November/December 2020
ISSN:0193-4120 Page No. 46 – 58

"Review and Analysis of CNN Approach for Lung Cancer Detection and Classification Using Deep Learning"

Dr Ajay Saini[1] Dr. Sumit Gupta[2] Pradeep raika [3] Pankaj Kala [4]
[1]Arya Institute of Engineering and Technology,Jaipur
[2]Parishkar college of Global Excellence Jaipur
[3] Parishkar college of Global Excellence Jaipur
[4] Rukmani Birla Modern High School, Jaipur

Article Info
Volume 83
Page Number: 46 – 58
Publication Issue:
November/December 2020

Article History
Article Received: *25 October 2020*
Revised *22 November 2020*
Accepted: *10 December 2020*
Publication: *31 December 2020*

Abstract— Lung cancer is one of the most frequent malignancies, with 225,000 diagnoses, 150,000 deaths, and a $12 billion annual health-care expenditure. In this paper, we provide a computer-aided diagnostic (CAD) technique for classifying CT scans with unmarked nodules for lung cancer classification. To distinguish lung tissue from the rest of the CT data, thresholding was used as an initial segmentation strategy. The next best lung segmentation came via Thresholding. The original plan was to send segmented CT scans directly into 3D CNNs for categorization, but this proved insufficient. Instead, prospective nodule candidates in CT images were identified using an updated U-Net trained on LUNA16 data (CT scans with labelled nodules). Because the U-Net nodule detection produced a number of false positives when determining whether a CT scan was positive or negative for lung cancer, regions of CTs with segmented lungs where the most likely nodule candidates were located as defined by the U-Net production were fed into 3D Convolutional Neural Networks.

Networks of Neurons (CNNs). 3D CNNs were used to generate the Accuracy O Test Set. Our CAD system surpasses prior literature

November/December 2020
ISSN:0193-4120 Page No. 46 – 58

CAD systems by containing only three crucial phases (segmentation, nodule candidate identification, and malignancy classification), allowing for more effective training and detection, as well as more generalisation for diverse cancers.

Keywords—Lung cancer; computed tomography; deep learning; Convolutional neural networks; segmentation

I. INTRODUCTION

As we know that lung cancer is a one of the worst cancers; nationally, only 17 percent of people diagnosed with lung cancer in the country live five years after diagnosis, although in developed countries, the mortality rate is smaller. A cancer's level corresponds to how deeply it has metastasized. Stages 1 and 2 refer to cancers found in the lungs, and cancers that have spread to other organs refer to the later stages. Present screening procedures, such as CT scans, include biopsies and imaging. Early diagnosis of lung cancer (detection during the earlier stages) greatly increases the probability of survival, but early detection of lung cancer is often more difficult when less symptoms are present. [1].

In patient CT scans of lungs with and without early stage lung cancer, our job is a binary classification question to diagnose the existence of lung cancer. To create an accurate classifier, we aim to use techniques from computer vision and deep learning, particularly 2D and 3D convolution neural networks. An correct classification of lung cancer could accelerate and reduce the cost of screening for lung cancer, encouraging more universal screening.

Early identification and survival change. The aim is to build a computer-aided diagnostic (CAD) system that involves patient chest CT scans and outputs as an input, whether the patient has lung cancer or not. [2].

Although this job sounds simple, in the haystack dilemma it is really a needle. The CAD device will have to detect the presence of a small nodule (< 10 mm in diameter for early stage cancers) from a large 3D lung CT scan to determine whether or not a patient has early-stage cancer (typically around 200 mm 400 mm 400 mm). An example of an early stage nodule of lung cancer seen in a 2D slice of a CT scan is given in Fig. 1. In addition, a CT scan is packed with noise from nearby tissues, bone, air, so this noise will first have to be preprocessed for the CAD systems search to be successful. Therefore, image preprocessing, nodule candidate identification, malignancy classification are our classification pipeline.

In this article, we use systematic preprocessing procedures to extract specific nodules in order to increase the precision of lung cancer diagnosis. In addition, we conduct CNN end-to-end testing from scratch in order to understand the full capacity of the neural network, i.e. to acquire discriminatory characteristics. A dataset containing lung nodules from more than 1390 low dose CT scans is used for detailed experimental assessments.

Figure : 2D CT scan slice containing a small (5mm) early stage lung cancer nodule.

II RELATED WORK

Lung cancer is now the leading cancer cause of death in industrialized world [1]. A

November/December 2020
ISSN:0193-4120 Page No. 46 – 58

dismally low cure rate largely reflects the propensity of lung cancer to present as clinically advanced tumors: most lung cancers are discovered late during their clinical course, by which time the options for effective therapeutic intervention are limited. It is critically important that primary care physicians recognize the potential significance of such a lesion and know how to proceed with investigation. By raising an early red flag, lung cancer can then be identified at a time when it is far more curable. Recently, attempts have been made in china to apply helical computed tomography (CT) to lung cancer screening [2].

When the nodules are less profuse, it may be difficult to be distinguished from the vessels. Radiographic detection of calcification within a solitary pulmonary nodule usually indicates benignity. Not only can high resolution spiral CT scans identify very small tumors that could not be detected with less sophisticated technology, but machines also operate much more efficiently and with greater accuracy than earlier [3]. The aim of image processing and image segmentation in this paper is to auto-detecting nodules from the lung CT image [4-5]. Therefore, earlier and more certain detection with more effective screening methods can be expected to improve cure rates.

When the nodules are less profuse, it may be difficult to be distinguished from the vessels. Radiographic detection of calcification within a solitary pulmonary nodule usually indicates benignity. Not only can high resolution spiral CT scans identify very small tumors that could not be detected with less sophisticated technology, but machines also operate much more efficiently and with greater accuracy than earlier . The aim of image processing and image segmentation in this paper is to auto-detecting nodules from the lung CT image . Therefore, earlier and more certain detection with more effective screening methods can be expected to improve cure rates.

The paper presents that to detect tiny nodules from CT image, which may present the characteristic of early lung cancer and proposes an algorithm that incorporates newer imaging and diagnostic methods to facilitate the evaluation and management of removing the pulmonary nodules.

By far, lung cancer is the prominent cause of cancer deaths for both men and women around the world. In 2018, statistics from WCRF (Worldwide Cancer Research Fund) showed that of 2.09 million people are diagnosed with this disease (21.7% of all cancer diagnoses), 1.76 million people have died (36.5% of all cancer deaths) [1]. The survival rate increases if the cancer is detected in its early stages; however, approximately 80% of patients are accurately diagnosed during the intermediate or advanced stage of cancer [2]. Hence, the survival rate so low.

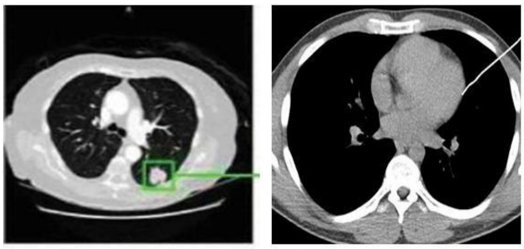
In the process of diagnosing potential cancerous lung nodules, the radiologists have two interconnected tasks: to detect an abnormality and categorize it as representing a specific type of disease. Diagnosticating is an arduous process, but pulmonary radiologists have a high degree of accuracy in diagnosis. In a study, the accuracy rate of the radiologists in the detection of lung cancer using CT scans revealed around 0.79%. Nonetheless, there remain problems in disease detection. Some of these problems consist of the miss rate for the detection of small pulmonary nodules, the

November/December 2020
ISSN:0193-4120 Page No. 46 – 58

detection of minimal interstitial lung disease, and the detection of changes in pre-existing interstitial lung disease. These problems are hard to overcome, even with high levels of clinical skills and experience.

For years CADx (Computer-Aided Diagnosis) systems have helped radiologists for early diagnose and increase of care services earlier, faster, and with higher accuracy. Traditionally, CADx systems identify tumors and different diseases of similar nature with complicated image processing techniques and segmentation methods. Consequently, making this process of extracting low-level features strenuous and intricating. However, in the recent research literature, machine learning and deep learning techniques (a subset of artificial intelligence) have been used to diagnose and classify cancer. Due to the nature of these methods, extracting low-level to high-level features from large amounts
of datasets and classifying cancer of different types has successfully proven to be somewhat accurate .

In this paper, by taking into consideration machine learning and deep learning techniques as some of the state-of-the-art methods in terms of automatic feature extraction and automatic detection, different algorithms and architectures of these techniques, specifically convolutional neural network, were investigated.



a. Nodules segmented from the primary chest CT scan image.   b. Pulmonary nodule with CT image.

There is a CT scan image of nodules segmentation, The segmentation of lung nodules is an important part of two different systems that are related to the prevention and diagnosis of lesions. The first system is the computer-aided diagnosis (CAD) system, which aims to improve the ability to detect the nodules and can help to classify the nodule as malignant or benign. Most of the patients with non-small-cell carcinoma (NSCLC) are obsolete because of locally improved diseases or remote metastases and thus radiation treatment remains the primary choice treat them. Actinotherapy is effective for treating carcinogenicity can pose the risks that require to be changed keep with the characteristics of each patient. A comparative study analyses how the variable affects the formation procedures of a Deep Neural Network to identify carcinoma images. It's difficult to classify for' heavier' images with DNN. While images of the form CT are used primarily in medical imaging, unnecessary artifacts are going to be created
First, cancer is a group of abnormal cells that grows until they spread into neighboring tissues. These cells can grow almost everywhere in the body. The human body is made out of tens of trillions of cells. These cells, to survive and spread their genes, they divide into new cells. So when old cells get old, they die, and new ones take their place.

November/December 2020
ISSN:0193-4120 Page No. 46 – 58

However, with cancer cells, this process of cell reproduction stage does occur. They are abnormal; in the sense that those cells contain errors or mutations in their genes. The old abnormal cells survive when they should have died, and new ones are created even though not needed, and this is how tumors grow.

Tumors can be malignant or benign. The difference between the two is that malignant tumors advance in nearby tissue, making it possible for cancer to travel, through blood or lymph system to distant organs, and cause harm far from the original location (organ/ tissue). On the other hand, benign tumors do not advance in other organs or tissues. Even though they can get big, once removed, they do not possess any harm to the patient (unless the location of the tumor is in the brain) .

Lungs are two cone-shaped breathing organs found in the chest. The lung's primary role is to bring oxygen in and release carbon dioxide when breathing out. Based on the terminology explained above, lung cancer is the rapid expansion of abnormal cells in one or both lungs. These cells are not specialized to do a specific function, as healthy cells. Instead, they cluster into a tumor and interfere with the lung's sole task as a part of the respiratory system. Based on the size of the cell in which cancer starts, there exist two types of lung cancer: small cell lung cancer (SCLC) and nonsmall cell lung cancer (NSCLC).

Moreover, NSCLC is going to be the topic of exploration for this thesis, and it consists of three groups that are identified based on the initial location of cancer. First, the most common type of lung cancer that emerges in the glandular cells on the outer part of the lung is called adenocarcinoma. NSCLC can also start "in flat, thin cells called squamous cells. These cells line the bronchi, which are the large airways that branch off from the windpipe (trachea) into the lungs, as it can be seen in the picture below. This type of cancer is called squamous cell carcinoma of the lung. Large cell carcinoma is another type of nonsmall cell lung cancer, but it is less common. There are also several rare types of nonsmall cell lung cancer. These include sarcoma and sarcomatoid carcinoma."[9] Below, it is presented the most common type of cancer, adenocarcinoma.

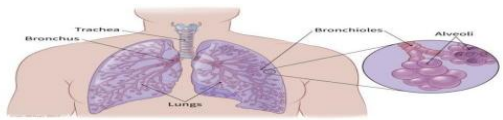


Figure : Non-small lung cancer - Adenocarcinoma

SCLC often starts in cells in the center of the lungs. The main types of small cell lung cancer are "small cell carcinoma and combined small cell carcinoma (mixed tumor with squamous or glandular cells)" .

For both types of cancer, the likelihood to cause lung cancer is increased by smoking (including cigars, tobacco, pipes, both past, and present). All studies and official websites point out the fact that smoking is the dominant risk factor for this disease. The longer the period of smoking, the earlier a person starts smoking, the more often the person smokes, the higher the risk. Additionally, other risk factors include exposure to secondhand smoke, asbestos, arsenic, chromium, beryllium, nickel, soot, tar, radiation, living in polluted air, having a family history of lung cancer, being infected by human immunodeficiency virus (H.I.V.), and many

TEST Engineering & Management
November/December 2020
ISSN:0193-4120 Page No. 46 – 58

more. [10].

In Nikita Pandey, Sayani Nandy Proposed A novel approach for detection of cancerous cells from Lungs CT scan images. This work proposes a method to detect the cancerous cells effectively from the CT scan images by reducing the detection error made by the physicians‟ naked eye for medical study based on Sobel edge detection and label matrix. Sobel operator helps to find the edges in an image; it does so by finding the image gradient. Image gradient is the change in the intensity of the image. Prof. Samir Kumar Bandyopadhyay provides a method using Computer Aided Diagnosis System (CAD) for detection of edges from CT images of lung for detection of diseases. Fatm Taher, Naoufel Werghi and Hussain Al-Ahmad [deals with filtering thresholding algorithm for extracting the sputum cell from the raw sputum image for lung cancer early detection. Qinghua Ji, Ronggang Shi This paper presents a new method of image segmentation using watershed transformation. To use morphological opening and closing operations to process the gradient image aim to eliminate the over segmentation areas, and reconstruction of the morphological gradient can maintain the shape of gradient image. The proposed method can simplify gradient image while maintaining the contours of the exact location of the dividing line, eliminating the root causes of the phenomenon have been split.

Jakimovski and Davcev proposed the Double convolutional deep neural network (DCDNN) for carcinoma stage prediction. within the training of both the CDNN and thus the regular CDNNs, they used CT (CT) scans. These topologies are experimented against images of pulmonary cancer to assess the stage of cancer therein topologies can predict carcinoma.

Rodrigues et al .suggested the approach of systematical co-occurrence matrix (SCM) classifying nodes as malignant nodules or benign nodules. Data on nodule locations and malignancy rates are given within the X-radiation analysis of the pulmonary imaging and image database tools initiative. The SCM is implemented in four filters, specifically, medium, Laplace, Gaussian and Sobel, both grayscale and Hounsfield.

Chung et al. proposed some way of lung segmentation to chop back the matter of the juxta-pleural nodule, a preferred challenge within the applications. Initially, they used Chan-Vese (CV) model for active contours and followed the Bayesian approach supported theresults of a CV model that predicts lung image in an earlier frame or the neighboring image on the premise of the segmented lung contour.The false positives were removed by the concave detection of the points. Eventually, the lung contour was changed by applying candidates from the last word nodule to the results of the CV model. nodule that would support any computer- aided diagnosis (CAD) system using lung segmentation.

To beat the above survey, during this paper, the Adaptive Hierarchical Heuristic Mathematical Model (AHHMM) to predict carcinoma supported Deep learning using CAT images has been proposed. carcinoma is an abnormal cell disease that multiplies and becomes a tumor. The lungs, or the lymph fluid that covers the lung tissue, can hold cancer cells removed from the lungs. Deep learning is best known than conventional techniques for image classification.

Recently, deep artificial neural networks have been applied in many applications in

November/December 2020
ISSN:0193-4120 Page No. 46 – 58

pattern recognition and machine learning, especially, Convolutional neural networks (CNNs) which is one class of models .

HENG YU.at all	The technology that is currently developing as an alternative to a conventional method for early warning observation is image processing.	72% Accuracy
Gawade Prathamesh Pratap at.all	Image procssingwith Mat Lab	82% Accuracy
Lakshmanaprabu S.K	After the feature extraction, dimensionality reduction technique Method, feature reduction is utilized that is LDA.	80% Accuracy

Another approach of CNNs was applied on ImageNet Classification in 2012 is called an ensemble CNNs which outperformed the best results which were popular in the computer vision community. There has also been popular latest research in the area of medical imaging using deep learning with promising results.

HENG YU proposed in year 2020 the technology that is currently developing as an alternative to a conventional method for early warning observation is image processing. Data set Genreated using CNN.

Gawade Prathamesh Pratap in 2016 proposed a method to detect lung cancer using Image processing with watershed algorithm

Lakshmanaprabu S.K in 2018 proposed a method to find lung cancer using Optimal Deep Neural Network
(ODNN) and Linear Discriminate Analysis (LDA).

Yutong algorithm utilizes a GLCM-based surface descriptor, a Fourier-shape descriptor to portray heterogeneity of nodules and a DCNN to train the features of node

Comparison list

R. Golan proposed a framework that trains the weights of the CNN by a back propagation to detect lung nodules in the CT image sub-volumes. This system achieved sensitivity of 78.9% with 20 false positives, while 71.2% with 10 FPs per scan, on lung nodules that have been annotated by all four radiologists

Philippe Lambin proposed a framework Convolutional neural networks have achieved better than Deep Belief Networks in current studies on benchmark computer vision datasets. The CNNs have attracted considerable interest in machine learning since they have strong representation ability in learning useful features from input data in recent years.

Data and Methodology

Lung cancer is a disease of abnormal cells multiplying and growing into a tumour. Cancer cells can be carried away from the lungs in blood, or lymph fluid that surrounds lung tissue. Lymph flows through lymphatic vessels, which drain into lymph nodes located

November/December 2020
ISSN:0193-4120 Page No. 46 – 58

in the lungs and in the centre of the chest. Lung cancer often spreads toward the centre of the chest because the natural flow of lymph out of the lungs is toward the centre of the chest. Metastasis occurs when a cancer cell leaves the site where it began and moves into a lymph node or to another part of the body through the blood stream [1]. Cancer that starts in the lung is called primary lung cancer. There are several different types of lung cancer, and these are divided into two main groups: Small cell lung cancer and non-small cell lung cancer which has three subtypes: Carcinoma, Adenocarcinoma and Squamous cell carcinomas.

The rank order of cancers for both males and females among Jordanians in 2008 indicated that there were 356 cases of lung cancer accounting for (7.7 %) of all newly diagnosed cancer cases in 2008. Lung cancer affected 297 (13.1 %) males and 59 (2.5%) females with a male to female ratio of 5:1 which Lung cancer ranked second among males and 10th among females [2]. Figure 1 shows a general description of lung cancer detection system that contains four basic stages. The first stage starts with taking a collection of CT images (normal and abnormal) from the available Database from IMBA Home (VIA-ELCAP Public Access) [3]. The second stage applies several techniques of image enhancement, to get best level of quality and clearness. The third stage applies image segmentation algorithms which play an effective rule in image processing stages, and the fourth stage obtains the general features from enhanced **segmented image which gives indicators of** normality or abnormality of images.

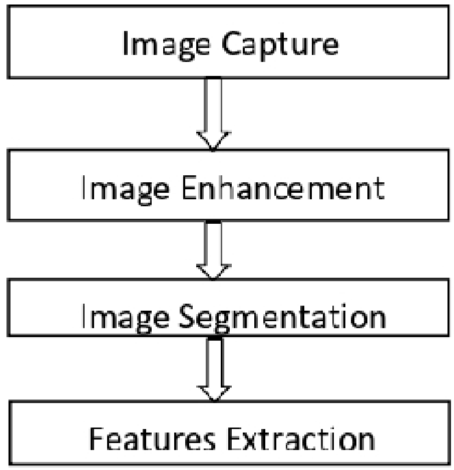


Figure : Lung cancer image processing stages

Lung cancer is the most dangerous and widespread cancer in the world according to stage of discovery of the cancer cells in the lungs, so the process early detection of the disease plays a very important and essential role to avoid serious advanced stages to reduce its percentage of distribution. The aim of this research was to detect features for accurate images comparison as pixels percentage and mask-labelling.

A Computer Aided Detection System (CAD) is one of the principal research streams in medical imaging and diagnostic radiology. A well-developed CAD helps in processing image for detection and extraction of abnormalities and also aids in classification of image features between normal and abnormal. In most of the cases CAD offers a very useful second opinion when radiologist examine patient at cancer CT screenings. A CAD system is instrumental in reducing the number of false negative diagnosis [3].The success of a CAD system is measured in terms of accuracy in diagnosis, speed and its degree

November/December 2020
ISSN:0193-4120 Page No. 46 – 58

of automation.

Amid various non-invasive medical imaging modalities such as X-RAY, Magnetic Resonance Imaging (MRI), Positron Emission tomography (PET), PET-CT, the Computed Tomography (CT) is widely used for spotting and diagnosing lung nodules. Since the invention in 1972 by G.N.Hounsfiled, the advancement in CT imaging has contributed in a great deal to the field of Pathology. CT has been influential in increasing the survival rate by diagnosing the life threatening ailments, mainly the cancer [4]. Recent CT scanners exhibit isotropic acquisition of whole chest with very high resolution within a single breath hold.

A typical lung nodule detection scheme is shown in Figure.1, which comprises of the following sequences of processing steps. Image Acquisition, Image Enhancement, Lung Parenchyma Segmentation, Candidate nodule Detection, False Positive Reduction and Nodule Classification. These steps are explained in the following sections.

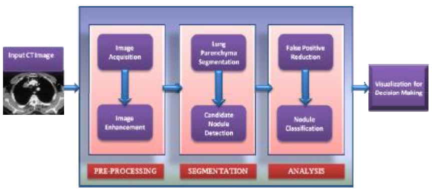


Figure .: Processing steps in typical Lung nodule detection scheme

Medical diagnoses - Machine learning techniques are used to detect and diagnose cancerous tissue. The financial industry and trading - many banks use ML in fraud detection and credit check. As can be seen from above, the variety of problems these techniques can solve is infinite. [12]

Primary dataset is the patient lung CT scan dataset from Cancer Hospital. The dataset contains labeled data , which is divide into training set of size 900, and test set of size 575. For each patient, the data consists of CT scan data and a label (0 for no cancer, 1 for cancer). Note that the dataset does not have labeled nodules. For each patient, the CT scan data consists of a variable number of images (typically around 100- 400, each image is an axial slice) of 512x512 pixel . The slices are provided in DICOM format. Around 70% of the provided labels in the dataset are 0, so we used a weighted loss function in our malignancy classifier to address this imbalance.

Dataset alone proved to be inadequate to accurately classify the validation set; we also used the patient lung CT scan dataset with labeled nodules from the Lung Nodule Analysis 2016 (LUNA16) Challenge to train a U-Net for lung nodule detection. The LUNA16 dataset contains labeled data for 888 patients, which we divided into a training set of size 710 and a validation set of size 178. For each patient, the data consists of
CT scan data and a nodule label (list of nodule center coordinates and diameter). For each patient, the CT scan data consists of a variable number of
images (typically around 100-400, each image is an axial slice) of 512 × 512 pixels.

LUNA16 data can be used to train a U-Net for nodule detection, one of the phases in our classification pipeline. The problem is to accurately predict a patient's label ('cancer' or 'no cancer') based on the patient's Kaggle lung CT scan. We will use accuracy, sensitivity, specificity, and AUC of the ROC to evaluate our CAD system's performance on

November/December 2020
ISSN:0193-4120 Page No. 46 – 58

the test set.

Typical CAD systems for lung cancer have the following Step: image preprocessing, detection of cancerous nodule candidates, nodule candidate false positive reduction, malignancy prediction for each nodule candidate, and malignancy prediction for overall CT scan .

These pipelines have many phases, each of which are computationally expensive and require well-labeled data during training. For example, the false positive reduction phase requires a dataset of labeled true and false nodule candidates, and the nodule malignancy prediction phase requires a dataset with nodules labeled with malignancy True/False labels for nodule candidates and malignancy labels for nodules are sparse for lung cancer, and may be nonexistent for some other cancers, so CAD systems that rely on such data would not generalize to other cancers. In order to achieve greater computational efficiency and generalizability to other cancers, the proposed CAD system has shorter pipeline and only requires the following data during training: a dataset of CT scans with true nodules labeled, and a dataset of CT scans with an overall malignancy label. State-of-the-art CAD systems that predict malignancy from CT scans achieve AUC of up to 0.83 . However, as mentioned above, these systems take as input various labeled data that is not used in this framework. The main goal of the proposed system is to reach close to this performance.

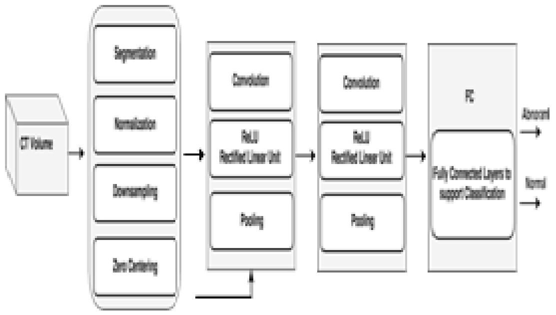


Figure : 3D Convolutional neural networks architecture

The CAD system starts with preprocessing the 3D CT scans using segmentation, normalization, down sampling, and zero-centering. The initial approach was to simply input the preprocessed 3D CT scans into 3D CNNs, but the results were poor. So an additional preprocessing was performed to input only regions of interests into the 3D CNNs. To identify regions of interest, a U-Net was trained for nodule candidate detection. Then input regions around nodule candidates de- tected by the U-Net was fed into 3D CNNs to ultimately classify the CT scans as positive or negative for lung cancer. The overall architecture is shown in Fig. 2, all details of layers will be described in the next sections.

A. Preprocessing and Segmentation

For each patient, pixel values was first converted in each image to Hounsfield units (HU), a measurement of radio density, and 2D slices are stacked into a single 3D image. Because tumors form on lung tissue, segmentation is used to mask out the bone, outside air, and other substances that would make data noisy, and leave only lung tissue information for the classifier. A number of segmentation approaches were tried, including thresholding, clustering (Kmeans and Meanshift), and Watershed. K-means and Meanshift allow very little super- vision and did not produce good qualitative results. Watershed produced the best qualitative

November/December 2020
ISSN:0193-4120 Page No. 46 – 58

results, but took too long to run to use by the deadline. Ultimately, thresholding was used.

After segmentation, the 3D image is normalized by applying the linear scaling to squeeze all pixels of the original un-segmented image to values between 0 and 1. Spline interpolation down samples each 3D image by a scale of 0.5 in each of the three dimensions. Finally, zero-centering is performed on data by subtracting the mean of all the images from the training set.

CONCLUSION

In the present Review paper, the results of the SMS with SLR on lung cancer based on deep learning are presented. Thirty-two articles were reviewed and various methods for diagnosing lung cancer were presented with deep learning, which resulted in different levels of accuracy and sensitivity. This article examines the universities and countries that have contributed to the publication of related articles. Most of the publications in this field are attributed to the IEEE. China, the United States, Australia, the United Kingdom, Japan, and Palestine had the largest number of articles, respectively.

Current article contains limitations since searches were done in the title of the articles only and by expanding this search, other articles can be accessed.

Focus of this study has been on English publications only and has not done any search on non- English papers.

Lung cancer is a dangerous disease, and early-stage detec- tion is therefore necessary. This paper presents deep learn- ing assisted Adaptive Hierarchical Heuristic Mathematical Model (AHHMM) to predict lung cancer on computed tomography images. This paper uses the Modified K-means algorithm and CNN Algorithm to pre-classify pictures into slices of images in the same image, where the DNN will concentrate on the image classification of images in similar images. The next thing is the convolution layer with filtered edges to scan for lung cancer thoroughly. Then, estimating the weighted mean function which replaces the pixel utilizing the cumulative distribution and likelihood distribution method improved the images quality. The injured portion has been segmented by a pixel-like value measurement after the image has been improved. Based on the similarity calculation, spectral- related features has been extracted. The proposed AHHMM system predicts computed tomography scanning images of lung cancer successfully. At the end of the system, you can say that the system is satisfying its desires. The findings of the evaluation showed that around 80% of the images has correctly identified. Such results show that DNN is useful in cyst diagnosis for classifying lung cancer. Hybridized Heuristic Mathematical Model will be implemented in future for predicting the lung cancer at earlier stage.

We also read a deep Convolutional neural network (CNN) architecture to detect nodules in patients of lung cancer and detect the interest points using U-Net architecture. This step is a preprocessing step for 3D CNN. The deep 3D CNN models performed the best on the test set. While we achieve state-of-the-art performance AUC of 0.83, we perform well considering that we use less labeled data than most state- of-the-art CAD systems. As an interesting observation, The first layer is a preprocessing layer for segmentation using different techniques. Threshold, Watershed, and U-Net are used to identify the nodules of patients.

November/December 2020
ISSN:0193-4120 Page No. 46 – 58

The network can be trained end-to-end from raw image patches. Its main requirement is the availability of training database, but otherwise no assumptions are made about the objects of interest or underlying image modality.

In the future, it could be possible to extend our current model to not only determine whether or not the patient has cancer, but also determine the exact location of the cancerous nodules. The most immediate future work is to use Watershed segmentation as the initial lung segmentation. Other opportunities for improvement include making the network deeper, and more extensive hyper parameter tuning. Also, we can make model parameters at best accuracy, but perhaps we could have saved at other metrics . Other future work include extending our models to 3D images for other cancers with Highest Accuracy . The advantage of not requiring too much labeled data specific to our cancer is it could make it generalizable to other cancers.

REFERENCES

[1] W.-J. Choi and T.-S. Choi, "Automated pulmonary nodule detection system in computed tomography images: A hierarchical block classifi- cation approach," Entropy, vol. 15, no. 2, pp. 507–523, 2013.

[2] A. Chon, N. Balachandar, and P. Lu, "Deep Convolutional neural networks for lung cancer detection," tech. rep., Stanford University, 2017.

[3] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision.," in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), pp. 253–256, IEEE, 2010.

[4] K. Alex, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems 25 (NIPS 2012) (F. Pereira, C. J. C. Burges,L. Bottou, and K. Q. Weinberger, eds.), pp. 1097–1105, 2012.

[5] H. Suk, S. Lee, and D. Shen, "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis," NeuroImage, vol. 101, pp. 569–582, 2014.

[6] G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao, and D. Shen, "Unsupervised deep feature learning for deformable registration of mr brain images.," Medical Image Computing and Computer-Assisted Intervention, vol. 16, no. Pt 2, pp. 649–656, 2013.

[7] Y. Xu, T. Mo, Q. Feng, P. Zhong, M. Lai, and E. I. Chang, "Deep learning of feature representation with multiple instance learning for medical image analysis," in IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, pp. 1626–1630, 2014.

[8] D. Kumar, A. Wong, and D. A. Clausi, "Lung nodule classification using deep features in ct images," in 2015 12th Conference on Computer and Robot Vision, pp. 133–138, June 2015.

[9] Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, "Chest pathology detection using deep learning with non-medical train- ing," Proceedings - International Symposium on Biomedical Imaging, vol. 2015-July, pp. 294–297, 2015.

[10] W. Sun, B. Zheng, and W. Qian, "Computer aided lung cancer diagnosis with deep learning algorithms," in SPIE Medical Imaging, vol. 9785, pp. 97850Z–97850Z, International Society for Optics and Photonics, 2016.

[11] J. Tan, Y. Huo, Z. Liang, and L. Li, "A comparison study on the effect of false positive reduction in deep learning based detection for juxtapleural lung nodules: Cnn vs

dnn," in Proceedings of the Symposium on Modeling and Simulation in Medicine, MSM '17, (San Diego, CA, USA), pp. 8:1–8:8, Society for Computer Simulation International, 2017.

[12] R. Golan, C. Jacob, and J. Denzinger, "Lung nodule detection in ct im- ages using deep Convolutional neural networks," in 2016 International Joint Conference on Neural Networks (IJCNN), pp. 243–250, July 2016.

[13] Kaggle, "Data science bowl 2017." https://www.kaggle.com/c/data- science-bowl-2017/data, 2017.

[14] LUNA16, "Lung nodule analysis 2016." https://luna16.grand- challenge.org/, 2017.

[15] M. Firmino, A. Morais, R. Mendoa, M. Dantas, H. Hekis, and R. Valentim, "Computer-aided detection system for lung cancer in computed tomography scans: Review and future prospects," BioMedical Engineering OnLine, vol. 13, p. 41, 2014.

[16] S. Hawkins, H. Wang, Y. Liu, A. Garcia, O.Stringfield, H. Krewer,Q. Li, D. Cherezov, R. A. Gatenby, Y. Balagurunathan, D. Goldgof, M. B. Schabath, L. Hall, and R. J. Gillies, "Predicting malignant nodules from screening ct scans," Journal of Thoracic Oncology, vol. 11, no. 12, pp. 2120–2128, 2016.

[17] M. S. AL-TARAWNEH, "Lung cancer detection using image processing techniques," Leonardo Electronic Journal of Practices and Technolo- gies, pp. 147–158, June 2012.

[18] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional net- works for biomedical image segmentation," CoRR, vol. abs/1505.04597, 2015