

Classification of Pulmonary Nodule using New Transfer Method Approach

Syed Waqas Gillani¹, Bo Ning^{2*}

College of Information Science and Technology
Dalian Maritime University, Dalian, Liaoning, China

Abstract—Lung cancer is among the world's worst cancers, and accounted for 27% of all cancers in 2018. Despite substantial improvement in recent diagnoses and medications, the five year cure ratio is just 19%. Before even the diagnosis, classification of lung nodule is an essential step, particularly because early detection can help doctors with a highly valued opinion. CT image detection and classification is possible easily and accurately with advanced vision devices and machine-learning technology. This field of work has been extremely successful. Researchers have already attempted to improve the accuracy of CAD structures by computational tomography (CT) in the screening of lung cancer in several deep learning models. In this paper, we proposed a fully automated lung CT system for lung nodule classification, namely, new transfer method (NTM) which has two parts. First features are extracted by applying different VOI and feature extraction techniques. We used intensity, shape, contrast of border and spicula extraction to extract the lung nodule. Then these nodules are transfer to the classification part where we used advance-fully convolution network (A-FCN) to classify the lung nodule between benign and malignant. Our A-FCN network contain three types of layers that helps to enhance the performance and accuracy of NTM network which are convolution layer, pooling layer and fully connected layer. The proposed model is trained on LIDC-IDRI dataset and attained an accuracy of 89.90 % with AUC of 0.9485.

Keywords—New transfer method; VOI extraction; feature extraction; classification; LIDC-IDRI dataset

I. INTRODUCTION

In this modern era of machine learning, doctors are finding some form of support that encourages their ability to analyze and diagnose CT images of patients easily and to identify extremely effective and accurate pathologies. Timely detection and classification of lung nodules enhances clinical results and improves the chances of survival rates [1]. Instead, cancer is now the ultimate common pale pathology that endangers anyone irrespective of age. Between various diseases, lung cancer is the unregulated cell growth in an overt body district [2]. The development of lung nodules is indicated for lung cancer and shows the clinical disease phase [3]. The nodules existing in the lungs equate to variants in lung tissue as from standard, which is round or fit as a 3 millimeter and 30 millimeter wide fiddle [4]. The human body comprises of several cells. The nodule is formed whenever cells develop feral outside of lung [5]. Computed tomography (CT) is mostly used for the detection of lung cancer [6]. Centered on the findings of a nationwide lung screening study performed in the USA [7], scanning for low dose CT scans decreased deaths from lung cancer by 18%. Hence, CT is

considered an appropriate diagnostic technique for the detection of cancer. Computer-aided diagnostic (CAD) methods are designed to help clinicians in analyzing health data and diagnosing the disease. Fig. 1 shows that 24% of CAD's are just used for lung cancer [23]. It means that the medical technology is being rapidly accepted and implemented quickly. CAD could be categorized into two kinds: the detection and diagnostic system (CADE and CADx). CADE's aim is to find the ROI to detect unique abnormalities and CADx offers doctor medical help to discern the form, frequency, level, development and disease deterioration. The CADx could just only use to the diagnosis purpose for example shape, thickness, and texture.

CNN began as an advancement of the deep neural network with the use of convolution algorithms to help interpret the input image. It is built on the basis of the biologically visual field and is therefore quite efficient for difficulties of image recognition, irrespective of size or volume. Ginneken et al. [8] evaluated CNN as well as food over fat as a clinical CAD tool for the identification of lung cancer tumors. LIDC CT-scanning images have been used to locate lesions. Over feat CNN collected characteristics of the lung nodules, and support vector machine methods were used to characterize the tumors. In fact, CAD-systems found the nodules. Analysis indicates that each approach can detect lesions with a sensitivity of 70%. Anthimopoulos et al. [9] Suggested that CNN identify and describe multiple lung tissues of respiratory problems. The CNN included five CNN layers, one layer of pooling, and three FC layers. The algorithm suggested was contrasted with many other CNN systems such as AlexNet and VGG Network. Results showed that the suggested CNN for tissue classification and identification was preferable to the other architectures. The new CNN obtained accuracy 85.61 %. Li et al. [10] suggested a CNN with just one convolution layer for patch detection on CT images with high resolutions. For this reason LIDC lung database is included. In addition, a pair of Support vector machine classifier with three feature extraction techniques their implemented CNN was contrasted to the mixture of a three-function extraction process and Support vector machine classification. Analysis indicates that CNN reached higher values of Sensitivity 0.88 and Precision 0.93 as compared to other methodologies. Shen et al. [11] suggested a multi-scale CNN to distinguish malignant and benign nodules in the lungs. The CNN classification methods have been used as support vector machine and random forest. In [12] a CNN is used to identify the lung CT scan images with an ILD dataset. CNN findings for the classification of good, ground glassy opacities, nano-nodules, reservoirs, of

*Corresponding Author

II. METHOD

ILDs were stated to be beneficial. In [13] used the updated edition of ResNet-18 as a classification system to label CT scanning images for the data science bowl and Kaggle lung. For certain of such fields, DNN may also attain near human intelligence [14]. CNN is the popular DNN model, integrates supervised learning methods that can help it to capture high-level features from unpasteurized images, and is promoted to handle lung nodule classifier accuracy improvements [15]. A new work adopts a deeper CNN with a one CNN layer for the classification of nodules and reveals a higher precision in comparison to the feature's traditional extraction methods. Owing to the customization of the CNN model [16], the transfer learning method often provides promising outcomes for the nodules classification [17]. Until the advent of deep learning, manual feature designing assisted by classifiers has been the basic nodule classification method. With the public availability of the LIDC-IDRI dataset [18], deep learning techniques [19] are already the dominant nodule classification system for research.

In order to obtain a suspicious-sensitive classification in CT images, we have to address at minimum two significant obstacles, the complexity of nodule depiction induced by a broad variety of nodule's morphology variants, and the problem raised by analytical models' radiological complexity to identify qualitative features as it is difficult to differentiate benign nodules from malignant nodules. So, to solve these challenges in this paper, we proposed a new transfer method (NTM) approach is used for lung nodule classification as shown in Fig. 2. First, volume of interest (VOI) extraction and feature extraction is used to extract the main features through CT image. Then for lung nodule classification task we proposed advance-fully convolution network (A-FCN).

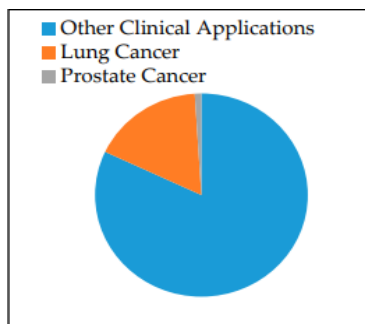


Fig. 1. Different CAD Applications.

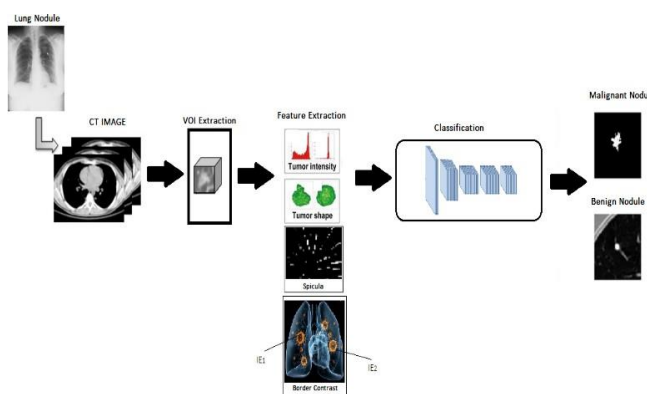


Fig. 2. Proposed New Transfer Method (NTM) Approach.

The proposed fully automated new transfer method (NTM) for lung nodule classification has two parts: (1) volume of interest (VOI) extraction and feature extraction; (2) advance-fully convolution network (A-FCN) which will automatically classified lung nodule between benign and malignant. In the proposed NTM technique the features are extracted by applying different VOI and feature extraction techniques then these features is being used as an input of advance-fully convolution network to classify the nodules among benign and malignant.

A. Volume of Interest (VOI) Extraction

The physicist has indicated the location and size of the nodule to be examined using standard CT scan. Thus the VOI segmentation across the pulmonary nodule is conducted for examination upon CT images. The center position of a VOIs obtained from the standard CT images are configured manually whereas the CT images were tested for MPR images. Firstly, there is the transaxial's image with the highest nodule region, and manually determined the central coordinates. Therefore the diameter of the image is fixed to the maximum, and denoted by MD_{xy} , in Nodule path x-y for transaxial plane. Then, even as deforming including its slice in the path of the axis direction of body, the value of a slice wherein the lesion is still visible is being acquired and establish as MD_z . The VOI is now extract from the original image by using pixel on 3 sides, that is, $2MD_{xy}$, $2MD_{yx}$ and $2MD_z$.

B. Feature Extraction

Feature extraction is the key of proposed model before the learning stage. One of the strategies for removal of dimensionality in image-processing is feature extraction of images. If the input image being examined is too complicated to process due to its repetitive features, therefore it is preferable to implement feature extraction technique. It effectively turns the immense collection of data into reduced range of feature. We extract CT image pixel intensity, shape, contrast of border and spicula for the lung nodule classification through CT image.

1) Features Explanation

a) *CT Image Pixel Intensity and Shape of Nodule:* Most malignant lesions in CT-images include a larger intensity of a pixel. Standard uptake or silly useless value (SUV) [20] of slightly earlier CT images has therefore been described as ESUV and DSUV. In addition, the gap between the deferred and early stages in SUV was specified as part of the SUV. Two approaches are developed in the measurement of SUV, namely SUV_{max} and SUV_{peak} . The CT values in the nodule central and the average CT value within the nodule is determined throughout the CT images. As with the nodules' shape, malignant nodules frequently have a small ball type of form, whereas benign has a line type structure. An approach was suggested to use a Hessian-matrix to compare the ball type and line types [21]. It the Hessian matrix was done by adding the 3D image differentiation in 2nd order.

b) *Contrast of Nodule Border and Spicula:* Sometimes the location of a malignant nodule is uncertain. Consequently

the boundary comparison was measured using the discrepancy between the CT values of the nodules' exterior and interior borderlines. The estimated CT values only at numbers corresponding here to interior edge IE_1 ($CTIE_1$) and the exterior area IE_2 ($CTIE_2$) are extracted in order to quantify this value, and the distinction among the two values, is established as either the contrast which is shown in Fig. 3 The image is binaries to acquire IE_1 and IE_2 , and the shape was generated by the Sobel system. On the outlining the range of pixels has also been specified as IE_2 . Consequently, morphological erosion with a systemic variable which includes a diameter of one pixel and the shape of the decreased area was derived in the same way as mentioned above; a collection of such pixels is utilized as IE_1 . The appearance of spicula across the nodule raises the chances for malignancy in the nodule. We used Gabor filter used to detect spicula [22]. Using Gabor Filter allows imagining line shapes and the orientation as shown in Fig. 4. Finally we got detected spicula is by using Gabor filter as represent in Fig. 5 and the amount of radial items and the ranges were determined as spicula, S_a and S_b characteristics.

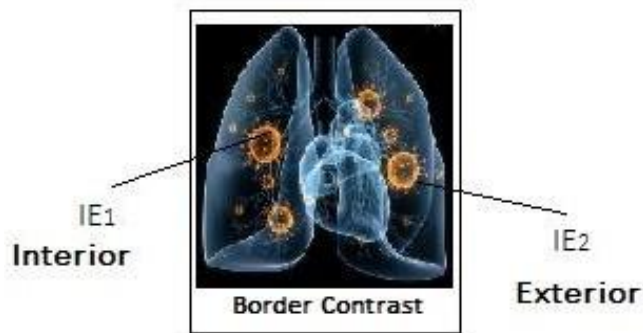


Fig. 3. Nodule Contrast.

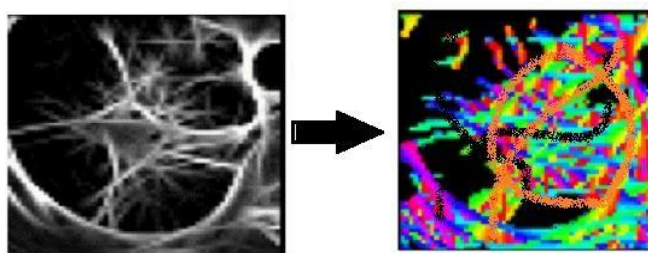


Fig. 4. (Left) Intensity, (Right) Angle using Gabor Filter.



Fig. 5. Detected Spicula.

C. Advance-Fully Convolutional Network (A-FCN)

We proposed a new method to enhance the classification accuracy of lung nodule. Our A-FCN is composed of three separate kinds' layers which are convolution layer (CL), pooling layer (PL), and fully connected layer (Fc). On every input image CL executes convolution operators. As during training process these all layers can obtain features from the input data. Deep layers could identify high abstractions features. A pooling layer operates on independent feature channel and analyzes the surrounding values into just one. It therefore decreases the number of training parameter and significantly reduces the training time effectively. An fc layer connects every neuron in the existing layer from all neurons through last layer. Fc layers perform less reliably as compared to convolution layers owing to the lack of structural connection in images. It often expands the number of training parameters and thus, elongates time needed in training. Dropout which is a regulation strategy to decreasing amount of neurons and interactions is proposed to fix certain issue. As shown in Fig. 6 there are fifteen layers in which nine are convolution one. We employ three convolution blocks: 3 CL and 1 Pooling layer and max operator pooling layer. Every layer is accompanied by a ReLU activation function. Due to its efficient computing efficiency the last that have been fully connected are aimed at resolving a problem of classification from the extracting features. A ReLU and a softmax activation function are observed, respectively.

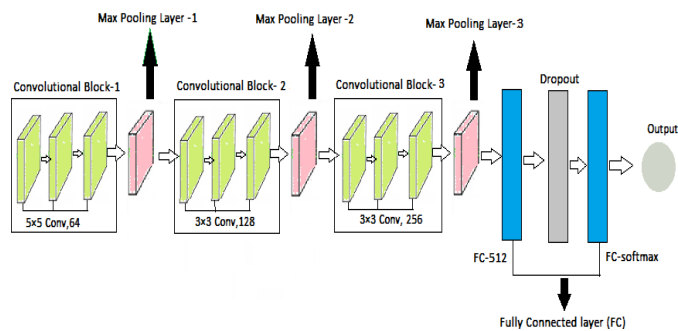


Fig. 6. Proposed Advance-Fully Convolutional Network for Classification of Lung Nodule.

III. EXPERIMENTS

This section provides descriptions of how the proposed model is implemented and evaluated and it also discusses the experimental findings. Section 3.1 shows about dataset which were described for our experimentation and also training and testing phase. The analysis on the LIDC dataset is shown in Section 3.2. It explains the contrast between our model and several recent methods.

A. Experimental Setting

1) *LIDC-IDRI Dataset*: The original edition of LIDC is 399 scans. Later, LIDC-IDRI has been extended to 1018 patient LIDC scans. In every CT scan, four professional radiologists used a LIDC-like labeling technique and an XML related file to report the data using a double-stage image classification process. Each radiologist interpreted and labeled every case separately at the first blind reading point. Growing radiologist studied its features and the attributes of any of the

three radiologists. Then they separately interpreted and recorded each case for the final decision. These two-phase marking will accurately classifies all pulmonary nodules, eliminating the need for pressured consensus. There were three types in the region of the nodule which are nodules ≥ 3 millimeter, nodules < 3 millimeter, and non-nodules ≥ 3 millimeter. If there were more than three annotated nodule meaning by and over two doctors, the nodule would be labeled malignant. Otherwise nodule is considered benign in comparison. Around 198 malignant nodules and 153 benign nodules were available in our dataset. The nodules are ignored by the same votes. To order to reduce computational uncertainty, we separate the key transect for every voxel. We then apply the Data augmentation (DA) approach to expand the data by applying types to a dataset in terms of reducing deep learning over fitting. We dynamically flip and magnify the image through zooming 0.2 in general. The translation step is chosen using just a voxel of $[-6, 6]$, and a rotation angle of $[90^\circ, 180^\circ \text{ and } 270^\circ]$ has been chosen at random. Eventually, 1958 malignant nodules, and 1867 benign nodules are available.

2) *Training and testing phase:* We employ Adam optimizer throughout the training process avoiding exponential decline from a learning rate=0.0001, and $\beta_1=0.9$, $\beta_2=0.99$ as a standard parameter. The k-cross validation approach is used. The procedure divides the data into the very same size k sections. The calculation is taken in k iterations; one component is used to test, whereas the rest is used for training. Iterations $K = 10$ is a popular approach of such a validation approach.

B. Experimental Result and Analysis

We evaluated the efficiency of our proposed new transfer method (NTM) model on the LIDC dataset and the ROC curve is also shown in Fig. 7. Our model achieved accuracy of 89.90% and area under curve (AUC) 0.9485. The efficiency of the suggest model is also contrasted with several other effective models in order to assess the effects of the appropriate methods, as seen in Table I. Liu et al. [24] presented a new hybrid convolution neural network in which LeNet as well as AlexNet were used and they combined both networks layers. Zhao et al. developed a multi-scale VGG-16 learning method to extract exclusion features from substitute stacked layers [25]. The latest 3D multi-scale convolution neural network framework for the classification of lung nodule has been created by Tafti et al. [26]. Yu Gu et al. [27] recommended the usage of a systemic method to identify a Deep convolutional neural network 3D multi-scalar predictor lung nodule. Moreover, it is suggested that an extra tiny nodule be observed using a classification approach with several scale cube clusters. Shen et al. [28] employed multi-scale to consider the complexity of a nodule, utilizing instead layered strata to isolate the discriminatory features.

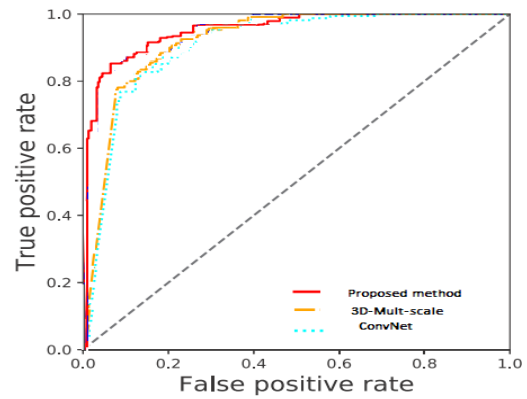


Fig. 7. ROC Curve of Proposed Method.

TABLE I. COMPARISON OF PROPOSED NTM WITH OTHER METHODS

Research	Methods	Sensitivity	Specificity	Accuracy %	AUC
Liu et al. [24]	CNN	-	-	82.23	-
Zhao et al. [25]	ConvNet	0.843	0.858	84.97	0.902
Tafti et al. [26]	3D-Multi scale CNN	-	-	83.75	0.926
Yu Gu et al. [27]	Multi scale 3D-DCNN	0.832	0.847	84.66	-
Shen et al. [28]	Multi scale CNN	-	-	86.84	-
Proposed method	NTM: VOI and feature extraction with Advance-fully Convolutional Network	-	-	89.90	0.9485

IV. CONCLUSION

In this paper the aim is to introduce a new architecture named: new transfer method (NTM) for lung nodule classification. The network is divided into two parts: (1) VOI extraction and feature extraction; (2) advance-fully convolutional network (A-FCN). First features are extracted from lung CT image by using two extraction techniques in which we used intensity, shape, contrast of border and spicula extraction to extract the lung nodule. After finding nodule then it transfer to the A-FCN for lung nodule classification. Our A-FCN part used three convolutional blocks in which we have convolutional layer, pooling layer and fully connected layer. We also use ReLU and softmax activation function in our classification phase to solve the complexity of lung nodule classification. Finally, our system is thoroughly equipped for benign and malignant lung nodule classification. The result of LIDC-IDRI indicates that the NTM framework has improved accuracy.

REFERENCES

- [1] S. Lee, A. Kouzani, and E. J. Hu, "Hybrid Classification of Pulmonary Nodules" Communications in Computer and Information Science, vol. 51, pp. 472-481, 2009.
- [2] Ur Rehman MZ, Javaid M, Shah SI, Gilani SO, Jamil M, Butt SI (2018) An appraisal of nodules detection techniques for lung cancer in CT images. Biomedical Signal Processing and Control 1(41):140-151.
- [3] Silva D, Giovanni LF, Thales Levi AV, AristófanésCS ACP, Marcelo G (2018) Convolutional neural network-based PSO for lung nodule false positive reduction on CT images. Comput Methods Prog Biomed 162:109-118.
- [4] Skourt BA, El Hassani A, Majda A (2018) Lung CT image segmentation using deep neural networks. Procedia Computer Science 127:109-113.
- [5] Badura P, Pietka E (2014) Soft computing approach to 3D lung nodule segmentation in CT. Comput Biol Med 53:230-243.
- [6] Sone S, Takashima S, Li F, Yang Z, Honda T, Maruyama Y, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet. 1998;351(9111):1242-5.
- [7] National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365(5):395-409.
- [8] Ginneken, B.V.; Setio, A.A.A.; Jacobs, C.; Ciompi, F. Off-The-Shelf Convolutional Neural Network Features for Pulmonary Nodule Detection in Computed Tomography Scans. In Proceedings of the IEEE 12th International Symposium on Biomedical Imaging (ISBI), New York, NY, USA, 16-19 April 2015.
- [9] Anthimopoulos, M.; Christodoulidis, S.; Ebner, L.; Christe, A.; Mougiakakou, S. Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. IEEE Trans.Med. Imaging 2016, 35, 1207-1216.
- [10] Li, Q.; Cai, W.; Wang, X.; Zhou, Y.; Feng, D.D.; Chen, M. Medical Image Classification with Convolutional Neural Network. In Proceedings of the 13th International Conference on Control, Automation, Robotics & Vision, Marina Bay Sands, Singapore, 10-12 December 2014.
- [11] Shen, W.; Zhou, M.; Yang, F.; Yang, C.; Tian, J. Multi-scale Convolutional Neural Networks for Lung Nodule Classification. Inf. Process. Med. Imaging 2015, 24, 588-599.
- [12] Bondfale, N.; Banait, S. Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. IJARCC 2017, 5, 9851-9856.
- [13] Data Science Bowl 2017. Kaggle. Available online: <https://www.kaggle.com/c/data-science-bowl-2017> (accessed on 5 February 2019).
- [14] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529, 484-489.
- [15] Szegegy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2015b). Rethinking the inception architecture for computer vision. CoRR.
- [16] Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., et al. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging, 35(5), 1299-1312.
- [17] Zhao, X., Liu, L., Qi, S., Teng, Y., Li, J., & Qian, W. (2018). Agile convolutional neural network for pulmonary nodule classification using CT images. International Journal of Computer Assisted Radiology and Surgery, 13(4), 585-595.
- [18] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In MICCAI, 2015.
- [19] S. G. Armato et al. The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. Medical physics, 38(2):915-931, 2011.
- [20] Keyes JW Jr. SUV: standard uptake or silly useless value? J Nucl Med. 1995;36(10):1836-9.
- [21] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44.
- [22] Rangayyan RM, Ayres FJ. Gabor filter and phase portraits for the detection of architectural distortion in mammograms. Med Biol Eng Comput. 2006;44(10):883-94.
- [23] Global Computer-Aided Detection (CAD) Market US\$ 2.2 Billion by 2023.
- [24] X., Liu, L., Qi, S., Teng, Y., Li, J., & Qian, W. (2018). Agile convolutional neural network for pulmonary nodule classification using CT images. International Journal of Computer Assisted Radiology and Surgery, 13(4), 585-595.
- [25] Zhao, D., Zhu, D., Lu, J., Luo, Y., & Zhang, G. (2018). Synthetic medical images using F&BGAN for improved lung nodules classification by multi-scale VGG16. Symmetry, 10(10), 519.
- [26] Tafti, A. P., Bashiri, F. S., LaRose, E., & Peissig, P. (2018, June). Diagnostic classification of lung CT images using deep 3d multi-scale convolutional neural network. In 2018 IEEE International Conference on Healthcare Informatics (ICHI) (pp. 412-414). IEEE.
- [27] Gu, Y., Lu, X., Yang, L., Zhang, B., Yu, D., Zhao, Y., ... & Zhou, T. (2018). Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs. Computers in biology and medicine, 103, 220-231.
- [28] Shen, W., Zhou, M., Yang, F., Yang, C., & Tian, J. (2015, June). Multi-scale convolutional neural networks for lung nodule classification. In International Conference on Information Processing in Medical Imaging (pp. 588-599). Springer, Cham.