Advances in the use of 3D Convolutional Neural Network for the detection of lung cancer

Kata raju gouda[a], John A[b], Sonu Mishra[c] , Prince Kumar[d]

[a,b,c,d]School of Computing Science and Engineering, Galgotias University, Greater Noida, India

Abstract: Lung cancer is one of the most prevalent cancer-related diseases with high mortality rate, and this is largely due to the lateness in detecting the presence of malignancy. Again, the conventional methods used in the diagnosis of lung cancer have had their shortfalls. While the effectiveness of computerized tomography in detecting this malignancy, the large volumes of data that radiologists have to process not only present an arduous task, but may also slow down the process of detecting lung cancer early enough for treatment to take its course. It is against this backdrop that computer-aided diagnostic (CAD) systems have been designed. One of such is the convolutional neural network, a method that best describes a group of deep learning models featuring filters that can be trained with local pooling operations being incorporated on input CT images in an alternating manner for the purpose of creating an array of hierarchical complex features. The need to have this type of data-driven technique is further informed by the attempt to ensure successful segmentation of lung nodules, a step that cannot be overruled when striving for a good model of detection or diagnosis. There are variations and models of the convolutional neural networks that have been effectively put to use in the lung nodule detection. The 2D CNN model has been utilized in the medical field for quite a while now, and as it has displayed its many strengths, so could the limitations not be hidden. It is in addressing these limitations and improving on the detection prowess of the convolutional neural network that 3D model is now fast gaining traction. The 3D models have been reported to return pronounced sensitivity and specificity in detection of lung nodules, but the issues of time-consumption, training complexities and hardware memory usage could make it difficult to implement the 3D model in the medical field. In this paper, review the advances that have been made in the area of adopting 3D CNN model in the diagnosis of lung cancer

Keywords: 3D Convolution Networks, Lung cancer, Nodules, CT images

1. Introduction

On a global scale, lung cancer is noted to be one of the most prevalent cancer-related diseases with high mortality rate [1]. The prognosis of the disease has not been very favorable and this is largely due to the lateness in detecting the presence of malignancy. It has been reported that patients who had lung cancer treated at stage 1 have better survival rate than those who are at the advanced stage of the disease [2]. It is estimated that rate of survival of lung cancer patients' increases by 20% as a result of early detection and subsequent proper treatment of the disease [3].Though low-dose computed tomography has shown its effectiveness in the diagnosis of pulmonary malignancy [4], the sheer (large) volumes of CT scans that radiologists always have to process while employing CT technique presents a great challenge. Radiologists face an arduous task in an attempt to detect small spherical nodules from CT images, and besides this, determining the malignancy of such nodules is yet time-consuming [5]. Therefore, it forebodes that the CT scans be examined through methods [like the computer-aided detection (CAD) technique that would adequately marked out lung nodules and non-nodules from CT scans based on predefined features[6][7][8].Using the CAD systems to complement computed tomography has been beneficial in detecting candidate nodules with small size, low contrast and those occupying areas with complicated anatomy[9][10] .The sensitivity of the CAD system in detecting candidate nodules is however affected or limited by the high probability of generating false positives [11] and owing to this, it is of utmost importance that such system be designed in suppressing the rate of false positives so as to have a precise and/or accurate assessment of candidate nodules [12] .To increase the sensitivity rating of a recognition/detection.

 model; the analysis and eventual reduction of false positives must be carried out through extraction of features and classification of lung nodule [9].

Segmentation of lung nodules

Segmentation of lung nodules poses a major threat to the accurate recognition of nodules in computerized tomography[13][14]. This is largely due to the presence of non-spherical, juxta-vascular nodules, as well as, nodules having non-solid and part-solid GGO (ground-glass opacity) which show significant and irregular intensity variations; and are known to have a greater tendency of expressing malignancy than the solid nodules [15]. However, it has to be stressed that part-solid nodules are more vigorous than non-solid nodules with respect to the rate of malignancy and doubling time.

Since malignancies and size of nodules are critical to the detection of lung cancer, it is usually recommended that non-solid nodules exceeding 5mm and part-solid nodules of all sizes be scrutinizingly investigated from the foregoing, it can be drawn that the failure to successfully segment lung nodules may ultimately mean missing out on the accurate diagnosis of lung cancer.
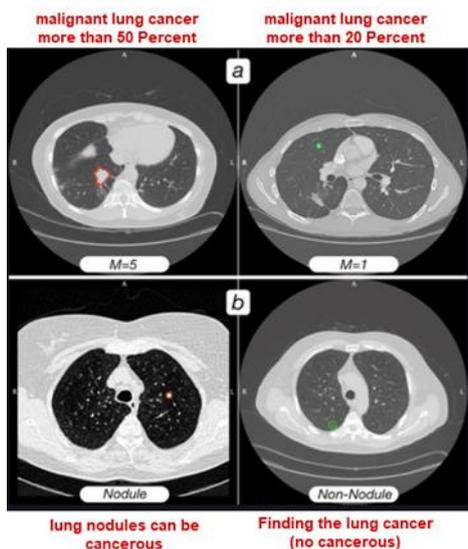
Figure 1: showing the variations in two pathological anatomy of the lung with the evident presence of alignant and benign nodules.

It is to this end that data-driven segmentation techniques as realizable under convolutional neural networks become vital in reducing inter-observer variability while eliminating the herculean task amounting from manual processing in the course of detecting candidate nodules from CT scan. Thresholding technique in tandem with voxel counting and the estimation of nodule shape in reference to the continuous space of the CT images are about the most commonly used methods of lung nodule segmentation in 3D techniques .

Convolutional Neural Network (CNN)

CNNs are typically deep learning models featuring filters that can be trained with local pooling operations being incorporated on input CT images - for example, CT scans - in an alternating manner for the purpose of creating a series of hierarchical complex features.

CNN has been aptly employed in the analysis of certain medical images that could possibly inform about specific conditions. With the right regularization method, CNNs can appreciably detect visual objects irrespective of the integration of hand-crafted features. Beyond the classification of images, CNNs have found viable usefulness in tracking and detection. CNNs, more than any other neural network, are vastly being utilized in detecting and classifying lung nodules that might be present on CT images [16,17] .

There are variations of the CNN architecture that have been effectively put to use in the lung nodule detection. For instance, the 2.5D CNN was proposed by for the representation of nodule volume features while classified nodules through the combination of contextual information at varying image scales using abstract-level representation implemented on their multi-scale CNN model. While all these aforementioned (CNN) models have achieved some degrees of success in lung nodule classification, focus has been majorly shifted to 3D CNN models which are robust for tackling the large variability issue that arises with the classification process. 3D models enable the estimation of the volume of the nodules from the voxels present within the regions of interest on CT scans [17,18] .
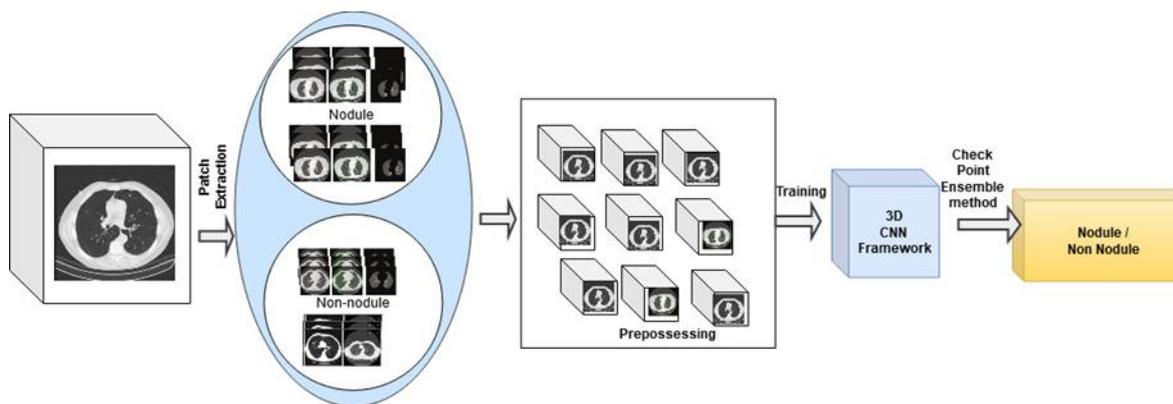
Figure 2: showing a 3D CNN framework through the pre-processing stage to the identification of nodule and non-nodules

2. Literature Survey

The area of research on the application of 3D convolutional neural network in detecting lung nodule is gradually becoming a beehive of activities with a good number of researchers having already directed their attention - and one can be certain that there are more to come - there. In view of this, it becomes pertinent to explore some of the works that have been done so far.

[5] designed a method of classifying lung nodule using 3D deep CNN and an ensemble technique. The 3D CNN models employed in this research were of two types - one with dense connections, and another with shortcut connections. These two connections allowed the direct and quick passage of gradient thus solving the gradient vanishing issue that is capable of making the training process less efficient due to poor back-propagation. More so, with these connections, it was easy to capture the unique and generic features of candidate nodules from the network. The researchers reported that their model gave a progressive performance in distinguishing nodules from non-nodules.

[3] proposed a multi-scale 3D CNN model using three different architectural frameworks that were trained and optimized for LUNA 16 challenge dataset. The extracted 3D patches with scales, 40 X 40 X 26, 30 X 30 X 10 and 20 X 20 X 6 for each of the candidate nodules. Label prediction values were incorporated from scale patches with a weighted sum, and weighting was manually achieved. This model was able to encode contextual information thus addressing the challenges of large variation feasible with lung nodules; this encoding measure specifically resulted in the model reflecting higher discriminatory capability which must have impacted its sensitivity - one of the variants was reported to have a sensitivity of 84% at 1 false positive.

Using 3D volumetric patches as input, [32] created a 3D CNN model that significantly reduced false positives. They designed the model's framework through the application of a deep CNN technique for the identification of nodules, and for detecting candidate nodules. This framework design entailed the design of a deconvolutional CNN architecture for the detection of candidate nodules on axial CT slices while a 3D deep CNN was directed at false positive reduction. The sensitivity of this model in detecting lung nodule was observed to be 91.3% - this is obviously traceable to the pronounced reduction in false positives.

An augmented method of lung cancer detection was designed. This model involved three phases namely segmentation, detection of nodule and classification of malignancy. To actualize the segmentation of lung tissue from CT scan, thresholding was used at the initial stage - though the researchers later used the watershed method in order to capture voxels that were not obtainable using thresholding since they (i.e. the voxels) were present at the edge of the lung. The Kaggles Data Science Bowl data set and Lung Nodule analysis 2016 (LUNA 16) were used for training the classifier; The latter (data set) was however found to be more effective for achieving a more accurate validation set. The preprocessing measure and the feeding of U-Net architecture - having 2D CT image slices segmentation of 256 X 256 - into the 3D CNN system appreciably took care of the issues of interference from other nodules; allowing for the accurate detection of candidate nodules. Regarding the results from the simulation experiment; the CAD method proposed here by [33] showed an accuracy rate of 86.6%, the false negative rate being 14.7% while false positive rate was 11.9%. The 3D image was then normalized using linear scaling model to obtain pixels that are squeezable to values between 0 and 1; this was

followed by the down-sampling of 3D image - each dimension measuring 0.5 in scale - and zero-centering.

Using a 3D multi-view CNN with chain and directed acyclic graphic (DAG) architectural frameworks, and CT dataset obtained from Lung Image Database Consortium/Image Database Resource Initiative (LIDC-IDRI), [4] was able to establish the superiority of multiview one-network strategy over the one-view-one network strategy with the former showing to be more significant in improving the output of the 3D convolutional neural networks.
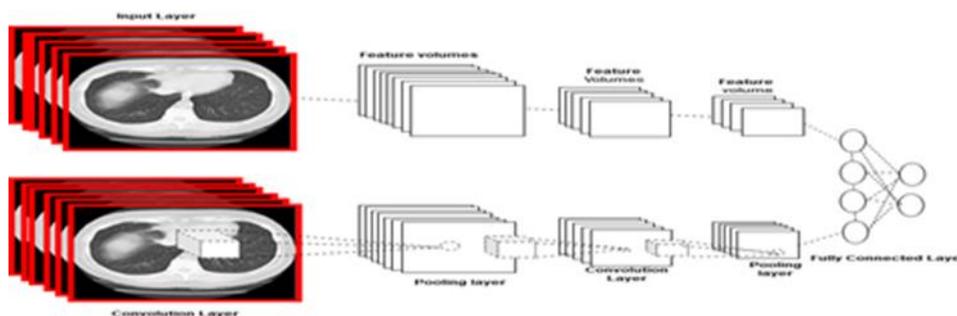
Figure 3: showing a representation of the 3D multi-view CNN framework

The authors classified the nodules based on two categories: Binary-benign and malignant; and ii) Ternary-benign, Primary malignant and metastatic malignant. Upon application of data augmentation and balancing strategy on the dataset; the authors discovered a total of 7,440 benign lesions (from 29 patient cases) and 7,080 malignant lesions (from 67 patient cases) under the binary classification while for the general classification; a total of 3,348 benign lesions (from 29 patient cases), 3,380 primary malignant cases (from 25 patient cases) and 3,368 metastatic malignant lesions (from 42 patients cases) were realised. The 3D multi-view CNN model proposed by [20] gave a specificity rate of 93.32 - 94.51 % and sensitivity rate of between 95.48 - 95.68% on the 3D multi-view CNN model with the DAG architectural framework while for the one with chain architecture gave a specificity rate of ranging between 89.73 - 93.94% and sensitivity of between 94.17 - 98.49%.

More specifically, the effectiveness of the multi-view one network strategy was reflected by the fact that one of its variants showed the lowest error it generated. With this model, the binary classification error rate was 4.59% while the ternary classification error rate was 7.70%.

[34] created the CCElargeCubeCnn, a 3D CNN model for the detection of lung nodule; this model is basically set on addressing the challenges arising from data resolution, hardware memory consumption and time-consumption, and the design is modeled after the Extended-Caffe framework. It features down-sampling and up-sampling sections; the combination of these sections enable the

capture of local and global information from the original input data. The researchers trained the network for 100 epochs and found it to be stable at 60 epoch, even as learning rate was reduced by 0.1 at epoch value of 50 and 80. In this model, the network was trained with LUNA 16 data set, and a combination of stochastic gradient descent and step-wise learning rate strategies was also used in the training process. 3D convolution was optimized in two different ways: through the single precision general matrix multiply (SGEMM) process and the Intel Math Kernel Library for Deep Neural Networks (MKL-DNN) which is renowned for its pronounced performance due to the influence of the Intel AVX512 architecture. Summarily, the performance of the CCElargeCubeCnn was quite notable as the Free-response Receiver Operating Characteristic (FROC) analysis gave a value of 0.833, and it was yet effective in detecting candidate nodules that are of small or large sizes. Breaking down further; a proportionality was observed between data resolution and performance - as better and/or improved performances were derivable from higher resolution. More so, it was reported that higher resolution brought about the conspicuousness and distinctiveness of the features of small candidate nodules. This, coupled with the FROC value, could highlight the importance of this model in detecting/predicting lung cancer at the early stage. Additionally, the efficiency of the model is brought to the fore in that training for 100 epochs of one-fold data set is actualized in nearly 61 hours which is about 5 times lesser what is attainable without optimization. Again, optimization of batch normalization using the MKL2017 resulted in 3.42 times higher performance ratio than the normalization that was not optimized. Another significant observation from this research concerns the variations in memory consumption as realized with both CPU and GPU; for the former, it was possible for the model to receive a maximum input data of 448 X 448 X 448 pixels consuming 378.7GB RAM with a batch size of 1 - this is a sharp contrast to the 128X128X128 pixels with memory consumption being 11.9GB possible on GPU with the same batch size.

[12] used series of a prior knowledge to propose a model that was effective for adequately detecting nodules from a limited data - the sensitivity rate was 90% while 5 false positives generated per CT scan - scans sourced from LIDC were used in validating the model. A Bayesian model was used to capture candidate nodules as the preset parameters filtered out contravening voxels - those whose clusters were not 23mm. The reduction of variability of the input fed into the 3D CNN was ensured by aligning the candidate nodules to principal directions through the application of intensity-weighted principal component analysis. The accuracy of this model was further boosted by the data augmentation and weight regularization - achieved through a hyper-parameter of 0.0005 - that were carried out in the developmental stage. The network was trained to 10 epochs and model selection was triggered when there was a loss on validation data set. Another interesting aspect of the observations from this research is the significance of the alignment to principal directions with marked reduction in detection performance reported when the model was not aligned; this further buttresses the role played by principal direction alignment in addressing the issues evolving from generalization errors. This model yet highlighted the influence of the dense evaluation of candidate clusters with single evaluation being more prone to errors at the stages of generation and extraction of candidate cube.

Table 1: Comparison of different CNN methods

Author	Method of detection	Result	Database system

Jenuwine et al.,(n.d)	3D CNN with architecture modeled on VGG-16 network	AUC: 0.722	LIDC-IDRI
Li et al. (2016)	Deep CNN (executed on the basis of the center of ROI	87.1% sensitivity with 4.62 false positives	LIDC-IDRI
Zhang et al. (2019)	Deep 3D CNN (implemented on Pytorch framework)	84.4% sensitivity; 83.0% specificity; AUC: 0.855	LUNA16 Kaggle dataset
Golan et al. (2016)	CNN (with weight training by back propagation)	78.9% sensitivity with 20 false positives	LIDC-IDRI
Alakwaa et al. (2017)	3D CNN (with U-Net architecture used in detecting ROI; threshold and watershed techniques for segmentation)	AUC: 0.833 Accuracy: 86.6% 11.9% false positive rate	LUNA 16
Jakimovski&Davcev (2019)	Double Deep CNN (trained over 100 epochs)	99.6% accuracy	LONI
Kang et al. (2017)	3D multi-view CNN (with chain architecture, and with DAG architecture)	For chain architecture: Sensitivity: 94.17 - 98.45% Specificity: 89.73 - 93.74% For DAG architecture Sensitivity: 95.48 - 95.68% Specificity: 94.32 - 94.51%	LIDC-IDRI

3. Conclusion

A good image recognition technique and/or model for medical use is built to deliver impressive specificity, sensitivity, and to a large extent, be very accurate in predicting certain disease conditions. For such a disease as lung cancer, the need for a diagnostic intervention that traverse through the complexities of anatomical structure and variability of nodules is one that cannot be overemphasized. As revealed in the reviewed literature, it is apparent that 3D that convolutional neural network holds great promises in the detection of lung nodules which could be highly valuable in the accurate prediction of lung cancer, and probably, the risk thereof. However, as it has also been observed different 3D CNN models gave varying results, meaning there is yet some room for improvement - even though the consistency in giving superior performance remains unquestionable. That said, there is a need to design 3D models that would considerably address issues of massive time-consumption and the requirement of large memory space in a bid to cut down the computational cost and propagate the application of 3D CNN models in clinical settings..

References

1. Siegel RL, Miller KD, Jemal A. (2016). Cancer statistics, 2016. CA Cancer J Clin. 66(1):7–30.
2. JC. (2000). Surgical management of early stage lung cancer. SeminSurgOncol 18:124–136. [PubMed: 10657914].
3. Liu X, Hou F, Qin H, Hao A (2018). Multi-view multi-scale CNNs for lung nodule type classification from CT images. Pattern Recognition; 77:262–275.
4. Jenuwine NM, Maheshb SN, Furstc JD, Raicu DS (n.d.). Lung nodule detection from CT scans using 3D convolutional neural networks without candidate selection.

5. J. Fan, W. Xu, Y. Wu, and Y. Gong (2010). Human tracking using convolutional neural networks. IEEE Transactions on Neural Networks, vol. 21, no. 10, pp. 1610–1623.

6. Teramoto A, Fujita H, et al. (2016). Automated detection of pulmonary nodules in PET/CT images: Ensemble false positive reduction using a convolutional neural network technique. Medical Physics, vol. 43, no. 6, pp. 2821–2827.

7. Shen D, Wu G, Suk HI (2017). Deep Learning in Medical Image Analysis. Annual Review of Biomedical Engineering, 19:221–248.

8. Kakinuma R, Noguchi M, Ashizawa K et al. (2016). Natural history of pulmonary subsolid nodules: a prospective multicenter study. J ThoracOncol 11:1012–1028

9. Alakwaa W, Nassef M, Badr A (n.d.). Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN).

10. Valente IRS, Cortez PC, Neto EC, Soares JM, de Albuquerque VHC, Tavares JMRS (2016). Automatic 3D pulmonary nodule detection in CT images: a survey. Comput Methods Programs Biomed; 124:91–107.

11. Mohan, Senthilkumar, et al. "Multi-modal prediction of breast cancer using particle swarm optimization with non-dominating sorting." International Journal of Distributed Sensor Networks 16.11 (2020).

12. Vallathan, G., et al. "Suspicious activity detection using deep learning in secure assisted living IoT environments." The Journal of Supercomputing (2020): 1-19.

13. John, A., P. Mathiyalagan, and Angelin Blessy. "Continuous Moving Object Clustering in Dynamic Road Network." 2020 International Conference on Inventive Computation Technologies (ICICT). IEEE, 2020.

14. Srivastava, Shriansh, J. Priyadarshini, Sachin Gopal, Sanchay Gupta, and Har Shobhit Dayal. "Optical character recognition on bank cheques using 2D convolution neural network." In Applications of Artificial Intelligence Techniques in Engineering, pp. 589-596. Springer, Singapore, 2019.

15. Srivastava, S., Priyadarshini, J., Gopal, S., Gupta, S., & Dayal, H. S. (2019). Optical character recognition on bank cheques using 2D convolution neural network. In Applications of Artificial Intelligence Techniques in Engineering (pp. 589-596). Springer, Singapore.

16. Vijayalakshmi, S., and J. Priyadarshini. "Breast Cancer Classification using RBF and BPN Neural Networks." International Journal of Applied Engineering Research 12.15 (2017): 4775-4781.

17. Noella, RS Nancy, Divyansh Gupta, and J. Priyadarshini. "Diagnosis of Parkinson's disease using Gait Dynamics and Images." Procedia Computer Science 165 (2019): 428-434.

18. Priyadarshini, J. "Machine Learning Algorithms for the diagnosis of Alzheimer's and Parkinson's Disease." International Journal 9.4 (2020)