Deep Convolutional Architecture for Block-Based Classification of Small Pulmonary Nodules

Ahmed Samy Ismaeil
Media Engineering & Technology
German University in Cairo
Cairo, Egypt
ahmedsamyismail119@gmail.com

Mohammed A.-M. Salem
Media Engineering & Technology
German University in Cairo, Egypt
Computer & Info. Sciences, Ain Shams University, Egypt
mohammed.salem@guc.edu.eg, salem@cis.asu.edu.eg

***Abstract*—A pulmonary nodule is a small round or oval-shaped growth in the lung. Pulmonary nodules are detected in Computed Tomography (CT) lung scans. Early and accurate detection of such nodules could help in successful diagnosis and treatment of lung cancer. In recent years, the demand for CT scans has increased substantially, thus increasing the workload on radiologists who need to spend hours reading through CT-scanned images. Computer-Aided Detection (CAD) systems are designed to assist radiologists in the reading process and thus making the screening more effective. Recently, applying deep learning to medical images has gained attraction due to its high potential. In this paper, inspired by the successful use of deep convolutional neural networks (DCNNs) in natural image recognition, we propose a detection system based on DCNNs which is able to detect pulmonary nodules in CT images. In addition, this system does not use image segmentation or post-classification f alse-positive r eduction t echniques which are commonly used in other detection systems. The system achieved an accuracy of 63.49% on the publicly available Lung Image Database Consortium (LIDC) dataset which contains 1018 thoracic CT scans with pulmonary nodules of different shapes and sizes.**

Keywords—**Medical Images Classification, Lung Nodules Detection, Deep Learning, Detection Systems, Supervised Learning**

I. Introduction

According to the World Health Organization [1], lung cancer is responsible for the largest number of cancer deaths worldwide. This is because of the poor prognosis for this cancer around the world. With an estimated number of 1.8 million deaths worldwide in the year 2018, lung cancer is responsible for 18.4% of the total cancer deaths. This percentage is more than the percentage of colorectal cancer which accounts for 9.2% and stomach cancer which accounts for 8.2% combined. In addition, the 5-year survival rate of lung cancer is one of the lowest compared to other types of cancers. However, the survival rate is 55% for early detected lung cancer that is at a localized stage which accounts only for 16% of the total number of cases. A thoracic CT scan is a type of 3D scan used in medical field [2]. The resolution of CT screening continues to increase as the technologies behind CT scans continues to advance. This puts a burden on radiologists who need to spend hours reading and analysing CT scans. Recently, convolutional neural networks have shown the ability to learn visual features beating manual image-processing algorithms [3], [4].

The objective of this paper is to introduce a new approach in dealing with medical images and CT scans with the help of sliding windows. Sliding windows are seldom used in medical Computer Aided Detection systems due to the complexity associated with their implementation, and the overhead of computational power that is needed as CT scans tend to be processed in the 3D space. However, sliding windows have great potential as they continue to show great results in other tasks such as object detection, face recognition and computer vision tasks [5]. So the aim was to create a new and valid system that is able to detect pulmonary nodules in the 2D slices, which would reduce the time needed to process a CT scan. Furthermore, this system does not need image segmentation or false-positive reduction techniques commonly used by other systems.

Several detection systems have been developed over the years to correctly classify pulmonary nodules. The systems that have been developed range between image-processing and machine learning techniques [6], [7]. The authors in [8] adopted a 3D neural network approach where they used a sliding window to train and test their neural network. The group achieved a sensitivity of 71.2%. Li et al. [9] worked on a DCNN architecture that showed a sensitivity of 89%. Gruetzemacher et al. Researchers [10] proposed two 3D deep learning models; one for candidate generation and another one for false positives reduction with a combined sensitivity 89.29%. Chakravarthy et al. Group [11] used a Probalistic Neural Network (PNN) for examination and classification of pulmonary nodules and their results indicated an accuracy of 90%.

Gong et al. [12] proposed a system that used 3D DCNNs and residual-network (SE-ResNet) that reached a sensitivity of 93.6%. In the research done by Huang et al. [13], a method based on an improved neural network achieved an accuracy of 81.7%. In another recent study, Jung et al. [14] introduced a 3D DCNN that makes use of shortcut and dense connections and their results were comparable with other high achieving methods. In study [15], the authors took advantage of Faster Region-based Convolutional Neural Network (Faster R-CNN) and achieved an accuracy of of 92.2%. Tang et al. [16] explored using a 3D DCNN that consists of two stages and the results showed a CPM score of 0.815. Group [17] worked on false-postive reduction techniques. Authors of [8], [9], [10], [11], [17] and [26] used the LIDC dataset while [12], [13], [14] and [27] used (LUNA16) dataset. However Jia et al. [15] used Tianchi competition dataset [23].

The rest of this paper is organized as follows: in section II, the methods and techniques that were tested at the first phase

are explained, the model is then described with the reasoning behind each layer. In section III, the results are discussed, along with the dataset and the pre-processing techniques which were used. In the last section, a conclusion to the work is presented and some ideas for future work are introduced.

II. Methodology

In this section we introduce a novel architecture for block-based classification of very small pulmonary nodules with the advantage of not using segmentation or false-positive reduction techniques which are used in most classification systems. A high-level block diagram of the steps that were used is shown in Figure 1. At first, the dataset slices are fed into a pre-processing system which generates a binary mask for each of the slices before generating the training data. These data are then passed to the model to extract features and tune its parameters. The last step is then querying this model with a slice to decide whether a nodule is present or not.

A. Pixel-Based Classification Approach

This section discusses the work done using the UNet and the ResUNet models, as pixel-based models, which take the value of the pixel's color and its proximity to other pixels into account.

1) Experimenting With The Popular UNet Model: U-Net is a convolutional neural network that is based on the autoencoder architecture [24]. It is used widely for medical pixel-based image segmentation. Although it is widely used with many modifications [25], it is very time- and memory-consuming.

To train the model, labeled or mask images were created using the annotation files given with the dataset. The mask images are binary images which show the position and shape of the pulmonary nodules in each slice if exist. The first attempt was to use the already tested and popular model UNet [18] which is suitable for image segmentation tasks. UNet uses pixel-based classification approach, that means that in our case it determines for every pixel whether that pixel contains a nodule or a non-nodule. The UNet model takes the input slice and its mask to train on. The input images are given to the model in small batches of 2 images and resized to 128×128. The model is trained for 20 epochs.

At the training phase, the model reached an accuracy of 99%. However, at the testing phase, the model ignored the pixels where there are nodules. The model showed that it could achieve 99% accuracy by classifying all pixels as non-nodules. These results are mainly due to two reasons; the first is that the number of slices that contain nodules is relatively small compared to the number of slices that do contain non-nodules. The second reason is that the region of the nodules in some slices is relatively small compared to the size of the slice. The smaller regions for the nodules makes the number of pixels that contain nodules much smaller than those that contain non-nodules. Due to these reasons it seemed that the UNet model did not apply well to our case.

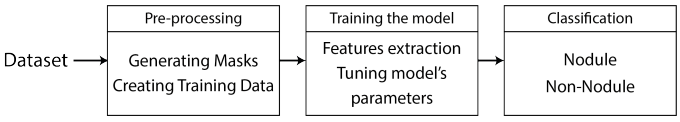


Fig. 1: Block diagram of the proposed approach.

2) Experimenting With The Popular ResUNet Model: The next model that was tested was the ResUNet model which is another popular pixel-based classifier. ResUNet was introduced by [19] and it has been shown to achieve good results on similar tasks [20]. The same approach that was used for UNet model was also used for ResUNet model; the input images were fed to the model in small batches of 2. Different batch sizes, different number of epochs and different sizes of resized images were tested. However, the same problem that appeared in the UNet model appeared once again in the ResUNet model; the model does not detect nodules' pixels and learn that it can achieve highest accuracy by just classifying all pixels as non-nodules. This is due to the same reasons that lead to the failure of the UNet model.

B. Block-Based Classification Approach

After experimenting with UNet and ResUNet it seemed that there was a major class imbalance between the number of pixels that contained nodules and those that contained non-nodules. Therefore, it was apparent that pixel-based classification approaches needed important steps for class balancing otherwise this approach fails. Thus, we headed for block-based detection approach. After the ground-truth masks for each slice was generated, the next step was to create the training data that the model would use to detect the pulmonary nodules. By inspecting [8], who used sliding windows, a similar approach was used. Rather than detecting the pulmonary nodules from the whole slice we detect them in a smaller region of the slice.

The first step in implementing the sliding window was to choose the size of this window; as all slices in the dataset have a size of 512×512 pixels, a window size of 128×128 was chosen because it gives a relatively good localization and so that the model could have good estimation of the surroundings of the nodules. After pre-processing each slice, each slice along with its mask were iterated over with the window size of 128×128. The maximum value of the group of pixels which the window covers in the mask was checked, if that maximum value was 1, that meant that the corresponding part that this window covered in the image contained a nodule, otherwise this part contained no nodules. These windows were then cropped and the cropped blocks were stored in one of two folders; the first if they contained a nodule and the second if they did not contain a nodule. The images of both the nodules and non-nodules were further split into two sub-folders to be used for training or for testing.

A stride of 32 pixels was used to generate images for the nodules; this is due to the small number of slices that actually contained a nodule, so a stride of 32 would make the nodule images augmented. However, a stride of 128 pixels was used to generate the image that contained non-nodules. A stride of 128 pixels was used to increase the diversity of the generated

images, also a threshold of 1500 generated images per scan was set to contain images from many different scans and thus the training dataset would not be too large. The training data was generated from the first 150 scans and the testing dataset from scans 151 to 200.

C. Proposed DCNN For Block-Based Classification

The training and testing processes would be discussed in the following results section after the architecture is introduced.

For our proposed model, some of layers' architecture too advantage of Truong et al.'s work [21] to capture small nodule features. The block diagram of our proposed model is shown in Figure 2. Our model contains a total of 18 layers; 4 convolution 2D layers, 4 Max Pooling layers, 6 Dropout layers and 4 Dense layers. The convolution layers were used to extract the features from the training images resulting in what is called a "Feature Map". A small even number for the number of filters was used in the convolution layer; so 8 filters for the first layer were used, 16 for the second, 32 for the third and 64 for the fourth. All the layers used filters of size 3×3 and the activation function used was "ReLU".

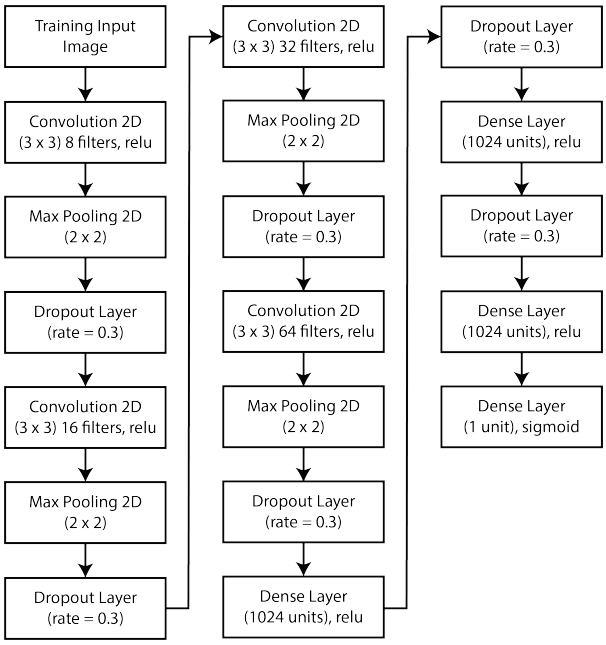


Fig. 2: The block diagram of the proposed model.

Max Pooling layers were used after every convolution layer to decrease the size of feature maps while reserving the extracted features. The size of the Max Pooling filters used was 2×2. Another technique that was used is Dropout. Dropout is a technique used to decrease over-fitting that appear in models. Dropout deactivates a percentage of nodes of the previous layer randomly at each step so that the layers won't be very dependent on their previous layers. Dropout layers were used to decrease over-fitting and increase the accuracy of the model. After every Max Pooling layer and after every Dense layer except the last one, a Dropout layer with a 30% deactivation rate was added. A Dropout layer wasn't added right before the last layer because the network has no ability to "correct" errors induced by dropout before the classification happens.

The last layers used were the Dense layers. 3 Dense layers of 1024 nodes were used to give as much data to the model without too much noise that would affect the performance. "ReLU" activation function was used for the Dense layers. The last layer used was the output layer which was only 1 unit and used an activation function of "Sigmoid". The model was compiled using "Adam" optimizer and "Binary cross entropy" was used as loss function.

III. Results and Discussion

A. Dataset

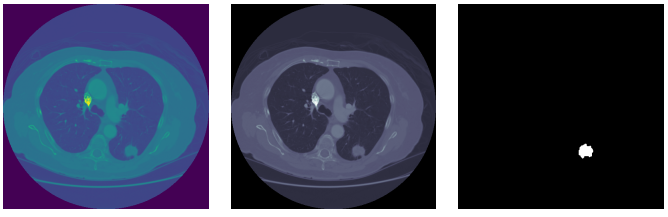
The dataset that is used in this paper is the publicly available Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI) dataset. LIDC dataset is the largest publicly available annotated dataset for pulmonary nodules. This dataset was collected under tightly specified inclusion criteria to serve as an international research resource for the development, training and evaluation of detection systems. The dataset was collected from five different institutions from across the United States and contains 1018 thorasic CT scans. A subset of 200 scans was used in this research. LIDC dataset originated from a total of 1000 patients and it contains the scans of patients at different stages of their disease.

B. Pre-Processing

CT scans come in DICOM format which uses Hounsfield scale to represent pixel values. Hounsfield scale values do not fall between the RGB range of values; so as a first step the values were normalized and a color-map was applied. To prepare the data for the model, a binary mask for each slice was needed. This mask has one-to-one pixel correspondence whether that pixel contains a nodule or not. The first radiologist's annotations were chosen and each pixel that contained a nodule was marked with white while those that contained non-nodules were left as black. These steps are shown in figure 3.

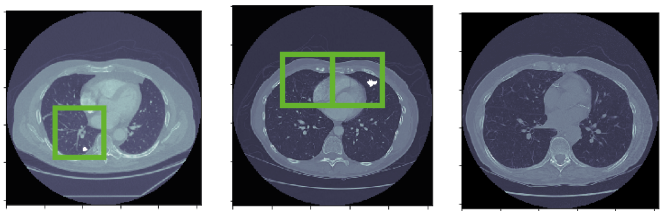
C. Experiments And Evaluation

Our model was trained for 15 epochs, as it has been found that more training epochs would lead to over-fitting. The model's training accuracy started from about 70% and kept getting better until it reached an accuracy of about 98%. The model's testing accuracy started from about 66% and kept getting better until it stabilized at an accuracy ranging from 80% to 84%.



(a) (b) (c)

Fig. 3: (a) A slice before applying color-map. (b) The slice after the color-map has been applied. (c) The ground-truth mask for the slice.



(a) (b) (c)

Fig. 4: (a) a slice that contained a nodule (highlighted with white) that the model classified correctly. (b) a slice in which the model has detected a correct nodule (highlighted with white within the right rectangle) and also has a false positive (left rectangle). (c) a slice that contained no nodules that the model classified correctly.

TABLE I: The number of blocks for each category.

Metric Name	No.	Per. relative to ground-truth
True Positives Blocks	217	52%
True Negatives Blocks	43808	64%
False Positives Blocks	25114	36%
False Negatives Blocks	197	48%

To visualize the detected blocks, a green rectangle was drawn around the regions that the model detected as containing nodules. Some results are shown in figure 4. In image (a), a a case with true positives and true negatives is shown, the model detected a nodule (highlighted with white) in the lower left region so a green rectangle was drawn there. Image (b) shows a case where the model has classified two regions as containing nodules (right rectangle and left rectangle), the right region contains correctly a nodule (true positive) while the left region does not contain a nodule (false positive). Image (c) shows a case with with true negatives only.

Table I shows the testing evaluation results represented as: True Positive (TP); which is the number of correctly classified blocks that include nodule pixels (positive blocks). True Negative (TN); which is the number of correctly classified blocks that include no nodule pixels (negative blocks). False Positive (FP) and False Negative (FN); which represent the number of wrongly classified blocks as either positive or negative respectively.

The overall accuracy as defined in [22] is given by:

$$Accuracy = \frac{TP + TN}{NB}$$

By substituting:

$$Accuracy = \frac{217 + 43808}{69336} = 63.49\,\%$$

The results are mapped to the slice-level instead of block-level. In this case, we count the number of slices that include nodule pixels. The detected slice rate is 67%.

IV. Conclusion and Future Work

In this paper we proposed a model that is able to replace current systems that detect pulmonary nodules in 3D CT scans by a more efficient one able to deal with 2D slices. Furthermore, this model does not need image segmentation or false-positive reduction techniques commonly used by other systems, which makes it more feasible to be implemented in a real-world system. This research used a subset of 200 slices of the LIDC dataset, where 150 slices were used for training, and 50 slices were used for testing. At first, ground-truth masks were generated for each slice. After generating the ground-truth masks, the popular UNet and ResUnet were tested. Both models are pixel-based classification models and they did not apply well to our case. Pixel-based classifiers did not show good results due to two reasons; the first is that number of slices that contained nodules was relatively small compared to the number of slices that contained non-nodules, the other reason was that there were slices where the nodules were very small so the number of pixels that contained nodules were much smaller. All these reasons made pixel-based classifications a challenge to get good results.

We then headed for block-based classification; A sliding window was implemented that created the training and testing data required for our model. The data was categorized according to two types (nodules and non-nodules). The model was then trained and tested and the results after that were validated. The experiments showed an accuracy of 63.49% which is significant considering no image segmentation or post-classification false positive reduction techniques were used. For the experiments done, only 200 slices from the dataset were used, which were divided to training and testing data, our proposal is that with more training data the model could achieve much better results. For future work, lots of techniques could be used to further enhance our method; one way is to train models that are similar to ours but with different input on different sizes of sliding windows. That way much better localization could be achieved with a much better accuracy.

References

[1] World Health Organization, "latest global cancer data", "https://www.who.int/cancer/PRGlobocanFinal.pdf", 2018, Accessed:2019-05-24.

[2] Monkam Patrice, Shouliang Qi, He Ma, Weiming Gao, Yudong Yao, and Wei Qian, "Detection and classification of pulmonary nodules using convolutional neural networks: A survey", IEEE Access, vol. PP, pp. 1–1, 062019.

[3] Salsabil El-Regaily, Mohammed Abdel-Megeed Mohammed Salem, and Mohammed Aziz, "Survey of computer aided detection systems for lung cancer in computed tomography", Current Medical Imaging Reviews, vol. 13, 06 2017.

[4] El-Regaily, S.A., Salem, M. A.-M., Aziz, M.H.A., Roushdy, M.I., "Lung nodule segmentation and detection in computed tomography", IEEE 8th International Conference on Intelligent Computing and Information Systems, ICICIS 2017, 2018-January, pp. 72-78.

[5] Zhengxia Zou, Zhenwei Shi, Yuhong Guo, and Jieping Ye, "Object detection in 20 years: A survey" 2019.

[6] Guobin Zhang, Shan Jiang, Zhiyong Yang, Li Gong, Xiaodong Ma Zeyang Zhou, Chao Bao, and Qi Liu, "Automatic nodule detection for lung cancer in ct images: A review", Computers in Biology and Medicine, 2018.

[7] Andrew Murphy, Matthew Skalski, and Frank Gaillard, "The utilisation of convolutional neural networks in detecting pulmonary nodules: a review", British Institute of Radiology, 2018.

[8] Golan Rotem, Jacob Christian, and Denzinger Jorg, "Lung nodule detection in ct images using deep convolutional neural networks", 978-1-5090-0620-5/16,IEEE, 2016.

[9] Wei Li, P. Cao, Dazhe Zhao, and Junbo Wang, "Pulmonary nodule classification with deep convolutional neural networks on computed tomography images", Computational and Mathematical Methods in Medicine, vol. 2016, pp. 1–7, 01 2016.

[10] Ross Gruetzemacher, Ashish Gupta, and David Paradice, "3D deep learning for detecting pulmonary nodules in CT scans", Journal of the American Medical Informatics Association, vol. 25, no. 10, pp. 1301–1310,08 2018.

[11] Chakravarthy and Harikumar Rajaguru, "Lung cancer detection using probabilistic neural network with modified crow-search algorithm", Asian Pacific journal of cancer prevention : APJCP, 2019.

[12] Li Gong, Shan Jiang, Zhiyong Yang, Guobin Zhang, and Lu Wang, "Automated pulmonary nodule detection in ct images using 3d deep squeeze-and-excitation networks", International Journal of Computer Assisted Radiology and Surgery, vol. 14, pp. 1–11, 04 2019.

[13] Wenkai Huang, Yihao Xue, and Yu Wu, "A cad system for pulmonary nodule prediction based on deep three-dimensional convolutional neural networks and ensemble learning", PLOS ONE, vol. 14, no. 7, pp. 1–17, 072019.

[14] Hwejin Jung, Bumsoo Kim, Inyeop Lee, Junhyun Lee, and Jaewoo Kang, "Classification of lung nodules in ct scans using three-dimensional deep convolutional neural networks with a checkpoint ensemble method", BMC Medical Imaging, vol. 18, no. 1, 12 2018.

[15] Ding Jia, Li Aoxue, and Wang Zhiqiang Hu, Liwei, "Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks", MICCAI 2017, pp. 559–567, 2017.

[16] H. Tang, D.R. Kim, and X. Xie, "Automated pulmonary nodule detection using 3d deep convolutional neural networks", in 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), April 2018, pp. 523–526.

[17] Soudeh Saien, Hamid Abrishami Moghaddam, and Mohsen Fathian, "A unified methodology based on sparse field level sets and boosting algorithms for false positives reduction in lung nodules detection, "International journal of computer assisted radiology and surgery, vol. 13, no. 3, pp. 397—409, March 2018.

[18] nikhilroxtomar, "unet segmentation in keras tensorflow", "https://github.com/nikhilroxtomar/UNet-Segmentation-in-Keras-TensorFlow/blob/master/unet-segmentation.ipynb,2019", Accessed: 2019-05-24.

[19] Diakogiannis, Foivos I. et al. "ResUNet-a: A Deep Learning Framework for Semantic Segmentation of Remotely Sensed Data." ISPRS Journal of Photogrammetry and Remote Sensing 162 (2020): 94–114. Crossref. Web.

[20] nikhilroxtomar, "resunet segmentation in keras tensorflow", "https://github.com/nikhilroxtomar/Deep-Residual-Unet/blob/master/Deep%20Residual%20UNet.ipynb", 2019,Accessed:2019-05-24.

[21] Thanh-Dat Truong, Vinh-Tiep Nguyen, and Minh-TrietTran, "Lightweight deep convolutional network for tiny object recognition", In Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods, 2018.

[22] Aur elien Geron, Hands-On Machine-Learning with Scikit-Learn & Tensorflow, O'Reilly Media, 2017

[23] Tianchi platform, "https://tianchi.aliyun.com/dataset/dataDetail?dataId=42", Accessed: 2020-08-24

[24] , O. Ronneberger, P.Fischer, T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", In Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI),LNCS, volume 9351, Springer, Pp 234–241, 2015.

[25] Almotairi, S., Kareem, G., Aouf, M., Almutairi, B., Salem, M. A.-M., "Liver Tumor Segmentation in CT Scans Using Modified SegNet", Sensors (Basel, Switzerland), 20(5), 1516, 2020, https://doi.org/10.3390/s20051516.

[26] El-Regaily, S.A., Salem, M.A.M., Abdel Aziz, M.H., Roushdy, M.I., "Multi-view Convolutional Neural Network for lung nodule false positive reduction", Expert Systems with Applications 113017, 2019.

[27] Mohamed, A. S. E.-D., Salem, M. A.-M., Hegazy, D., Shedeed, H. A., "Improved watershed algorithm for CT liver segmentation using intraclass variance minimization", Communications in Computer and Information Science, 756, pp. 164-176, 2017.