideally, the input units' size should be n/p. This will help the algorithm to retain n units for its use after the removal of various units from the network while training the data set. As the Dropout removes random units from the sample, this ends up removing some key units of the input and simultaneously retaining some of the noise units of the input, this in turn manifests some amount of noise in the gradients. To help the algorithm overcome this problem, the author suggests using 10-100 times the learning rate which works well for a simple neural network and also, by using a high momentum value. A large learning rate and momentum values can make the neural network weights to become very huge, to prevent this, the author suggests to use Max norm regularization for input weights of a unit.

The Generalizability of this algorithm for a set of input data is found to be extremely high. It uses different parameters that can be adjusted as per the size of the input to get the same level of high performance. The training of this algorithm is very time consuming but is worth the results which are performance and generalizability.

**Image Semantic Segmentation Using Deep Learning**

The purpose of this research is to elaborate image segmentation using deep learning. Image segmentation is widely used in different image processing techniques and detection of a particular feature like edges between a foreground object and the background scenery in the image. It helps in image recognition, detection and classification. It can be performed using deep learning. Image segmentation is done using a deep convolutional neural network model where the model is partitioned into several segments on the basis of similarity of pixels with the neighbouring pixels nearby. In deep neural networks, various parameters are trained iteratively to get the minimum error loss using the back-propagation method and hence are slow as they learn connections between every input like human brains do. Hence, hardware such as GPUs is also progressed to support deep learning algorithms for high-end computations. Image segmentation automation is in research and we are considering interactive segmentation to simplify the task.

The research proposes various models for image segmentation. The first one is FCNs (Fully Connected Layers) which is a modified convolutional neural network where fully connected layers or decision making layers at the end are replaced by a convolutional layer with big receptors. It predicts features very accurately and until the last layer as it supports various size images. It helps in generating a heat map and highlights the object in the image due to the kernels present. The drawback is that it requires high computational time. The second one is U-Net (an extension of fully connected layers). This is the extension of FCN. It enables known elements in the network to pass into higher-level high-resolution layers and in this process, the model takes a U shape. It reduces distortions giving increased resolution images but high training time is needed. The third model includes Atrous convolution (for extracting features deeply). It classifies objects in different levels of representations. It uses many successive pooling and stepping layers increasing the training time. It reduces resolution but leaves the heatmap. The efficiency of such a network depends on the dilation rate. The fourth one is SegNet that represents deep fully CNN for pixel-based segmentation. It comprises of three networks. An encoder network for convolution, decoder network for deconvolution, and a classification layer. It helps to map low-resolution images to input images precisely providing smooth segmentation. It conserves the boundary details of even small shapes. It is cheaper than other networks and has a high accuracy of training.

This research implementation a new efficient Mask R-CNN that is an extended and faster version of R-CNN. The main aim of the R-CNN method was to detect objects and define their edges. There are separate models to create the area, classify the image, and to optimize the area. The optimization is done using linear regression. It has to train each model separately and was slow. Therefore a fast version was created which uses a single model only with a softmax layer in place of SVM and in parallel linear regression layer is used that reduces a lot of computational time. In a faster version, a feature map is used to interpret the region proposals by using RPN. Mask R-CNN uses the idea of the faster version. It is done in 2 stages. In the first stage, region proposals are found using RPN and in the second stage, it predicts the class of each object and a binary mask which puts 1 wherever the object is present otherwise 0. It uses pixel-level recognition that results in misalignment. ROI alignment is done to solve the problem.

The study concludes that image segmentation can be done using deep convolutional neural networks which help to extract the details and features in an efficient way with minimum error. The best performing and efficient model is found to be R-CNN. It identifies the objects and segments the images on the basis of a given input object and the training dataset. But in the medical field and for large datasets, U-Net is preferred for imaging and is broadly used. Atrous deep architecture is used for real-life designs where high-resolution images are needed such as automated cars. FCNs make image multi-channels into segments and helps to distinguish the objects from their background. The CNNs performed best within time and space complexities.

**Speeding up Deep Neural Network Training by Decreasing Covariate Shift using Batch Normalization**

Training of the deep neural network using stochastic gradient descent (SGD) is hard due to the fact that the distribution of the inputs in each and every layer depends upon the parameters of previous layers. It requires very small learning rates and needs a lot of attention to initialize the parameters. It results in the slowing down of training. This phenomenon is also known as ‘Internal Covariate Shift’. To decrease the internal covariate shift and achieve the fixed distribution of the input, the layers can be whitened at every step. In the whitening method, the parameters are dependent upon the activation values of the network and the network requires activation after every parameter update in the entire training. The optimization steps such as gradient descent require normalization itself to be updated every time. So it becomes costly and is not differentiable everywhere. Hence, we need an alternative to whitening which does not require analysis of training at every step.

The research proposes a method of batch normalization to overcome this problem where the inputs of each layer are normalized in small training batches. The hypothesis is that normalization enhances the gradient backpropagation. The idea of the study is to normalize each feature independently by taking mean as 0 and variance as 1 as long as the elements of batches are tested from the same distribution. It uses the entire training set to normalize activations. The mean and variance of each activation are calculated by each batch in SGD. The batch normalization transform algorithm is discussed in this study. Each normalized activation is input to a sub-network that is comprised of linear transforms. Each Subnetwork training speeds up resulting in speed up of the entire network. For training, the Batch Normalization transform is applied to each element of the activation subset. During training, gradient loss from the transformation is backpropagated and follows a chain rule establishing a differentiable transformation where the model is learning from less covariate shift input distribution. Batch normalization standardizes the model because the training network doesn't develop deterministic values for particular training.

The experiments were done on the MNIST dataset. The comparison was also performed between batch-normalized and baseline networks where it is found stable. Batch normalization is also applied to the inception convolutional network that is trained on ImageNet classification. To take full advantage of batch normalization networks and their parameters are modified and experimented to speed up the BN networks at its full potential. The experiments were performed on both, Single network as well as an ensemble of 6 networks to get the better result.

The study concludes a new mechanism 'batch normalization' for speeding up the network by reducing the covariate shift from the internal network activations. It permits higher learning rates and needs less care for initializing the parameters. It requires 14x lesser training steps and performs better by speeding up the training. The goal of achieving a fixed distribution of activation values throughout training has been achieved. The normalization is done in small batches and gradients are back-propagated through the normalization parameters. Therefore, It also preserves the representation ability of the network. It guarantees that the normalization can be handled by any optimization technique which is used to train the network. Multiple models trained with batch normalization are combined to get better results on ImageNet. The future work may include finding all capabilities of batch normalization and applying it to recurrent neural networks where covariate shifts and gradient disappears harshly.

**Deep Residual Learning framework for Image Recognition**

Image recognition and accurate detection of the object has been the biggest challenge in the field of deep learning. The deep convolutional neural network combines different level features (low/mid/high) and classifiers in an end-to-end multilayer fashion, where each level can be seen as a number of stacked layers representing the depth. The highly in-depth representation of deep network can improve the image recognition tasks. Such deep representations are possible if deeper neural networks are trained properly which are hard to train. For example, in image recognition, VLAD is a representation which has been also discussed in the thesis thoroughly.

This research proposes a ‘residual learning’ framework for easy training of such networks that are deeper than those that are used recently. In residual learning, the layers are reformulated as learning residual functions with reference to the layer inputs, forming a residual network. The residual networks can be optimized easily and give accuracy because of deep layering. The thesis focused on the ImageNet dataset to evaluate the residual network with 152 layers that is 8 times deeper than a VGG network but still having the lower computation complexities. Residual networks achieve 3.57% error on this dataset. One more dataset, CIFAR-10 with 100 to 1000 layers was also analyzed.

The problem arises when a deeper network starts converging as the network depth increases. This convergence leads to degradation of accuracy of the training resulting in increased error. It shows that all systems cannot be optimized in a similar way. The hypothesis is that residual mapping is easier to optimize than the unreferenced original mapping. The problem can be solved by letting each layer fits a residual mapping, f(x) = H(x) -x, rather than fitting an underlying mapping, H(x). The original mapping is recast into F(x) +x which can be represented by feed-forward neural networks with shortcut connections. Shortcut connections skip one or more layers and perform identity mapping adding their outputs to the output of stacked layers. This connection does not increase any computational complexity. The entire network is trained end-to-end by stochastic gradient descent (SGD) with back-propagation. SGD is used in small batches of 256 and learning rate starts from 0.1.

The Experiments were done on the ImageNet 2012 dataset that consists of 1000 classes to show the degradation problem and evaluating the residual learning method. It was found that identity mapping was enough for labelling the degradation problem. The 18-layer and 34-layer of both, plain networks and residual networks were evaluated. In case of the plain network, the deeper 34-layer model has higher validation error than 18-layer model. While, in residual network the situation was reversed with a lower training error. It showed that deeper layer was better than shallower one.

The 152 layer-model was the deepest network presented on this dataset. It presented that deep residual networks are easy to optimize with lesser training error than their normal counterparts when depths or layers are increased. These networks accurately trained from the deep network. The group of residual networks achieves 3.57% error on ImageNet dataset. CIFAR-10 dataset was also analyzed that successfully trained the deep models with hundreds and thousands of layers. This research has won 1st place in the ILSVRVC 2015 classification competition, ImageNet detection, COCO detection, and other visual recognition fields.

**Search for Discovering Activation Function**

The activation function is responsible for the successful training of Deep Neural Networks and major job performance. All deep networks go through a linear transformation that follows an activation function f(.). The broadly used and successful activation function is the Rectified Linear Unit (ReLU), which is defined as f(x) = max(x, 0). The challenge in this area is to find an effective alternative for this function.

This research tried to support automatic search techniques that help to discover various novel activation functions. This was possible due to the combination of exhaustive and reinforcement learning-based search. The search was then verified and evaluated with the best newly discovered activation function “Swish” that performs better than the most successful activation function ReLU on deep learning models on demanding datasets. The Swish is a sigmoid function that is defined as f(x) = x. Sigmoid ( Bx ). It is simpler and easier to use than ReLU.

The methodology of the research includes the Search space that contains the activation function. Search space needs to be balanced in size as limited space will not be able to contain a new activation function and it will be difficult to search effectively in a very large space. The activation function is composed of multiple repetitions of the core unit that consists of 2 scalar inputs that are passed through a unary function which results in 2 unary outputs and the binary function combines both of them resulting in a scalar output. The aim of the search algorithm is to search for effective choices for the unary and binary functions. The prediction of the activation function component from each step of the RNN controller is given as feedback in the next step and repeat until every component is predicted. This results in a predicted string that is used to construct the candidate activation function. After the generation of candidate activation function by the search algorithm, a child network is trained with it and the validation accuracy is recorded to update the search algorithm each time. The dataset used is the ImageNet classification dataset, consisting of 1000 classes and 1.28 million training images. A list of the top-performing activation function is maintained. To maximize the validation accuracy, the RNN Controller is trained with reinforcement learning. Search is expensive with respect to time and computation. Hence, every child network is parallelized using a distributed training scheme.

However, the top-performing activation function works well for small networks but the discovered functions generalize to larger models. Hence, an exhaustive search is chosen for the small search space and RNN controller for the larger ones. Search founds that discovered activation functions utilize periodic functions such as sin and cos. Complicated functions did not perform well. Functions using addition and subtraction perform better while functions using division tend to perform poorly when denominator was near 0. Top-performing functions were also tested for the robustness using 3 models, ResNet-164 (RN), Wide ResNet 28-10 (WRN), and DenseNet 100 in the TensorFlow. The results are favorable but it is still undecided whether newly discovered activation function will be able to replace ReLU completely.

The research concludes that the performance of novel discovered activation function Swish matches or over-performed the standard on all tasks against other newly discovered functions and ReLU also. The results were aggregated by comparing performance to a variety of models and found improvement as statistically significant. Swish used a lot of different automated search techniques. The smoothness, simplicity, similarity, and robustness of this function replaced the ReLU in any network.

**Warm Restarts in Stochastic Gradient Descent**

In the field of deep learning, image recognition is of great importance which needs improvement in classification. This improvement can be done by training the model in the right way which is hard even on high performing GPUs. The model with millions of parameters is connected in a function that takes some input and produces some prediction as output. While training the model, predictions start to deviate. Hence, we try to update the randomly taken parameters to minimize the loss function. And this updation method is gradient descent. But in doing so we may be glued in some local minima. And where the stochastic gradient descent (SGD) comes into the picture that resets its learning rate every number of steps to come out of that problem early out. This improves the situation as we go through more iterations while going through the dataset. Hence, we need to increase the time it takes for decreasing the learning rate. This is what happens in a warm restart and increases the performance of the SGD.

Gradient-free optimization commonly uses restart to handle the multimodal functions by covering search space with dynamically allotted openings of local optimizers, but these only works for small search spaces. So, moving towards Gradient-based optimization improves the convergence rate in large scale searches instead of handling multimodality. This study uses warm restarts in SGD to increase its performance on the CIFAR-10, CIFAR-100, EEG (electroencephalographic recordings of brain activity for classification of the right as well as left hand and foot movements) and a subset of the ImageNet dataset. The research proposes to periodically mimic the warm restarts in SGD, wherein each restarts the learning rate is initially having some value and is organized to decrease. The gradient descent restarts whenever the angle between the momentum and the negative gradient is indistinct. Here, restarts require two to four times lesser epochs than the currently-used learning rate. It improves the results for every dataset.

The experiment was done by initially simulating the restarts by increasing the learning rate. The amount of this increase controlled to which extent the momentum is used. But the performance got worsed. They considered the Wide Residual Neural Networks (WRNS) with depth d and width k on the CIFAR-10 and CIFAR-100 datasets. The best results were obtained with a WRN-28-10 model with 28 layers (d) and 10 times more filters per layer than used in the original (k). For image preprocessing, global contrast normalization and ZCA whitening were performed. For data enlargement, horizontal flips were considered and images were cropped randomly and filled the missing pixels with the reflections of the original image. SGD is trained with Nesterov’s momentum with an initial learning rate of 0.1 which further dropped. The single model result achieved good performance by training 2 times broader WRN (WRN-28-20) next time. By making ensembles of the models test errors can be improved.

The conclusion of the research was that the warm restart in SGD accelerated the training of deep networks. It achieved competitive results on CIFAR datasets up to four times faster and also achieved new state-of-the-art results using broader WRNS and ensembles. The results on the subset of the ImageNet dataset recommends that SGDR can also lessen the problem of learning rate choice if restarts of SGDR assess a range of learning rates. The future scope should examine warm restarts for other popular training algorithms such as Adam and AdaDelta.

**Learning Numerous Layers of Features using Tiny Images**

This research proposes the training of a multilayer network model on images that are more natural. To train the model, classifiers have to be trained beforehand to classify the images that use a labelled subset of the CIFAR-10 dataset. And for the experiment purpose, the dataset was made by collecting the images using a web search with almost every English noun present in the WordNet lexical database. Lots of search engines were used and duplicates were removed. Also, the images with a large portion of white pixels were not considered. This dataset was named ‘tiny images’. Moreover, the experiment was completely performed using unsupervised learning.

The methodology focuses on two kinds of networks due to their capabilities of extracting meaningful features from the data. First is RBN (Restricted Boltzmann Machines), which is a graphical model with a visible and hidden set of nodes. Both sets of nodes are connected to each other. The model uses hidden nodes for efficient sampling, generalizing, and extracting of features. On the other hand, DBNs (Deep Belief Networks) train the layers in a greedy sequential manner. DBNs support to extend the RBMs to multiple hidden layers. Sampling the hidden units and then visible units and repeating it infinite times make it forgets the starting point. It can be done in a finite number of steps using a contrastive divergence (CD) learning process. Also, the real images have the property of a stronger correlation of pixels to the close pixels and weaker to the distant ones. Tiny image dataset exhibit the same. The hypothesis is that if the model chose the higher-order correlation over two-way correlations, it can find the interesting commonness in images. And this can be done by a data pre-processing step called the ZCA whitening transformation using the whitening filters. The transformation yields the edge information which assigns pixels of regions on relatively uniform color to 0. The RBM with binary values as visible inputs cannot model for real values. Hence, the Gaussian-Bernoulli RBM came into the picture.

During the training, the weights were arranged in 32 X 32 square grids and the filters were visualized. Intensity indicated the strength of the weight. Since the yellow region indicated a strong negative relation to blue pixels, it was considered that yellow is a negative-blue. For measuring the performance, a reconstruction error was used. The weights were learned by RBM to start the feed-forward networks which were trained with a back propagation algorithm.

The main motivation was to learn the multilayer model of images that have a strong relationship with a human visionary. Since neurons in human eyes can detect top-level features like edges. The RBMs with small units were trained on patches and merged it with newly trained RBMs with bigger units. This model behaves very well and successfully trained. The next goal was to train the RBM on un-whitened data. However, learning standard deviations has a negative consequence on filters even if it improves the reconstruction error. They also described the parallelization of the RBM training where more than one core per machine gave faster results but the price was little high due to updating weights twice on different machines which required little extra computation. The algorithm was implemented on C and Python. For machine communication, TCP sockets were used, and for core communication, shared memory rather than sockets.

Thesis implemented their algorithm on a large problem and the model was communicated between machines without slowing each other using parallelization. Double precision was slower but scales better than the single precision. Performances got double too when machines were doubled.

**ImageNet Classification with Deep Convolutional**

**Neural Networks**

The purpose of this research was to train a sufficiently large deep convolutional neural network to classify nearly 1.2 million variable resolution images in a contest named ImageNet LSVRC-2010. Classification is very helpful in object recognition and can be achieved using machine learning techniques. The main aim was to improve the performance of the visual recognition model and in order to do so, model must be trained with larger dataset to learn about that much objects from millions of images to be able to recognize accurately. The datasets such as ImageNet and LabelMe consists of lot of high resolution images in various categories of over 22,000.

Convolutional neural network model is one such large learning capacity model in which it’s capacity can be controlled by varying their depth and breadth. They are more accurate in making assumptions about nature of images like name of the image, locality of pixels, etc than feed-forward neural network models with same size of layering. CNNs are expensive to apply large number of high resolution images. Recent datasets like ImageNet contains enough labelled examples to train models without causing overfitting.

The implementation involves ImageNet as it’s dataset that is having 15 million labelled high resolution images with 22,000 categories. ILSVRC, an annual competition uses a subset of ImageNet with only 1000 images in each of 1000 categories. Totally, 1.2 million training images, 50,000 validation images and 1,50,000 testing images were there. The experiment is performed on ILSVRC 2010 version in which test dataset labels were available. ImageNet consists of variable resolution images. Hence, the images were down sampled to a fixed resolution of 256 X 256. The images were used without pre-processing on a raw RGB pixel values.

The architecture of the network was made up of eight learned layers with five convolutional and three fully connected. Deep CNN with ReLU trains the model a lot faster, but the primary concern of this model with ImageNet dataset was to prevent overfitting. Since, faster learning can highly affect the performance of large models on such large datasets. The ReLUs do not require input normalization to prevent saturation. The model was also trained on multiple GPUs. Since, single GTX 580 GPU with 3 GB memory limits the size of networks on which it can be trained. Hence, two GPUs were used in parallel. It takes slightly lesser time than single GPU. The overfitting was reduced in many ways. Adjacent pooling layers in CNN have their output of neighbouring group of neurons to not overlap to prevent overfitting. The easiest method to reduce overfitting is to extend dataset using transformations with labels preserved. Transformed images are coded in Python on CPU and they are not stored on the disk. Other technique for overfitting reduction was dropout technique which dropouts the neuron with probability 0.5. The model was trained using stochastic gradient descent. The batch size was kept as 128, the momentum as 0.9, weight decay of 0.0005 and the learning rate as 0.01 initially which reduces three times until termination.

The network achieved top-1 and top-5 test set achieved error rate of 37.5% and 17%. The result shows that a large deep CNN is able to break the record at ILSVRC-2010 competition which was 47.1% and 28.2%. The model uses purely supervised learning. The future scope can include large and deep CNNs on video sequences which will be able to cover helpful information that cannot be found in static images.

**Appropriateness of clinical severity classification of new WHO childhood pneumonia guidance**

In this paper, a huge and retrospective study of kids under sixty months who were admitted to Kenyan hospitals due to having pneumonia between 1st March 2014 and 29th February 2016 has been reported. In these age groups, the pneumonia case fatality rates (CFR) i.e. the rate of death in pneumonia have not been reported. In 14 Kenyan hospitals, the pneumonia cases of the kids reached about 16031. In that time, Kenya used both pneumococcal conjugate vaccine and Haemophilus Influenzae type b vaccine. About 832 deadly cases (5% of 16031) were informed. In this matter, the death rate is very high, but it is usual in sub-Saharan-Africa. Among those 832 kids, WHO would have classified at least 322 (39%) who has non-severe-pneumonia needing treatment at the home according to revision-2013. 5% of 16031 kids were died in spite of getting treatment by the hospitals, but the point is that the WHO Integrated Management of Childhood Illness-(IMCI)-revision would have mentioned them home-based increases many thoughtful questions. The review of those kids who were probably wrong classified as non-severe-pneumonia (by reason of their deadly result) exposed that the risk of death is related to pallor and malnutrition (commonly unrecognized). On the other hand, it may be also show several diagnostic-confusion among pneumonia and malaria, although the risk factor is even presented in the zone of non-malaria. Undoubtedly, recognition of acute-malnutrition is a vital issue. This group has a strong link between the lower chest wall in drawing and death. A huge size of sample is taken from several hospitals through the whole country caused in accurate estimations which are approximately represents the amount of the kids with pneumonia admitted in several hospitals in African sub-Saharan area. The superiority of the obtained result is visually established on the basis of mean (SD), median (IQR), and statistical significance test. The Results are presenting the statistical significance near to p = 0.05. A discrete investigation is performed to discover the consequence of nutritional-status well-defined via the Weight-For-Age Z (WAZ) scores on the clinical result. WAZ score is frequently used to define nutritional status in the analysis of kids aged less than five years. In low and middle income countries (LMICs),kids with pneumonia were at risk of severe pallor and underweight death and it might be vital to consider the best setting for the management. Obtained results suggest that the presence of the weight-for-age z scores either a lesser amount of –3 SD or any pallor degree between the kids with non-severe pneumonia classified by the WHO, would be considered together with the criteria of WHO for the admission attention in present-day African populations: this finding gives warranty for advance analysis. Though this result is unable to establish proper proof that the lives will be saved by using these factors, these outcomes suggest a risk of present outpatient treatment.

**Document Recognition implementing Gradient-based Learning Method**

Speech recognition and handwriting recognition are the growing applications of pattern recognition. And better are the ones which are learning automatically by directly using the machine learning techniques into the neural network that perform on pixels directly. The extraction of manual and hand-designed features (traditional way) has been replaced by Graph Transformer Network that globally trains the multi-module networks using a Gradient-based method to minimize error and increase the performance. For handwriting recognition, two modules have been integrated. First, the Feature Extractorwhichdoes not fluctuates when there are modifications and distortions in patterns of the input. Another one is the Classifier, which decides the category based on a suitable set of features. It decides the accuracy of pattern recognition. The OCR structures also use Neural Network that is trained with back propagation.

The research focuses on the recognition of patterns from characters taken individually to the variable-length words and phrases in any document. It states that combining multiple trained modules will lessen the all-around error. If in a network, the multi-modules can be trained by manipulating the directed graphs then it is known as the trainable Graph Transformer Network.

Research Methodology includes the learning of data from a gradient-based method and back propagation. The adjustable parameters like weight minimize the loss function. To measure the performance, the accuracy level of testing data is analyzed. The difference between the expected test set error and training set error is called a gap. The main focus is to minimize the training error as well as the expected gap. The characters are separated from their acquaintances within a word or phrase called segmentation. Segmentation can be removed by the concept of recognizing every possible location on the image and detecting the character even in between other characters. The practical recognition system is possible if information transmitted between every module is illustrated by graphs with information represented numerically on the arcs of graphs. The entire practical model has to be trained to minimize the global error to reduce the character misclassification. Each module in the network takes many graphs as input and produces an output which is also a graph forming a Graph Transformer Network. Using convolutional network LeNet-5, multiple features are extracted at various locations. After feature extraction, there is no need of knowing the location of the character. The weights are learned with back propagation and hence able to extract their own features. The loss function used here is the mean squared error.

The database used here is a modified NIST database which contains binary images of the handwritten digits. The error rate decreases on each pass of training. Overtraining can result in increasing error after reaching some threshold. But the database taken in this research was large enough to tackle this problem. More training examples were generated using modifications in original images but still error decreases. The classifier chosen for the model was compared with other classifiers also like SVM, PCA, etc.

The neural network trained with back propagation found to be the successful Gradient-Based Learning method. The network uses the global training combined with Graph Transformer Networks for accuracy. This is used in reading bank checks accurately.

**Back Propagation Technique for learning characterization**

The research proposes a unique method called back propagation that is a simpler learning method for the neural networks especially the layered ones. It tries to adjust the weights of the associations in a network in order to minimize the error. The error is the estimate of the actual and the desired output of the network. During the adjustments, hidden components, which are neither portion of input nor output, seem to characterize important features of the job and commonness due to their interactions. It is having the ability to create new features that a traditional method does not have. The goal of this study is to formulate an internal representation for a particular job by using some transformations in the neural network so that the network can become self-governed.

The hypothesis is that if there is a direct connection between input and output components, it will be easy to find learning methods that adjust it to minimize the difference between actual and desired output but it will not be able to learn representations or characterizations. And it becomes hard when it comes to hidden components that are not dependent on tasks. There must be a method to decide when should hidden units be active and by what amount. The learning procedure can haveinput, intermediate, and output layerswith connections in between thembut connections can hop intermediate layers. Though, the research aims to use a linear function for integrating the inputs to a unit, finding a batch of weights that assure that for each input, the actual output produced remains the same as the desired output. To minimize the error, gradient descent is computed with the partial derivative of the error with regard to each weight in thenetwork. The sum of these partial nets is computed for each pair of input-output. The forward pass in each layer determined by the input they receive from units in lower layers. The backward pass propagates from a higher output layer to the lower input layer. Hence, any change in overall input to an output affects the error and it can also be done by changing weights and sums.

The benefit is that no distinct memory is expected for the derivatives. This method does not connect as faster as the second derivatives method does, but it is simple and can easily be executed in hardware parallelly. It can be improved using an acceleration technique. The symmetry cannot be detected by just connecting the input to the output. And for the detection of the same, it is necessary to use the intermediate layer. Since a single input component gave no proof of symmetry or non-symmetry until now. Therefore, using individual input is of no use and hence used two intermediate units.

The most noticeable drawback of this learning method is that there may exist local minima. Hence, the gradient descent method here is not assured to find a global minimum. The evidence is that the network rarely got stuck in local minima making the situation worse than the global minimum. Such unwanted behavior is only found in networks with only having enough connections for a job. Therefore adding little more connections gives extra weight space. It concludes that applying this method to various jobs illustrates that internal characterizations or representations can be made by gradient descent.

**ImageNet: A Large order Hierarchical Image Database**

A large number of image data available on the internet has the strength to facilitate powerful advanced large-scale image search algorithms and models that can help the user to record, retrieve, organize, and interact with those images. This research proposes a new database of the images called ImageNet which is constructed using the hierarchical arrangement of WordNet. WordNet consists of multiple words and phrases called Synonym-Set or Synset. The goal of the ImageNet dataset was to describe each Synset with approximately 1000 images. This dataset is more diverse and accurate than any other dataset existing. The uses of this dataset include image recognition, object detection, and classification. The properties of the dataset were analyzed on the basis of hierarchy, precision, and diversity as densely populated, 99.7% precise, and diverse due to positions, appearances, and picture points respectively. The existing version of ImageNet having twelve subtrees including birds, fish, fruits, mammals, flowers, etc.

The methodology of the research includes the construction of the ImageNet dataset. It is done in two steps. First, 'collecting' the nominee image. Second, 'cleaning' the nominee image. The collection of nominee images is done for each Synset. The images are collected from the internet by examining numerous image search engines. For each Synset, WordNet synonyms are taken as a set of queries. To extend the dataset, queries are translated into other languages like Italian, Chinese, etc. After the collection stage, the cleaning is done using Amazon Mechanical Turk, an online arena where users are given tasks of labeling the data for which they get paid and the task has to be completed in time. It is convenient for large scale labeling. Human mistakes and limitations are considered during cleaning. The solution included multiple users labeling independently until a threshold reaches.

The ImageNet is compared with other datasets like a Small image dataset that is found not as challenging as ImageNet. ImageNet consists of 20 times more classes and 100 times more images than this dataset. TinyImages contained larger noise and lower resolution images than ImageNet. ESP included basic level image labeling, due to the acquisition of image labels from online gaming and human players within a certain limit of time. It was also not publically available. LabelMe and LotusHill consisted of images from researchers while ImageNet images were collected using Internet crawling.

Some of the ImageNet applications are Non-parametric image recognition which uses the nearest neighbor method and gives high resolution and clean images, Tree structure-defined Image classification which uses tree max classifier, and automated object localization which gives extra and extended information about each image.

The existing ImageNet dataset consists of 10% of Synsets from WordNet. The ImageNet is in the development stage and to speed up this process more AMT labels taken from users can ve analyzed. To accurately verify the images, the repetitions should be optimized. Future work includes completion of ImageNet with around millions of neat, diverse, and high-resolution images overall Synsets and making it available and accessible online to the public and researchers. Also storing in the cloud enables the fruitful distribution of the dataset and expanding ImageNet with more information and bringing up an online platform or a community-based setting where anyone can contribute and get benefitted as well.

**Automated Lung Disease Detection using Artificial Neural Network Classification and Chest Radiograph Tool**

Lung disorders such as Pneumonia, Tuberculosis, etc are crucial health risks and late diagnosis can lead to death. Tuberculosis (TB) is a contagious and deadliest disease. Pneumonia is an inflammatory lung disease. In Lung cancer, the sudden growth of bronchi cell linings leads to the formation of a tumor. For the timely diagnosis of such diseases, imaging technique is important. One such tool is Chest Radiograph (x-ray) that is used to diagnose such lung diseases. It is inexpensive than other techniques like CT scans. Chest radiographs render the exact location of the tumor in the lungs.

The research proposes a methodology for automated lung disease detection and classification. It goes through four stages: Preprocessing, Segmentation, Feature Extraction, and classification. The first step in automated detection of lung disorders is preprocessing of the images that remove extraneous and unwanted data bits available on x-ray, recover helpful information, and strengthen the area of interest and simplify features. It is done in two steps: image enhancement and image filtering. For image enhancement, the histogram equalization technique is used which uniformly distributes the image intensity pixel and adjusts the contrast. Image filtering removes undesirable noise existing in x-ray images using high pass filters that accentuate fine details and sharpen the image. The second step includes Lung segmentation that obtains region of interest in the lungs. It helps in identifying lung boundaries. There are two methods for segmentation. The first one is an intensity orientated method that uses a threshold value, below which the image pixels are set to 0 and becomes black while above the threshold value, the image becomes white. The other one is the discontinuity based method (Edge Detection method). The research used a Sobel operator, an edge detection method for lung boundaries detection. The discontinuities are the irregularity of the image pixels. It helps in edge detection. These boundaries are helpful in the third step of extracting important features like geometric features (area, parameters, etc) and statistical features (mean, standard deviation, etc). The final step is image classification for accurate identification of lung disease which is done by the feed-forward artificial neural network.

The overall Artificial Feed Forward Neural Network model developed in this paper consisted of 3 layers: input, hidden, and output. Seven variables were given as input to the network. The input layer was defined from the extracted features. The three hidden layers represent lung diseases and one output layer denotes the classification of the lung disorder. The neuron was activated by applying a sigmoid function. The output layer infers the chest x-rays as lung disorder positive or negative. For error reduction, the backpropagation method is used. Chest radiographs of 80 patients from the database have experimented. It was done using a neural network toolbox in Matlab.

This research developed an automated detection system of major lung diseases and the results were 92% precise. Nonetheless, the proposed method is limited as little changes in height and degree of the position of chest x-ray disturbed the result and it can be overcome by the digitization of x-ray images. Another limitation is that Chest Radiographs are alone adequate for early diagnosis but in the case of lung cancer, a CT scan is the best solution. It can be concluded that with the help of automated classification and accurate detection in Artificial neural network, the death rate can be lowered.

**Lung cancer Detection using Computer-Aided Diagnosis CT Scan Model**

In the early stages of cancer, patients are not conscious of the symptoms which can lead to a late cure and even death. Computer-aided diagnoses (CADs) are very much useful in detecting diseases from medical-based images and thus are helpful for radiologists and other medical specialists. Chronic diseases like lung cancer can be detected using this. It starts when the bronchi cells in the lung got mutated or lose their ability to stop growing further, forming small and round lumps also called nodules. All nodules are not cancerous. Smaller nodules that stop growing after some limit is benign while those who do no stop are malignant. There are various screening tests to detect lung cancer like MRI, CT, PET, etc. CT scan images work productively in actual time because of effective price, quick treatment, and lesser radiation. They produce detailed and high-resolution images and can detect accurately. But in a real-world scenario, the radiologists do manual detection that is not enough. This reason motivates CAD development.

The central motive of this research is to formulate a CT scan based CAD for lung cancer detection utilizing several image processing techniques. The detection involves three important methods: segmentation, feature extraction, and classification. A common method for segmentation is Water Shed. Lung parenchyma also includes other non- nodules that can be mistaken for cancerous nodules. Therefore, False Reduction is done before the classification stage. For the classification purpose, some of the methods are support vector machines, Neural networks, etc.

The methodology of the research includes downloading the CT scan images from the public access database (cancer imaging archives) which will be acting as input for the proposed system. They are in the DICOM structure that is converted to JPEG. Segmentation is performed into two portions. First, Parenchyma (present in lungs that include alveoli, blood vessels, etc) is segmented first using image masking where only pixels of the region of interest is set to 1 and others as 0. Next, the nodules are segmented using the easy threshold technique. The reconstruction of edges of parenchyma is done whenever the nodules were disappeared due to the lung walls. It is done using edge detection. To remove the false detection, trachea at the center was erased and from the segmented nodule, other non-nodule objects that are lesser than 1 mm (already present in lung parenchyma) were also removed. Features were extracted for classifying the nodule into benign and malignant such as geometric features, radiology features, statistical features, and texture-based features. The final stage of classification was attained using a multilayer perceptron (MLP) which is a feed-forward neural network that consists of multiple layers. The network is trained using the backpropagation method. It was trained with ten extracted features, thirty hidden layers, and two output categories.

This results in a CAD model that detects lung cancer from numerous sequel of CT scan images. It was accomplished using the Lung Image dataset. Classification preciseness was 95% due to some limitations. There were classification errors like benign cancer is mistakenly detected as malignant and vice versa. The average processing time was 20 seconds. The model was generated in MATLAB. The future scope includes extracting features from three-dimensional images to get decent results.