Original Article

Pulmonary Nodules Classification In Computed Tomography Images Using Faster R-CNN

Anitha. S[1], Hemavathi. P.V[2], Sethu. S[3], Subbulakshmi. M[4], Packialatha .A[5], Vijitha. S[6]

[1]sanitha.se@velsuniv.ac.in
CSE ,Department Vels Institute of Science, Technology& Advanced Studies
[2]pvhemavathi.se@velsuniv.ac.in
CSE ,Department ,Vels Institute of Science, Technology& Advanced Studies
[3]sethu.se@velsuniv.ac.in
CSE ,Department ,Vels Institute of Science, Technology& Advanced Studies
[4]subbulakshmi.se@velsuniv.ac.in
CSE ,Department, Vels Institute of Science, Technology& Advanced Studies
[5]packialatha.se@velsuniv.ac.in
CSE ,Department, Vels Institute of Science, Technology& Advanced Studies
[6]vijitha.se@velsuniv.ac.in
CSE ,Department, Vels Institute of Science, Technology& Advanced Studies
DOI: 10.47750/pnr.2022.13.S10.177

Abstract

Pulmonary nodules are the primary symptom of lung cancer, which has a high mortality rate anywhere in the globe. Radiologists are able to spend less time on false positives and missed diagnoses because to automatic pulmonary nodule identification. We suggest using a more efficient variant of the Faster R-CNN algorithm to identify these pulmonary nodules in early stages. A more efficient Faster R-CNN algorithm is able to identify pulmonary nodules, and this is shown by using the training set. Theoretically, modification of the parameters involved in a network may lead to both structural and detection improvements. This research proposes an enhanced and optimized approach for identifying pulmonary nodules, which, on average, increases detection accuracy by over 20% compared to previous, more conventional techniques. The Faster R-CNN-based technique of pulmonary nodule detection demonstrated high accuracy in our experiments, suggesting it may have practical use in the diagnosis of pulmonary illness. This technique may be of additional use to radiologists and academics working on improving the pulmonary nodule's detection lesion.

Keywords: Pulmonary nodules; deep learning; faster R-CNN; computed tomography; computer-aided diagnosis

1. INTRODUCTION

Humans suffer from several forms of cancer, including brain, lung, bone, liver, breast cancer, etc. Lung cancer has a survival rate of little about 17% after 5 years throughout all stages of the disease. If lung cancer is found early, 52% of people will still be alive after five years. The survival rate has improved dramatically, but it must be increased even more than its present level. So that the undesired lesion may be eliminated, an inside view of the body is necessary. In the year 2020, it is anticipated that there would be around 200000 newly diagnosed cases and 135,700 mortalities in the United States. Combined estimates for colon, pancreatic, and prostate cancers are lower fatalities, therefore this is a significant increase above those. Yet it has been shown that early diagnosis and treatment may significantly reduce mortality rates. A total of 2.09 million new instances of lung cancer will be identified in 2020, making it the most frequent disease globally, as reported by the World Health Organization (WHO) [1]. Worldwide, 1.76 million people

lose their lives to lung cancer every year, making it the disease's primary cause of death. Fig.1 shows the top 10 countries with highest lung cancer rates.

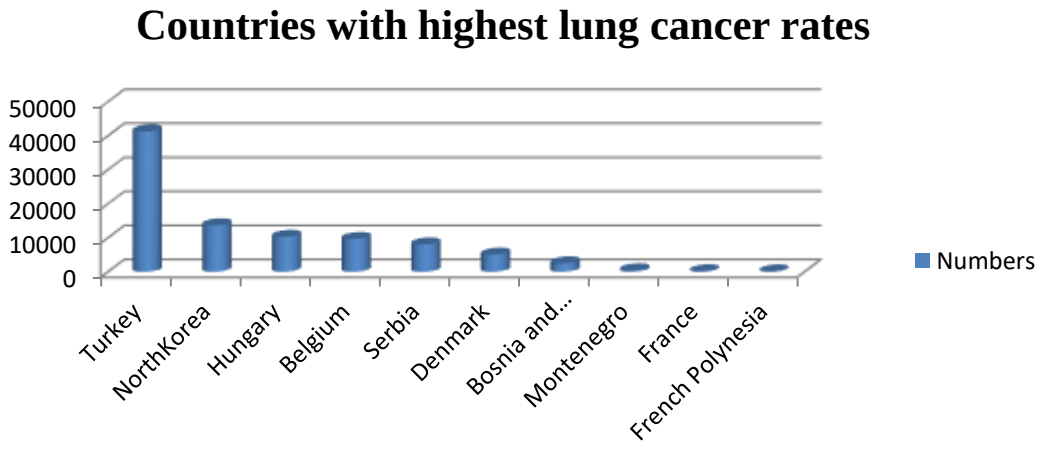


Figure 1: Countries with highest lung cancer rates

To successfully cope with this challenge, early diagnosis and treatment are essential. Early detection consists of two parts: screening and early diagnosis. Various methods like, X-rays, Computed Tomography (CT) scans, Magnetic Resonance Imaging (MRI), etc., are used to get images from the inside of the body. The most suggested procedure for the lungs is the CT scan, which gives three-dimensional pictures of the lungs. The lesions may be automatically recognized by increasing the lung picture. Recent advances in Magnetic Resonance Imaging (MRI) have enabled for its application in the diagnosis of lung illness. Even while CT is better than MRI in finding microscopic nodules, MRI still has its benefits. Unlike CT scans, MRI does not expose patients to any radiation. Neither method was noticeably better than the other in spotting malignant nodules. MRI has shown potential as a screening method for lung cancer and the identification of malignant nodules. Functional information, including physiological, pathophysiological, and molecular data, may also be obtained from an MRI scan, expanding the technology's use beyond its purely diagnostic uses. Therefore, MRI is useful for identifying lung diseases. Each MRI scan has numerous slices, and because MRI may use various modalities, this means that each case's dataset is quite large. In addition, the radiologist's level of expertise and visual tiredness has a role in the sensitivity of the diagnosis. This makes it difficult to use Computed Aided Detection(CAD) software on MR pictures. Measurement and categorization of pulmonary nodules depend on their detection. There is a pressing need for research towards automated lung nodule detection[2].

Recent widespread use of CT has enabled for dramatic progress in the early detection of malignancies in cases of lung cancer. The use of low-dose 3D chest CT images is now one of the most accurate methods for manually diagnosing patients. While advancements in radiology technology have made cancer nodule detection easier, the process is still time-consuming and error-prone, since several images must be examined in detail, often by many radiologists[1,2]. Many different kinds of CAD systems have been created to help solve this problem and lighten the load on professionals. Analytical techniques include manually extracting the shape, volume, hardness, and other common properties of malignant or benign nodules. Not very long ago, deep learning algorithms became the go-to for detecting and classifying nodules, and their results have been fascinating.CAD technologies are being built to assist medical professionals, namely radiologists and doctors, in the process of identifying and diagnosing pulmonary nodules.A detection system, known as CADe, and an aided diagnostic system, known as CADx are often both components of CAD.The purpose of the CADe test is to determine if patients for pulmonary nodules really have nodules or not, while the purpose of the CADx test is to determine whether the nodules that have been discovered are malignant or not. Numerous lung nodule detection techniques for CT scans have been investigated. A few detecting algorithms have been implemented for MR images. CT provides greater image contrast and fewer artefacts than MRI.

Due to the variations between CT and MR pictures, CT nodule identification algorithms may not be applicable to MR imaging [3].

Typically, nodule detection systems include, preprocessing, segmentation, identification, and false positive (FP) reduction. Based on the location of growth, the thorax and lung parenchyma were first removed. To improve the parenchyma shape, a rolling-ball algorithm-based approach for lung reconstruction was developed. At last, adeep learning classifier was trained to categorize the lung images as either nodule or FP. On 33 CT tests, this approach yielded 0.43 FPs per scan and 0.17 FNs/scan value. A technique for detecting internal and juxta pleural nodules has been developed. The first potential nodule was identified and segmented using a 3D mass–spring model. Then, seven characteristics were extracted based on their form, size, and intensity. Finally, a method for FP reduction was developed.This approach was evaluated using 90 scans. A 3D shape feature descriptor [4] was suggested to identify lung lesions on CT imaging. Initially, lung nodules were segmented by combining thresholding with different component labeling of related components. They created a technique for detecting tiny lung nodules. The pulmonary parenchyma was first segmented using a process called region growth. Combining the Gaussian mixture models with the Hessian matrix yielded the nodule candidate. This technique was evaluated using 28 CT images. The detection system has a sensitivity of 90.6% and a specificity of 85%. The pulmonary parenchyma is first removed and rebuilt. Then, candidates for nodules were identified. A deep learning classifier was then trained once form and texture characteristics had been retrieved.

Deep learning is now regarded as one of the finest solutions for several computer vision and pattern recognition issues, including image processing, audio processing, video processing and natural language processing, etc. For very precise object recognition and classification, many researchers turn to Convolutional Neural Networks (CNNs), which utilize convolutional features in each of their layers. Because of these advantages of deep learning, pulmonary nodule detection and classification systems are made to be simpler. Most prior nodule identification approaches need manual extraction of features [5]. In several pattern recognition tasks, CNN that automatically learns features has recently shown promising results. This motivated the use of CNN to the automated identification of lung nodules. They first found the potential nodule using conventional methods. Then, for each candidate, a patch was removed and CNN was used to categorize it as normal or nodular. Most deep learning-based approaches for CT nodule identification consist of two phases. In this study, Faster R-CNN was built for pulmonary nodule identification to address these above issues. Rapid R-CNN accepts the whole picture as input and does not need feature extraction. Additionally, it can identify objects of various shapes due to the use of several anchors. Transfer learning is used in parameter optimization to circumvent the issue of overfitting. Faster R-CNN, the detection results include several False Positive (FP) areas. A false positive reduction approach based on the anatomical properties of pulmonary lesions is meant to decrease FPs and maintain real nodules [6].

The suggested deep learning approaches may be split into two ways. The first method identifies many suspicious spots that contain a nodule, and the second module to confirm or deny the existence of a lesion. This technique identifies and categorizes nodules from start to end. In CT images, a CAD method based on **Faster-RCNN** is utilized to locate nodules. Similarly, the object identification framework Overfeat detects nodules directly from 2D slices derived from CT images. Given the variety of available object detectors, it might be difficult to choose which to utilize as a starting point for the construction. We hope that our research will help set a standard for figuring out which deep learning approaches for finding nodules are most likely to work [7,8].In this research, we investigate the topic of identifying candidates for lung nodules in CT scans as nodules or non-nodules. We will demonstrate that our CNN network with focused loss is an effective way for classifying lung nodules.On the other hand, there are still numerous obstacles and constraints to overcome. In order to better grasp CAD technology for pulmonary nodule identification, categorization, this review is helpful to researchers and radiologists. We provide a concise overview of the efficacy of existing methods, discuss some of the obstacles standing in the way, and suggest avenues for further groundbreaking study.

2. RELATED WORK

In order to decrease the number of false positives in volumetric CT images used for automated pulmonary nodule diagnosis, researchers have developed a unique 3D CNN. Because of their hierarchical design, which is trained using 3D samples, 3D CNN are superior to their 2D counterparts in terms of their ability to encode more spatial features and extract more relevant features. To address the issues posed by the wide range of pulmonary nodule appearances and the difficulty in differentiating them from one another, we also offer a simple yet efficient method of encoding multi-level contextual information. Making use of ensemble learning with various 3D CNNs, this system can determine whether or not a CT image has micro-nodules. A total of 35,000 image samples were collected from the LIDC database, consisting of 13000 nodules and 22000 non-nodules. These samples were gathered from around 1000 CT scans. Five alternative images sizes are used to examine potential lung nodules. Then, using candidates for a single size of nodule, different three dimensionalCNN models are constructed and deployed. The final classification results are obtained by using an Extreme Learning Machine network to merge the outputs of the various 3DCNNs. Performance-wise, these approaches much outperform two dimensional CNNs, a single three dimensional CNN model, and traditionalmethodologies trained on same database. Findings show that it is possible to create an autonomous system that can distinguish between nodules and non-nodules in CT scans, which opens the door to research into micro-nodules in lung cancer. This method might be useful in the creation of additional trustworthy clinical-decision support systems [9]. Our performance in LUNA16, where we got the highest possible score for reducing false positives using the Competition Performance Metric (CPM), was a big vote of confidence in the suggested method. The framework can be used to find 3D objects in volumetric medical images in many different ways [10]. Due to their 3D structure, nodules require the use of 3D CNNs in both the training and inference phases of the CAD system. In the first phase, we specifically develop a segmentation-based 3D CNN with a hybrid loss to classify nodules. Nodule candidates are created using a threshold technique and connected component analysis based on the probability maps provided by the segmentation network. Second, to cut down on erroneous positives, we use three distinct 3D CNNs trained on various classification tasks and fed with a variety of inputs. To make greater segmentation network, we incorporate hybrid inputs in addition to the traditional raw data input. The use of data augmentation and batch normalization in studies helps prevent overfitting [11].

Lung nodule detection system using a methods such as a unique SVM classifier, random undersampling, ROI, and SMOTE.By achieving a balanced training sample and removing noise and duplicate information, the SVM method for pulmonary nodules classification performs better under imbalanced data. Eight features are retrieved for use in training and classification, and they include both 2D and 3D features [12]. The local image properties are used in conjunction with two sequential k-nearest-neighbor classifiers to identify potential structures inside the lung volume in a totally automated manner. Comprehensive metrics are utilized to evaluate the system's performance, such as the percentage of correctly identified lesions as malignant and the accuracy rate for nodules of varying sizes in the databases [13].Researchers used advanced CAD system, named Faster R-CNN model with an anchor box, to identify pulmonary nodules. Improved detection accuracy is achieved by learned anchors as a hyper parameter. To mitigate the FPs in the results generated by Faster R-CNN, Densenet, it is suggested to use a residual CNN. Our comprehensive 3D CNN model for nodule generation is based on the 3D U-Net architecture. It is provided in an effort to reduce the FPs nodules. This topology improves feature transmission and facilitates the reuse of nodal characteristics [14].

Although micro-nodules are considered the first signs of lung cancer, they have received very less attention in the few automated methods for detecting pulmonary nodules that have been developed. As a consequence of not being able to tell the difference between micro-nodules and other particles, the FPs rate of the majority of the currently available technologies is rather high. The suggested system beats other current detection techniques in accuracy 95.13, sensitivity 94.35, and F-score 94.20, AUC 0.96 and with a CPM score of 89%. Experimental outcomes showed the significance and efficacy 3D CNN architecture for automated pulmonary nodule diagnosis in volumetric CT data [15]. Using 1000 CT scans from the publicly accessible IDRI-LIDC and SPIE-AAPM datasets; we conduct an evaluation of the system and find that, ours obtains the best performance with just one false positive per scan on average. One hundred and fifteen CT images from nearby hospitals are also analyzed for additional information [16].85% sensitivity with an average of 5.1 FPs/scan is attained in a sample of 850 scans from the screening study. However, this method of object recognition struggles due to the varying nodule sizes present in training datasets. In spite of their rarity,

subsolid lung nodules have a far higher malignancy rate than their solid counterparts. Several parts of the CAD system, as well as the testing reports, were used to fine-tune the system's performance. A complete set of 128 features has been developed to aid in the identification of possible subsolid nodule candidates. In addition to the usual luminance, form, and texture, we now have access to new contextual factors. Data from several lung cancer screening centers is used to train and enhance the CAD system [17].

Numerous CAD systems have been developed for automatic pumonary nodule identification, which is a highly researched area. The majority of CAD programmes have two distinct phases: the identification of potential nodule candidates and the mitigation of false positives. Morphological features, voxel clustering, and pixel thresholding are examples of the kinds of characteristics that must be manually produced for use in traditional techniques. In the detection phase, a nodule segmentation technique based on enhancement filters and features is designed to localize cluster cores. Classification comes next, and invariant characteristics established on a gauge coordinate system help distinguish between true nodules and blood vessels that might cause false positives. Both lungCAM and the voxel-based neural method approaches start with a three dimensional region-expansion algorithm that removes the trachea and separates the lungs into two sections [18].

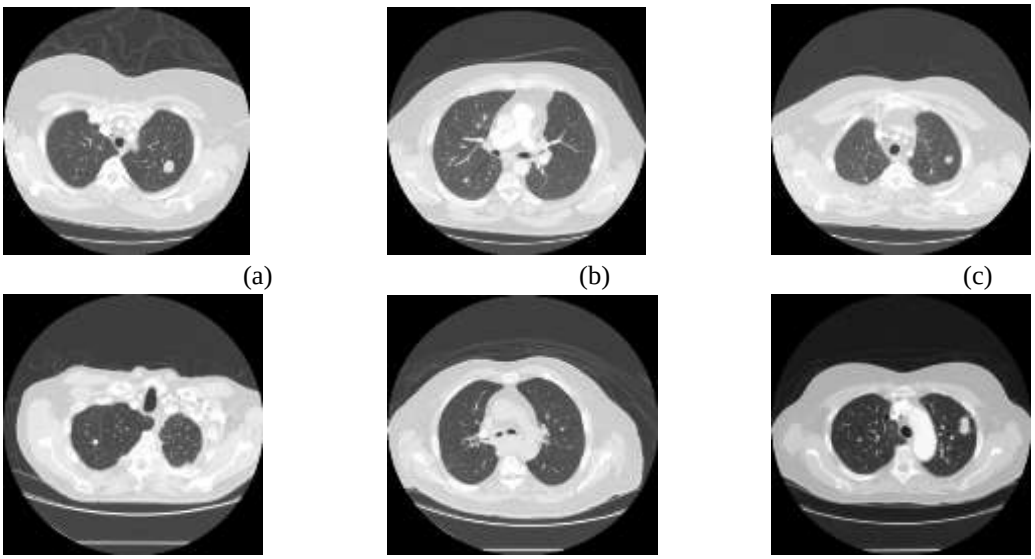
In order to reduce false positives and improve candidate discovery, ZNET employs convolutional neural networks. A 512-by-512 square is extracted from the input slice. When U-Net is deployed to each axial slice, it produces a probability map that is used for candidate identificationmaximizing the number of discovered nodules. Partial volume effects are then eradicated using a linked component analysis is used to categorize the applicants. Locations of potential candidates are shown by their components' centers of mass. ZNET employs broad residual networks to reduce false positives, and 64 x 64 patches are taken for each candidate. The extensive residual networks handle each patch independently. The final forecast is calculated by averaging the network's projected values at the output for each of the three separate patches. JianPeiCADutilises a rule-based screening on many scales to identify potential nodule locations[19].

Overfitting is mitigated during training via data augmentation in a 3D CNN with large channels dedicated to reducing false positives. In order to calculate the volume of the lungs, MOT M5Lv1 employs a three-dimensional growth region technique, excluding the trachea and separating the lungs. Nodules are separated using a combination of morphological processing and thresholding at several grey levels. In order to reduce false positives, a total of 15 characteristics are computed. These include geometrical parameters and intensity features (radius, sphericityaverage, standard deviation, maximum, entropy). Feed-forward neural networks are used for classification. With its proposed 3D CNN nodule identification architecture, Resnet first uses a fully convolutional network to filter out low-probability regions before retrieving those with a higher likelihood. As a further step toward reducing false positives, they suggest incorporating a series of 3D CNNs with varying receptive field widths to better handle the wide range of pulmonary nodule changes and differentiate them from their hard counterparts. Using an iterative method including the deployment over a specified threshold, M5LCAD is able to first segment the lung's internal components. The ant colony relocates to a new location and begins releasing pheromones in accordance with a predetermined set of regulations. New ant colonies are placed in unexplored voxels, while existing colonies leave voxels they have already explored. The pheromone maps are thresholded iteratively to get a shortlist.Recently, DCNN and feed-forward neural network [20] have shown their ability to efficiently extract picture information for accurate classification in a wide range of various tasks or datasets without modifying the network topology. Many well-known object identification frameworks have been presented up to this point, including Faster R-CNN, Alexnet, Densenet, SSD, Resnet, and R-FCN, all of which create candidate bounding boxes in a single step using more accuracy and less time.

3. DATASET

We acquired 3000 lung scans from a publicly accessible data called LIDC to supplement the training samples (2,700 images) and further validate the algorithm (300 images). The filtered back projection of the lung kernel was used in order to rebuild images with a dimension of 512 by 512. The slice thickness varied between 1.25 and 1.5 millimeters. All 2700 training scan images must be annotated, which is laborious and time-consuming. To shorten the overall

labeling time, the pre-trained model was then applied to 2700 images with a low scoring threshold of T1 D 0:1 that was utilized to produce candidates. For a particular nodule with a diameter < 70 mm among the possibilities, each nodule slice had a corresponding two-dimensional bounding box marked with a number that represented the nodule kind. In the end, a total of 10000 nodules were recovered, of which about 85% were small nodules measuring less than 7mm in diameter. Approximately 17% of the cases ranged size from 7 mm to 30 mm, while 5% were bigger than 30 mm. The proportions of three different nodule types were about 24.0%, 6.5%, and 72.6%, respectively. Separate object-detection models are trained and optimized for each detection class to avoid dataset imbalance. A conceptual test is conducted in time environment, therefore obtaining an automated assessment of the product's quality, removing the need for a human expert's interaction, and eliminating visual analysis mistakes. In some cases, the conceptual test gave great results, with a 98 percent accuracy rate.



(a) (b) (c)

(d) (e) (f)

Figure 2: Lungs with nodules

4. PROPOSED ARCHITECTURE

4.1. CNN Architecture

Our Faster R-CNN receives a 224 by 224 RGB picture during training. In pre-processing, we remove the mean RGB value from each pixel. The image is processed through a stack of convolutional (conv) layers using 3-by-3-pixel filters. In one arrangement, we use 1 * 1 convolution filters, which change input channels linearly. The convolution step size is 1 pixel, and the input space padding for 3 * 3 convolution layers is also 1 pixel. To improve the classification of pulmonary nodule or non-nodule, we propose Faster R-CNN, a novel 2D DCNN architecture. The three distinct layers that make up **Faster R-CNN** are the convolutional layer, the pooling layer, and the fully connected layer. Each input image undergoes a series of 2D convolution operators in a convolutional layer. During the training process, these layers may use the input image to extract relevant information. Features at a higher level of abstraction may be detected by networks with deeper layers. Convolutional layer defined as in Equation (1)

$$x_{i,j} = \sum a \sum b \, w_{a,b}^{(l-1,f)} y_{i=a,j+b}^{(i=1)} + bias^{f} \quad (1)$$

Where,

$y_{i,j}^{l}$ – Output layer of l

$w_{a,b}^{(l-1,f)}$ – Weight of filter f

The five max-pooling layers, which follow part of the convolutional layers, are responsible for the spatial pooling that is performed. A pooling layer may combine the values of neighboring features into a single value by operating on separate feature channels. As a result, it shortens the training time, limits overfitting, and minimizes the number of trainable parameters. When a layer is "completely linked," all of its neurons are coupled to one another and to those in the layer below it. Due to the inherent spatial connection in pictures, convolutional layers perform better than fully connected layers for 2D images. The training time of a network is increased because the number of trainable parameters is higher when using fully linked layers. An approach called "dropout," which regulates the number of neurons and connections, is presented as a solution to this issue. Dropout is defined as in Equation (2)

$$E_R = \frac{1}{2} \ (t - \sum_{i=1}^{n} p_i \ (1 - p_i) \ w_i^2 \ I_i^2 \qquad (2)$$

Grayscale images with a resolution of 224 * 224 pixels are fed into the network. There are a total of 15 layers in our network, nine of which are convolutional. Every one of our three convolutional blocks has three convolutional layers and a pooling layer that makes use of the max operator. Late network nodes extract more characteristics at a higher abstraction level than early ones. As you go down the blocks, the filters change from 5 * 5 to 3 * 3. The RELU activation function is used after each convolutional layer because it improves computation performance more than other nonlinear activation functions. It is defined as in Equation (3)

$$y = \begin{cases} x & if \ x \geq \ 0, \\ 0 & otherwise, \end{cases} \qquad (3)$$

Last is soft-max. All networks have identical completely linked layers. Hidden layers have ReLU non-linearity. None of our networks use local response normalisation, which doesn't enhance speed but increases memory and calculation time. With the characteristics retrieved by the earlier convolutional blocks, the last two fully connected layers attempt to classify the data. The output of the preceding layer is fed into a normalized exponential function called softmax, which generates a categorical distribution. Nodular probability is determined in our network using softmax. In order to prevent overfitting, we use a dropout method after the first completely linked layer. To maintain the network depth, the first and final three convolutional layers are padded to provide a high output size after the first three convolutional blocks.Softmax is defined as in Equation (4)

$$\sigma(\overrightarrow{Z}) = \frac{e^{Z_i}}{\sum_{j=1}^{K} e^{Z_j}} \qquad (4)$$

4.2. Focal Loss

The initial objective of the focus loss function is to solve the issue of excessive balance between foreground and background classes during object identification. Focal loss is primarily used for object identification, but we demonstrate that its sparse-specific properties are equally suitable for unbalanced dataset classification problems.

Focal loss begins with the cross-entropy loss function for classification,shows in Equation (5)

$$CE(p, x) = \begin{cases} -\log(p) & if \ x = 1, \\ -\log(1 - p), & otherwise, \end{cases} \qquad (5)$$

Where,

y - positive and negativeground truth values

p - estimated probability for the class with label x = 1

A weighting factor $\alpha \in [0, 1]$ is provided in a similar notation to balance the impact of positive and negative data:

$$\alpha_t = \begin{cases} \alpha & if\ y = 1, \\ 1 - \alpha & otherwise, \end{cases} \quad (6)$$

A modulation factor m with an adjustable focusing value $\gamma \geq 0$ is added to the cross-entropy loss function to reduce the loss contribution from readily identified data.

$$m = (1 - p_t)^{\gamma} \quad (7)$$

Equation (8) becomes the focused loss function with these two extra components.

$$FL(p_t) = -\alpha_t (1 - p_t)^{\gamma} \log (p_t) \quad (8)$$

It is important to take note that and $\alpha\ and\ \gamma$is two factors that indicate how sensitive it is to the samples that may be readily identified. Within the scope of this study, we want to apply this focused loss function to the very last stage of the 2D CNN architecture that we propose. We will show that our idea helps make CAD systems better at putting lung nodules into the right category.

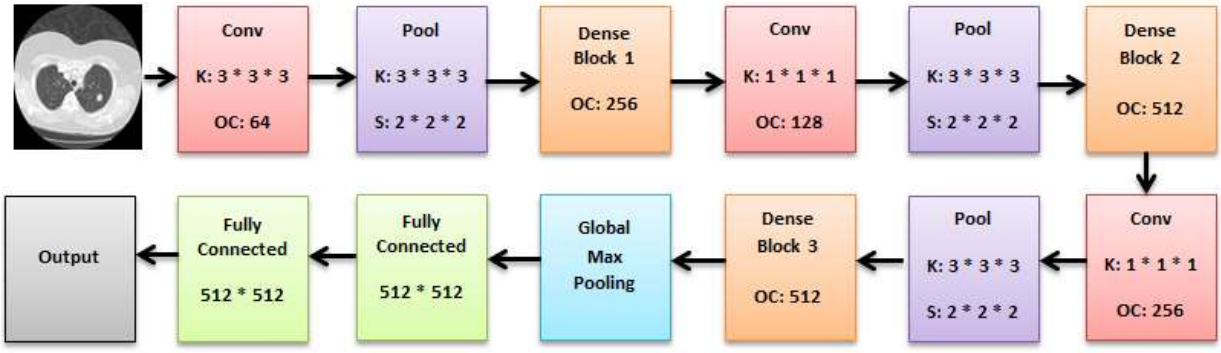


Figure 3: Architecture of Faster R-CNN model

4.3. Nodule Detection using Faster R-CNN

Candidate bounding boxes for lung nodules are what nodule detection is all about. Our nodule identification model is based on Faster R-CNN since it offers good performance in object detection. We add ISODATA to Faster R-CNN's proposal network for automatic anchor box configuration, which boosts the detection performance. For each lung nodules images, we first combine its two nearby slices as an input image to fulfill pre-trained CNN layers and leverage the 3D information of CT images. The pictures are fed into the improved Faster R-CNN model, which has a Region Proposal Network (RPN) to suggest possible nodule candidates and a classification network to separate real nodules from background noise.

4.4. Region Proposal Network.

When given an image to process, Region Proposal Network will provide a list of candidate areas. Each area is a bounding box that has been projected to contain a lung nodule and given an objectness score that indicates the degree to which that prediction is supported by the data. CNN layers are used for feature extraction from the input pictures.

Afterwards, feature maps of these CT scans are obtained. Sliding windows of size 3 * 3 are used on the feature map by Region Proposal Network to construct recommended areas. After the sliding windows have been flattened, they are used in class score prediction and bounding box regression in two fully-connected layers. Anchor boxes are a kind of proposal that may be predicted by each sliding window. These anchor boxes are made with different size, scales, ratios and shapes.

The initial Faster R-CNN model generated anchor boxes of various sizes and aspect ratios. Suitable anchors have been shown to have a significant influence in the detection stage. Each anchor box has its own unique set of features that, taken together, attempt to match objects of varying sizes. However, because of the wide range of possible nodule sizes and ratios, configuring them needs a lot of manual labour. Therefore, we present an automatic anchor box configuration approach that uses an efficient clustering algorithm to discover the optimal anchor box settings. **ISODATA**, the iterative successor to K-means, is a flexible clustering method that may divide and merge clustering results based on the distribution of input data. A flexible anchor box size configuration algorithm is give below.

Algorithm 1: Anchor Boxes Size Configuration Algorithm

Input: n nodule sizes $x_i(w_{i,}h_i)$, threshold T
Output: Flexible anchor box sizes $\hat{Y}[\quad]$
function flexible anchor box

for $a = 0$ **to** $n - 1$ **do**

$$l = 0$$
$$mode\ [a][k] = x_i;$$

repeat
mode$[a][l+1]$ = nextMode(mode$[a][l]$);

$$k = k + 1;$$

untildist (mode$[a][l]$, mode$([a][l-1]) < T$,
mode$[a][0]$ = mode$[a][l]$;
end for

$\hat{C}[\quad]$;.
for $a = 0$ **to** $n - 1$ **do**
if mode $[a][0]$ not in $\hat{C}$ then
$\hat{Y}$+= [mode$[a][0]$];
end if;
end for

return $\hat{Y}$**;**
end function

4.5. ROI Classification Network

The purpose of the region-of-interest (ROI) classification algorithm is to ascertain whether or not a certain area contains a lung nodule. After the feature map is extracted and areas are offered, a ROI pooling layer is used to convert the various sized areas into uniformly sized feature maps. Next, two fully connected layers transfer the feature maps into a vector, and then a bounding box regression and bounding box classifier are employed for position regression and classification between pulmonary nodules.Predited boundary boxes values described in Equation (9)

$$L(\hat{p}, p, t^{\hat{p}}, t^{p}) = L_{cls}(\hat{p}, p) + I(P \geq 1)L_{reg}(t^{\hat{p}}, t^{p}) \quad (9)$$

Where,

$\hat{p}$ – Predicted anchor value
p – True probability of an anchor value
$t^{\hat{p}}$ – predicted boundary boxes
t^{p} – True boundary boxes
$I(P \geq 1)$ – Regression loss
L_{cls} – Classification loss
L_{reg} – Regression loss

4.6. False positive Reduction

To that end, this stage seeks to significantly reduce the number of false positive candidates. A large body of research uses classification strategies to refine categorization outcomes. Multiclassifier systems were established by integrating the results of a neural network with those of a decision tree and a statistical classifier. The combinator's performance was far higher than that of the finest single-nodule detector. To arrive at the final classification, we combined the findings of several independent streams of 2D Convents using a novel fusion method. The experimental results showed that the system's performance was enhanced by the incorporation of more perspectives into the design. In this study, we deploy a collection of 3D CNNs trained on classification data to model the 3D topology of lung nodules. We utilize three networks with different inputs to incorporate different types of data and maximize the output probability maps from the prior segmentation network so that we may make use of them and perform a comparison.

$$F_{fl} = \begin{cases} -\alpha(1-\hat{y})^{\gamma}\, log\hat{y}, & y = 1 \\ -(1-\alpha)\hat{y}\log(1-\hat{y}), & y = 0 \end{cases} \quad (10)$$

Where,

$\hat{y}$ – predicted value
α – Weighted factor
γ – Tunable focusing parameter

4.7. Performance Metrics

The correctness of the classification may be evaluated in terms of accuracy (the capacity to properly discriminate nodule and non-nodule instances), sensitivity (the capacity to correctly detect nodule cases), and specificity (the capacity to correctly decide non-nodule cases). These measures are often used for classification issues, and their formulations are given below:

$$F1\ score = 2.\frac{P.R}{P+R} \quad (11)$$

$$Precision = \frac{TP}{TP+FP} \quad (12)$$

$$Recall = \frac{TP}{TP+FN} \quad (13)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN'} \quad (14)$$

Where, TP - number of instances where nodules were correctly identified
FP - number of instances where nodules were incorrectly identified
TN - number of instances where nodules were correctly identified
FN - number of instances where nodules were incorrectly identified

5. RESULTS AND DISCUSSIONS

As a model, it scored 93.42% accuracy, 0.994% AUC, 94.73% F-Score, and 93.93 Sensitivity. We may infer that a deeper network can be highly beneficial in enhancing nodule identification performance using region proposal approaches, as compared to the variation with another CNN models. As a result, two images are generated, each of which contains a segmented pulmonary nodule. This method is about as accurate as Resnet -50 on average, but it is significantly faster. Nonetheless, deeper networks do worse. Performance-wise, Efficient Nets are a step up from SSD and CenterNet and on par with Resnet-50 using Faster-RCNN. Higher versions of Efficient Net were also tested; however, they were still unable to identify pulmonary nodules. Due to this, we have chosen not to provide their performance in this comparison. As a conclusion, it seems that no one variant of YOLO is particularly well-suited to the job of nodule identification, and all of them achieve around the same average accuracy. However, these findings should not be taken at face value since network performance may vary depending on training hyper-parameters. Advancing the optimization of hyper-parameters is a focus of our future research. The findings as a whole show that direct nodule recognition is very difficult, even for cutting-edge deep object detection methods. Tab.1. shows the various CNN model's performances in lung nodule detection

Table 1: Classification performance with different neural networks

CNN Models	Accuracy	AUC	F-Score	Sensitivity
Faster R-CNN	**95.42**	**0.994**	**94.73**	**93.93**
DenseNet	89.54	0.917	85.43	82.58
Resnet	78.75	0.960	92.26	91.94
Alexnet	80.90	0.979	91.73	90.80
Googlenet	91.22	0.971	92.21	91.42
R-CNN	90.65	0.928	83.43	80.58
Fast R-CNN	79.86	0.971	90.26	89.94
PSPNet	90.91	0.911	89.73	88.8
PFN	92.32	0.982	90.21	89.42
SSD	90.00	0.922	89.73	88.8
YOLO	92.33	0.985	90.21	89.42
CenterNet	90.56	0.958	89.73	88.8
EfficientNet	92.33	0.969	90.21	89.42

The Faster R-CNN method is compared with DenseNet, Resnet, Alexnet, Googlenet, R-CNN, Fast R-CNN, PSPNet, YOLO, CenterNet and EfficientNet are presented in Fig.4.

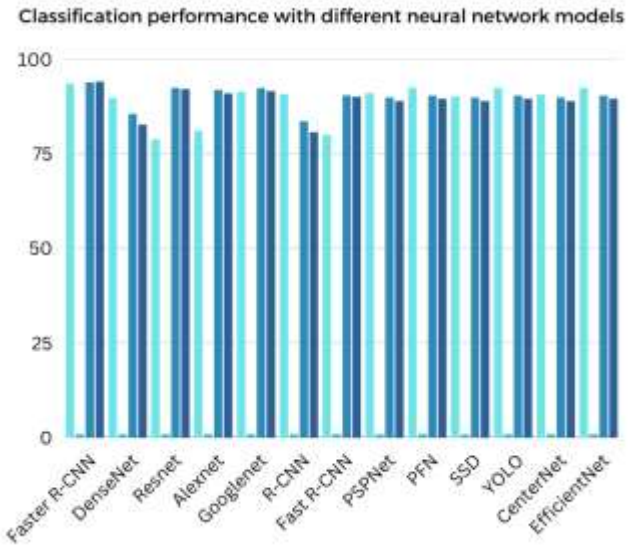


Figure 4: Classification performance of CNN models

A comparison between the various Loss functions and optimizers are shown in Tab.2. Logcosh loss function with Adam optimizer performs better among others.

Table 2: Proposed model trained with different loss functions and optimizers

Loss Functions– Optimizers	**Accuracy**	**AUC**	**F-Score**	**Sensitivity**
Kullback-Leibler (KL) divergence loss - Gradient Descent	93.52	0.976	70.56	80.57
Mean Absolute Error (MAE) - Stochastic Gradient Descent	90.39	0.958	85.65	69.32
Mean Absolute Percentage Error (MAPE) - Adagrad	89.96	0.957	72.15	55.34
Mean Squared Error (MSE) –Adadelta	78.75	0.838	74.68	61.25
Mean Squared Logarithmic Error (MSE) - RMSprop	80.80	0.823	69.48	72.64
Binary cross entropy loss – Adam	84.29	0.912	71.23	86.42
Sparse Categorical cross entropy loss - Gradient Descent	81.53	0.941	82.89	85.91
Categorical Hinge loss - Stochastic Gradient Descent	87.26	0.941	85.31	71.82
Cosine Similarity –Adagrad	82.38	0.936	74.59	76.38
Logcosh– Adam	**94.67**	**0.988**	**90.11**	**89.71**
Huber cosh–RMSprop	85.10	0.966	81.18	77.68
Poison loss –Adadelta	79.27	0.925	68.90	58.94

The ratio of false classification produced by different methods are measured and presented in Tab. 3. and in Fig.5. The proposed Faster R-CNN model has produced less false ratio compare to other approaches.

Table 3: Performance of False Reduction

Methods	**1000 Images**	**2000 Images**	**3000 Images**
EfficientNet	33	36	40
DenseNet	30	33	37
Resnet	28	31	35
Alexnet	27	30	34

Googlenet	27	30	34
R-CNN	26	29	33
Fast R-CNN	25	28	32
PSPNet	22	25	29
PFN	22	25	29
SSD	21	24	28
YOLO	18	21	25
CenterNet	17	20	24
Faster R-CNN	**13**	**16**	**20**

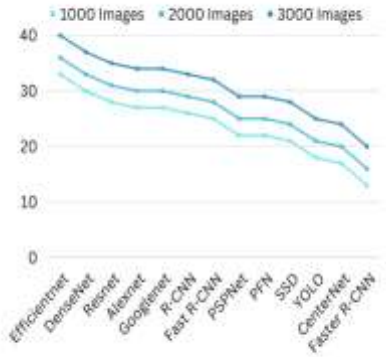


Figure 5: Analysis on false classification ratio

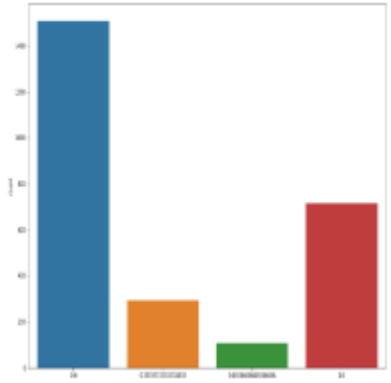


Figure 6: Loss function probability

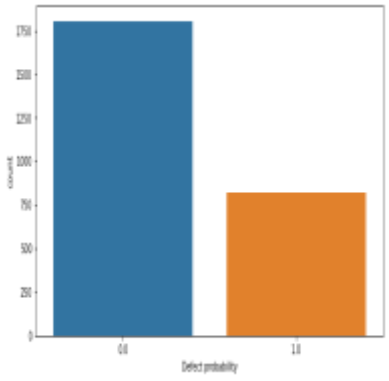
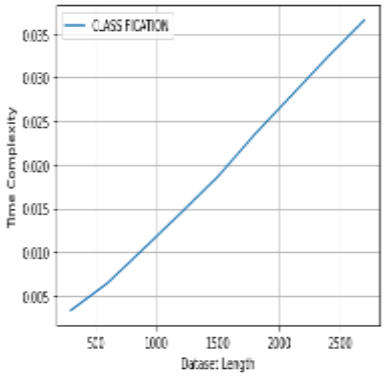


Figure 7: Defect probability



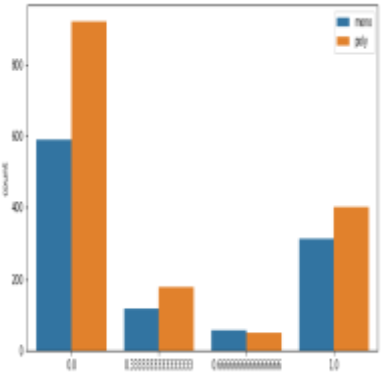


Figure 8: Classification of Faster R-CNN Figure 9: Classification of dataset

Fig.10. shows the results of training accuracy and validation accuracy of the pulmonary nodule dataset and Fig.11 shows the results of training loss and validation loss of the pulmonary nodule dataset

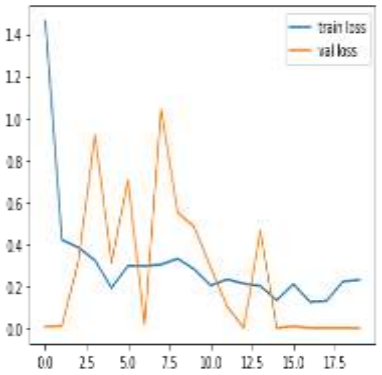


Figure 10: Train acc and Val acc

Figure 11: Train loss and Val loss

CONCLUSION

For this research, automated pulmonary nodule identification was expanded to include micro-nodules, which play a crucial role in the detection of lung cancer at an early stage. To solve the issue of identifying CT image candidates for lung nodules as nodule or non-nodule, we present a novel 15-layer 2D deep CNN architecture called Faster R-CNN. The network is divided into an automated feature extractor comprised of 3 convolutional blocks and a classifier comprised of fully connected layers. To further improve our suggested network's classification results for pulmonary nodules, we use the focus loss function. Our deep learning approach with focused loss is a high-quality classifier, as shown by the assessment on the dataset, which reveals an accuracy of 95.42%, sensitivity of 93.93%, and specificity of 94.73%. This study demonstrates that 3D convolutional neural networks (CNNs) are better to their 2D counterparts when used for analyzing volumetric image data, and that it is crucial to take into account a wide range of contextual information by using numerous models with varying receptive fields. Including such methods in the data preparation phase may further improve the deep CNN's classification accuracy provided here. Finally, utilizing bigger datasets for training might be a promising study route to continue increasing our classification accuracy of lung nodules, since in deep learning, classification accuracy frequently rises with the quantity of data used for training grows. Furthermore, the findings show that the suggested approach and methodology may be very useful in creating trustworthy clinical-decision support systems.

REFERENCE

1. Ding, Jia, et al. "Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2017.
2. Traoré, Abdarahmane, Abdoulaye O. Ly, and Moulay A. Akhloufi. "Evaluating deep learning algorithms in pulmonary nodule detection." 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2020.
3. Tran, Giang Son, et al. "Improving accuracy of lung nodule classification using deep learning with focal loss." Journal of healthcare engineering 2019 (2019).
4. Li, Yanfeng, et al. "Lung nodule detection with deep learning in 3D thoracic MR images." IEEE Access 7 (2019): 37822-37832.
5. Akram, Sheeraz, et al. "Pulmonary nodules detection and classification using hybrid features from computerized tomographic images." Journal of Medical Imaging and Health Informatics 6.1 (2016): 252-259.
6. Naranjo-Torres, José, et al. "Disease and Defect Detection System for Raspberries Based on Convolutional Neural Networks." Applied Sciences 11.24 (2021): 11868.
7. Monkam, Patrice, et al. "Ensemble learning of multiple-view 3D-CNNs model for micro-nodules identification in CT images." IEEE Access 7 (2018): 5564-5576.
8. Dou, Qi, et al. "Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection." IEEE Transactions on Biomedical Engineering 64.7 (2016): 1558-1567.
9. Jacobs, Colin, et al. "Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images." Medical image analysis 18.2 (2014): 374-384.
10. Murphy, Keelin, et al. "A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification." Medical image analysis 13.5 (2009): 757-770.
11. Nguyen, Chi Cuong, et al. "Pulmonary Nodule Detection Based on Faster R-CNN With Adaptive Anchor Box." Ieee Access 9 (2021): 154740-154751.
12. Qin, Yulei, et al. "Simultaneous accurate detection of pulmonary nodules and false positive reduction using 3D CNNs." 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2018.
13. Gu, Yu, et al. "A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning." Computers in biology and medicine 137 (2021): 104806.
14. Sui, Yuan, Ying Wei, and Dazhe Zhao. "Computer-aided lung nodule recognition by SVM classifier based on combination of random undersampling and SMOTE." Computational and mathematical methods in medicine 2015 (2015).
15. Tan, Man, et al. "Pulmonary nodule detection using hybrid two-stage 3D CNNs." Medical physics 47.8 (2020): 3376-3388.
16. Tong, Chao, et al. "Pulmonary nodule detection based on isodata-improved faster rcnn and 3d-cnn with focal loss." ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 16.1s (2020): 1-9.
17. Su, Ying, Dan Li, and Xiaodong Chen. "Lung nodule detection based on faster R-CNN framework." Computer Methods and Programs in Biomedicine 200 (2021): 105866.
18. Wang, Jun, et al. "Pulmonary nodule detection in volumetric chest CT scans using CNNs-based nodule-size-adaptive detection and classification." IEEE access 7 (2019): 46033-46044.
19. Xie, Hongtao, et al. "Automated pulmonary nodule detection in CT images using deep convolutional neural networks." Pattern Recognition 85 (2019): 109-119.
20. Zhu, Wentao, et al. "Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification." 2018 IEEE winter conference on applications of computer vision (WACV). IEEE, 2018.