International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES),(UGC APPROVED)
Impact Factor: 5.22 (SJIF-2017),e-ISSN:2455-2585
Research Symposium on
"Advancements in Engineering, Science, Management, and Technology"
Volume 5, Special Issue 04, April-2019.

Lung Cancer Detection on CT Scan Medical Image

Endalew Simie Muche[1], Dr. Mandeep Kaur[2]

[1]*School of Engineering and Technology, Department of Computer Science and Engineering Sharda University,*
[2]*School of Engineering and Technology, Department of Computer Science and Engineering Sharda University,*

Abstract— Currently, lung cancer is one of the deadly diseases in the world. Various investigations are underway to reduce this disease. The lung disease imaging system produces large amounts of medical images (such as CT scan image, X-ray, MRI images etc.) that containing vast amounts of information. It is very difficult for doctors and clinicians to interpreting and identifying this information accurately. Medical image analysis has become a major research area in the healthcare centre. The purpose of lung cancer detection system is able to detect and provide reliable information to doctors and clinicians from a medical image. To minimize this problem, many systems have been proposed by using different image processing techniques, machine learning, and deep learning techniques. Machine learning is one of the most popular techniques that used in computer-aided diagnosis CAD) and medical image analysis in the classification of the objects such as lesions to certain lesions (example lesions or non-lesion, and malignant or benign) based on input features obtained from segmented objects. The recent advent of deep learning has replaced many other machine learning methods, because it avoids the creation of hand-engineering features, thus removing a critical source of error from the process. Deep learning is a breakthrough in machine learning techniques that have overcome the field of pattern recognition and computer vision research areas. Deep learning provides the machine learning that high-level abstraction features from a medical image but not use handcrafted features. Deep learning is providing exciting solutions for medical image analysis problems and is seen as a key method for future applications. In the medical image analysis, there are various application areas including detection, segmentation, classification, and computer-aided diagnosis. The main aim of the paper is reviewing the recent literature and finding the gaps of the proposed system that related to lung cancer detection in medical image and providing the future work.

Keywords— Medical Image, Lung Cancer, Cancer detection, Segmentation, CAD

I. INTRODUCTION

Lung cancer is one of the dangerous diseases in the world that taking human life rapidly. The death of the people is increasing exponentially because of lung cancer. In order to reduce the disease and save a human's life, the automated system is needed. Lung cancer is one of the widespread diseases and the leading causes of cancer-related death worldwide. As reported in global cancer statistics in 2012, nearly 1.83 million new cases of lung cancer occurred and the estimated deaths are over 1.5 million [20].

Medical imaging system produces a large amount of medical image containing the relevant information related to diseases. The medical image is one of the major research areas in medical problem domains to detect and diagnosis numerous disease. Medical image analysis is used to analyzing and solving medical problems by using different medical image analysis techniques to extract relevant information or knowledge from medical images. There are various medical imaging modalities such as Magnetic resonance image (MRI), Computed Tomography (CT scan), Positron emission tomography (PET), mammography, X-ray, ultrasound and so on, that used for early detection and diagnosis of disease [12]. But one of the best imaging technique is Computed Tomography (CT scan) imaging are reliable for lung cancer diagnosis because it can disclose every suspected and unsuspected lung cancer nodules [1]. Recently, to assist radiologists and doctors detect diseases accurately computer Aided Diagnosis has become a supplement and promising tool [1]. There are various ways that used to Lung cancer detection such as image processing, pattern recognition, and artificial neural network to implement the computer-aided diagnosis (CAD).

Image processing techniques have a major role in medical image analysis and computer-aided diagnosis (CAD). Many researcher have been proposed and implement several systems for supporting radiologist and clinicians. Medical image processing steps implemented on various research papers like preprocessing, edge detection, morphological processing, and feature extraction. In lung cancer detection preprocessing is a processing of reading lung cancer Images (CT scan, X-ray, MRI etc.) that performed. That medical image is gathered from the different medical database on the internet. However many researchers use CT scan images in the DICOM (Digital Imaging and Communications in Medicine). In the image preprocessing image enhancement is also performed.

Interpretation of the resulting images requires sophisticated image processing methods that enhance visual interpretation, and image analysis methods that provide automated or semi-automated tissue detection, measurement, and characterization. In general, multiple transformations will be needed in order to extract the data of interest from an image, and a hierarchy in the processing steps will be evident, e.g. enhancement will precede restoration, which will precede analysis, feature extraction and classification [5]. Image processing technique has a great role in medical image analysis and it makes the classification processes more accurate. This image process is overwhelming by machine learning techniques on the detection phase due to the detection problem. In machine learning, one develops and studies methods that give computers the ability to solve problems by learning from experiences. The goal is to create mathematical models that can be trained to produce useful outputs when fed input data. Machine learning models are provided experiences in the form of training data and are tuned to produce accurate predictions for the training data by an optimization algorithm [6]. It has a great role in image-based disease detection, diagnosis, and disease treatments in order to support radiologist and clinicians. Machine learning has several algorithms used for classification and clustering related problems.

In recent years Machine learning techniques have played important role in medical field like medical image processing, computer-aided diagnosis, image interpretation, image registration, image segmentation, image-guided therapy, image retrieval, and analysis. Techniques of ML extract information from the images and represents information effectively and efficiently. The ML and AI facilitate and assist doctors that they can diagnose and predict accurate and faster the risk of diseases and prevent them in time. These techniques enhance the abilities of doctors and researchers to understand how to analyze the generic variations which will lead to disease. These techniques composed of conventional algorithms without learning like Support Vector Machine (SVM), Neural Network (NN), KNN etc. and deep learning algorithms such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short term Memory (LSTM), Extreme Learning Model (ELM), Generative Adversarial Networks (GANs) etc. [11]. Machine learning is algorithms are limited in processing the natural images in their raw form, time-consuming, based on expert knowledge and requires a lot of time for tuning the features.

Deep learning algorithms are fed with raw data, automatic features learner and fast. These algorithms try to learn multiple levels of abstraction, representation, and information automatically from a large set of images that exhibit the desired behavior of data. Although automated detection of diseases based on conventional methods in medical imaging has been shown significant accuracies around for decades, new advances in machine learning techniques have ignited a boom in deep learning. Deep learning-based algorithms showed promising performance as well as speed in different domains like speech recognition, text recognition, lips reading, computer-aided diagnosis, face recognition, drug discovery. Now a day deep learning algorithm has got great interest in each and every field and especially in medical image analysis due to the representation of multiple levels of abstraction and extraction of features from large dataset automatically.

More specific uses of deep learning in the medical field are segmentation, diagnosis, classification, prediction, and detection of various anatomical regions of interest (ROI). Compared to traditional machine learning, deep learning is far superior as it can learn from raw data and has multiple hidden layers which allow it to learn abstractions based on inputs. The key to deep learning capabilities lies in the capability of the neural networks to learn from data through general purpose learning procedure [13].

II. PAGE LAYOUT

An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.

Many lung cancer detection system and diagnosis system have been proposed to the help of radiologist and clinician to detect and classifies the disease with a better result by using different approaches of image processing, machine learning, and deep learning. We summarize these systems based on the methods they adopt.

A. Image processing approaches

Image processing technique has a great role in medical image analysis and it makes the classification processes more accurate. Image processing techniques have been greatly explored in nodule classification for lung CT images. Numerous studies adopted segmentation, morphological operations, and contour filter approaches for better nodule detection [2]. By using those approaches several researchers have proposed and implement lung cancer detection system with good accuracy. Neelima Singh et al. [8] has been proposed lung cancer detection systems by using image processing techniques. The system is used image preprocessing, and after preprocessing of an image, a canny filter is used for Edge Detection. Superpixel Segmentation has been used for segmentation and Gabor filter is used for Denoising the medical images.

B. Machine learning approaches

Machine learning is studying methods that give computers the ability to solve problems by learning from experiences. The goal is to create mathematical models that can be trained to produce useful outputs when fed input data. Machine learning models are provided experiences in the form of training data and are tuned to produce accurate predictions for the training data by an optimization algorithm [16].

In recent years Machine learning techniques have played important role in medical field like medical image processing, computer-aided diagnosis, image interpretation, image registration, image segmentation, image retrieval, and analysis. These techniques composed of conventional algorithms without learning like Support Vector Machine (SVM), Neural Network (NN), and KNN etc.

Suren Makajua et al. [9] proposed a model that detect the cancerous nodule form CT scan image by using watershed segmentation for detection. In this proposed system Gaussian filter method is implemented in the pre-processing stages and using SVM for classification of the nodule as Malignant or benign. Qing. W et al. [8] proposed a system that detects small cell lung cancer (SCLC) form computed tomography (CT) scan images. The system proposed a novel neural-network based algorithm, refers to an entropy degradation method (EDM) and use vectorized histogram as training inputs.

C. Deep learning approaches

Machine learning is algorithms are limited in processing the natural images in their raw form, time-consuming, based on expert knowledge and requires a lot of time for tuning the features. Due to this limitation machine learning is overwhelming by deep learning techniques. Deep learning is fed with raw data, automatic features learner and fast. These algorithms try to learn multiple levels of abstraction, representation, and information automatically from a large set of images that exhibit the desired behavior of data. Deep learning-based algorithms showed promising performance as well as speed in different domains like speech recognition, text recognition, lips reading, computer-aided diagnosis, face recognition, drug discovery.

Now a day deep learning algorithm has got great interest in each and every field and especially in medical image analysis due to the representation of multiple levels of abstraction and extraction of features from large dataset automatically. Winona et al. [3] are proposed an Automatic Lung Cancer Detection and Diagnosis Using Handcrafted and Deep Learning Features. The system uses both handcrafted features such as the suspected nodule location, radius of the nodule, spectral signatures, taxonomic diversity and distinctness, and deep learning features were obtained with Inception-v3, a pre-trained network trained on ImageNet, which is currently the state of the art Convolutional Neural Network (CNN) architecture. This system was trying to compare the results based on their features including handcrafted features (Bag of frequencies and taxonomic indices), deep learning features and use both of them by concatenating. The result shows the promising results but it needs a room of improvement for both handcrafted features. The Combined results were obtained by concatenating the feature vectors (handcrafted and deep learning), but the extreme disparity in size led to lower accuracy than expected.

Allison et al. [4] are to develop a Deep Learning for Categorization of Lung Cancer CT Images. The system has been present ensemble methods of convolutional neural network (CNN) using multiple preprocessing methods to improving the accuracy of the system. This system uses both un-smoothing and smoothing images in separate networks to improve the classification. The result predicted by using both networks in a combined manner by applying voting systems rather than using average values. The smoothing network uses Gaussian filter methods to the testing and training images to create a second smoothed training-testing set.

Rotem et al. [10] have been proposed Lung Nodule Detection in CT Images using Deep Convolutional Neural Networks. The system was proposed by using the publicity available Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI) database, which consists of CT scan images with different shape and size, and the system uses a deep Convolutional Neural Network (CNN), which is trained, using the back-propagation algorithm to extract lung nodules in sub-volumes of CT images. The system has a 78.9% sensitivity (True positive rate) with 20 false positives (FPs) per scan. This proposed system has been achieved the result without using any segmentation methods and any FPS reduction process.

TABLE 1: SUMMARIZATION FOR RESULT OF REVIEWED ARTICLES IN THE DOMAIN OF LUNG CANCER DETECTION IN CT MEDICAL IMAGE USING VARIOUS APPROACHES

Ref.	Articles	Data source	Methods	Result
[1]	Lung cancer detection using CT Scan Image	LIDC	SVM	Accuracy 92%, sensitivity 82.5%, and specificity 50%
[8]	Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm	National Cancer Institute	Neural network (NN)	Overall accuracy 77.8%
14]	Detection of Lung Cancer from CT Image Using Image Processing and Neural Network	Private dataset	Neural Network (NN)	Overall accuracy 96.67%
[7]	A Computer-Aided Pipeline for Automatic Lung Cancer Classification on Computed Tomography Scans	LIDC-IDRI	probabilistic neural network (PNN)	95.91% accuracy, 97.42% sensitivity, and 94.24% specificity
[16]	A novel approach to a CAD system for the detection of lung nodules in CT images	LIDC	Support Vector Machine (SVM)	Sensitivity (91.65%), Accuracy is 96.22%.

TABLE 2: SUMMARY OF LITERATURE REVIEW

Ref.	Article	Data source	Method	Result
[3]	Automatic Lung Cancer Detection and Diagnosis Using Handcrafted and Deep Learning Features	LIDC/IDRI	Used Handcrafted feature and deep learning features(CNN) by concatenated	The result shows the promising results but it needs a room of improvement for both handcrafted features
[4]	Deep Learning for Categorization of Lung Cancer CT Images	CT scan image Form Kaggle dataset	CNN	Accuracy 97.5% and a low percentage of false positive (<10%)
[15]	Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images	LIDC-IDRI	Stacked Autoencoder (SAE), Deep Neural Network (DNN) and Convolution Neural Networks (CNN)	Models Accuracy Sensitivity Specificity CNN (84.15% ,83.96%, 84.32%), DNN (82.37%, 80.66%, 83.9%) and SAE (82.59%, 83.96%, 81.35%) respectively

[17]	Automated pulmonary nodule detection in CT images using deep convolutional neural networks	Lung Nodule Analysis 2016 (LUNA16)	Convolutional neural network 2D-(CNN)	sensitivity of nodule candidate detection achieves 86.42%
[18]	Automated Classification of Lung Diseases in Computed Tomography Images Using a Wavelet-Based Convolutional Neural Network	DICOM images of lung CT scans	CNN+SVM	the overall accuracy of 91.9%
[19]	Automatic Lung Nodule Segmentation and Classification in CT Images Based on SVM	DICOM images of lung CT scans	SVM	a sensitivity of 84.93%, specificity of 80.92% and an accuracy performance of 78.08%

III. CHALLENGES

A. Availability of data

Deep learning requires a massive amount of training dataset as classification accuracy of deep learning classifier is largely dependent on the quality and size of the dataset, however, unavailability of a dataset is one the biggest barrier in the success of deep learning in medical imaging. On the other hand, the development of large medical imaging data is quite challenging as annotation requires extensive time from medical experts especially it requires multiple expert opinions to overcome the human error. Furthermore, an annotation may not be possible due to unavailability of qualified expert or availability of sufficient cases are also issue in case of a rare disease. Another issue major issue is unbalancing of data that is very common in the health sector i.e. rare diseases, by virtue of being rare, are underrepresented in the data sets. If not accounted for properly, the class imbalance that ensues.

B. Privacy and black-box issues

Medical image needs to keep the privacy of patients, due to these getting the medical data is one of tedious work. The other issues are black box problem, many researchers use different methodologies for their work but due to the complexity of the model, some researchers don't know how the models work. Examples we can take the Convolutional neural network (CNN) deep learning model.

C. Complexity of contextual information among nodules

The nodule is a solid clump of tissue that exists in and around the lungs. These nodules are visible in CT scan images and have complex contextual information, they can be cancer (Malignant) or non-cancer (Benign). Because of their sophisticated contextual information, the existing cancer screening algorithm faces the problem of accuracy, and typically very tough for Doctors.

D. The inherent size and shape of cancer nodules

The Morphology of the cancer nodules and nodules, in general, are varying across the CT scan images of a lung. The inconsistency of cancer nodules shape and size will create ambiguity on radiologist and/or oncologists during screening. Due to this, radiologists and/or oncologists recommend for further action such as monitoring, blood test, and biopsy, and as a result, it becomes a burden for patients. Not only human, but it is also not easy for an algorithm to localize the cancer properties with their respective size and shape.

IV. CONCLUSIONS

Lung cancer detection is one of the crucial diseases which need medical treatment by radiologists and physicians. Many researchers have proposed and developed methods in order to reduce the human's death by using computer-aid diagnosis (CAD) systems. Still now there are many new innovation paper are published related to lung cancer, even though the system still need a room improvement due to the challenges that mentioned in the above. In order to reduce the above challenges deep learning techniques has a great role and achieved a good accuracy and performance, and for the future work in lung cancer disease and related medical imaging problem the deep learning approaches needs a good attention

REFERENCES

[1] Suren Makajua, P.W.C. Prasad*a, Abeer Alsadoona, A. K. Singh, A. Elchouemi. Lung Cancer Detection using CT Scan Images. Procedia Computer Science 125 (2018) 107–114.

[2] Jiaying Shi. Lung Nodule Detection Using Convolutional Neural Networks. Technical Report No. UCB/EECS-2018-27.

[3] Winona L. Richey, Naji Khosravan, Ulas Bagci. Automatic Lung Cancer Detection and Diagnosis Using Hand Crafted and Deep Learning Features. Senior Member, IEEE, Published 2016

[4] Allison M Rossetto and within Zhou. Deep Learning for Categorization of Lung Cancer CT Images. ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE) 2017, IEEE 272-273

[5] Image analysis in medical imaging: recent advances in selected examples. G Dougherty, Ph.D., FIET, SMIEEE. Biomed Imaging Interv J. 2010 Jul-Sep; 6(3): e32.

[6] Alexander Selvikvåg, Lundervold, Arvid Lundervold. An overview of deep learning in medical imaging focusing on MRI. ZEMEDI-10775; No. of Pages 26, 2018

[7] Emre Dandıl. A Computer-Aided Pipeline for Automatic Lung Cancer Classification on Computed Tomography Scans. Journal of Healthcare Engineering Volume 2018, Article ID 9409267, 12 pages, 2018

[8] Qing Wu and Wenbing Zhao. Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm. 2017 International Symposium on Computer Science and Intelligent Controls. 978-1-5386-2941-3/17 $31.00 © 2017 IEEE

[9] Brahim AIT SKOURT*, Abdelhamid EL HASSANI, Aicha MAJDA." Lung CT Image Segmentation Using Deep Neural Networks" in the First International Conference on Intelligent Computing in Data Sciences. Procedia Computer Science 127 (2018) 109–113

[10] Rotem Golan, Christian Jacob, Jorg Denzinger."Lung Nodule Detection in CT Images using Deep Convolutional Neural Networks". 2016 International Joint Conference on Neural Networks (IJCNN). 978-1-5090-0620-5/16/$31.00 c 2016 IEEE

[11] Razzak M.I., Naz S., Zaib A. (2018) Deep Learning for Medical Image Processing: Overview, Challenges and the Future. In: Dey N., Ashour A., Borra S. (eds) Classification in BioApps. Lecture Notes in Computational Vision and Biomechanics, vol 26. Springer, Cham.

[12] Dinggang Shen, Guorong Wu, and Heung-Il Suk (2017) Deep Learning in Medical Image Analysis: Annu Rev Biomed Eng. 2017 June 21; 19: 221–248.

[13] Baker, Mihalj; Radosav, Dragica. 2018. "Deep Learning and Medical Diagnosis: A Review of Literature." Multimodal Technologies Interact. 2, no. 3: 47.

[14] Md.B. A. Miah and Mohammad Abu Yousuf (2015). Detection of Lung Cancer from CT Image Using Image Processing and Neural Network. Conference: International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), At JU, Savar, Dhaka, Bangladesh. Volume: IEEE, pages: 1-6.

[15] Q. Song, L. Zhao, X. Luo, and X. Dou (2017) Using deep learning for classification of lung nodules on computed tomography images. Journal of Healthcare Engineering, 2017.

[16] M. Javaid, M. Javid, M. Z. U. Rehman, and S. I. A. Shah (2016), "A novel approach to the CAD system for the detection of lung nodules in CT images," Computer Methods and Programs in Biomedicine, vol. 135, pp. 125–139.

[17] Xie, Hongtao, Dongbao Yang, Nannan Sun, Zhineng Chen and Yongdong Zhang. "Automated pulmonary nodule detection in CT images using deep convolutional neural networks." Pattern Recognition 85 (2019): 109-119.

[18] Eri Matsuyama and Du-Yih Tsai (2018) "Automated Classification of Lung Diseases in Computed Tomography Images Using a Wavelet-Based Convolutional Neural Network". Biomedical Science and Engineering, 2018, Vol. 11, (No. 10), pp: 263-274.

[19] Elmar Rendon-Gonzalez and Volodymyr Ponomaryov (2016) "Automatic Lung Nodule Segmentation and Classification in CT Images Based on SVM" 2016 IEEE.

[20] L.A. Torre, F. Bray, R.L. Siegel, J. Ferlay, J. Lortet-Tieulent, A. Jemal, Global cancer statistic, 2012, Ca A Cancer J. Clin. 65 (2) (2015) 87–108.