Review on Lung Cancer Lesion Detection and Segmentation Methods

Kangjun Ji [1, *, †] and Haoyang Lin [2, †]

[1] Department of Biomedical Engineering, Imperial College London, London, UK

[2] Department of Information and Computing Science, Xi'an Jiaotong-Liverpool University, Suzhou, China

* Corresponding author: Haoyang.Lin20@student.xjtlu.edu.cn

[†]These authors contributed equally.

Abstract. Lung cancer itself and relevant detection and segmentation methods, in the modern society, becomes increasingly popular and significant topics. Scientists believe that people smoke positively may deteriorate their body health themselves, and people who breathe it in second hand may also suffer from this harmful environment. To help people with lung cancer lesions, there are several methods used for cancer treatment. Automated CT imaging can encircle suggested segmentation areas in a 3-D version, and it provides convenience with users when they feel tired after diagnosing for a whole day. Semi-automated CT deep learning model is another technique to detect particular regions in the lung by adjusting pixels. Additionally, few-shot learning based on advanced learning algorithm is an efficient method for lung cancer lesion detection, and Generative Adversarial Networks(GAN) can be used for lung cancer detection by using a small number of medical images as train datasets. However, CNN model cannot obtain global information; therefore, the integration of 2dcnn and 3dcnn solves this limitation in an effective approach.

Keywords: Convolution Neural Network, Lung Cancer Lesion Detection, CT Image Processing.

1. Introduction

It is significant to notice that lung cancer is one of the deadliest diseases in the world, and lung cancer treatment is a painful, terrible, and money-consuming process. In terms of modern treatment technology and medicine, nobody nowadays innovates the specific medicine for lung cancer. According to the recent researches, lung cancer death rate accounts for 18.4% of total cancer death rate worldwide in 2018, which equals the death rate proportion of breast and colon cancers combined [1].

In terms of factors of lung cancer, positive smoking or passive smoking should take the major responsibility, and tobacco is regarded a risky factor of lung cancer, stomach cancer, pancreatic cancer and kidney cancer [2]. The detection of lung cancer regularly is very significant in the early stage. computed tomography (CT), a kind of traditional and constant analyzing method in the past decades, makes a huge contribution to this field. However, it is difficult for beginners to use CT imaging, and they commonly will make mistakes.

To solve those problems, this essay applies automated CT imaging and diagnosis and semi-automated deep learning algorithm via CT. Automated CT can create convenience for users, especially for doctors, and provide them with suggestions when they need to make treatment planning. Also, the simi-automated deep learning (DL) model can adjust the pixel size of images, and analyze the different lesion parts of the lung in different pixels [3]. Few-Shot Learning is an application of the Meta Learning in the field of supervised learning. It has been shown to be an effective method for lung cancer detection, with the property that new classes can be classified quickly based on trained models, and it has been widely used in various fields, including object detection, image classification. Also, sing generative adversarial networks (GAN) provides a new idea to solve this problem.

2. Related lung cancer detection methods:

2.1. NSCLC Volumetric Segmentation via Automated CT Imaging

Machine learning and artificial intelligence techniques gradually positively influence every field in the world, and medical imaging field is one of the most affected industries [4]. The fully automated pipeline is designed for detecting and segmenting non-small cells of lung cancer, and this innovative approach is faster and more reproducible since this technology demonstrates better quantitative performance and abilities in various fields. For example, image slice thickness, and the size and location of tumor are relevant influenced fields [1]. Additionally, due to the application of machine learning and artificial intelligence techniques, this system can segment the risky tumor from organs in human bodies, and the high accuracy and precision can avoid serious medical accidents.

There are three main steps in the proposed workflow, which are image preprocessing, lung isolation, and automated detection and segmentation via CT imaging. It is the most crucial step to acquire various images from different scanners in different datasets, and responsible for reconstruction protocols. The first step is to extract a 3-D array with voxel intensity values from Digital Imaging and Communications in Medicine data [5]. After normalizing the space, a standard image that contains the structure of bone window setting is saved. The spatial normalization allows to optimize the storage and image processing based on the ways packing the images, which are converted into a significantly smaller than 8-bit integer range, and procedure can also filter high-frequency noise [1].

Secondly, the second step - lung isolation - allows the system to focus on the ROI and the input of the entire CT scans. The last step is to achieve automated tumor detection and segmentation, which employs the convolutional neural network [1]. The axial projection was used to train the network due to the higher resolution of image representation in this plane [6]. Automated CT detection and segmentation does not need the supports and helps from a human operation to do external validation [5]. As shown in Fig.1. and Fig.2., the red dash line shows the segmentation part predicted by the automated CT imaging model, while the blue line is the area that human beings indicate. Based on the diagrams, we can conclude that the accuracy of automated CT imaging is very high since the predicted segmentation area is approximately as large as human beings' diagnosis [1].

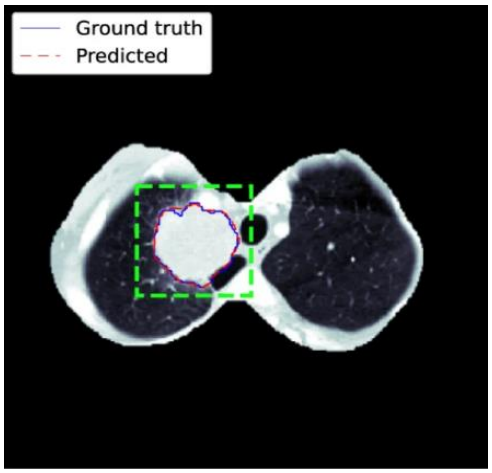

Figure 1. The image of segmentation via automated CT imaging [1]

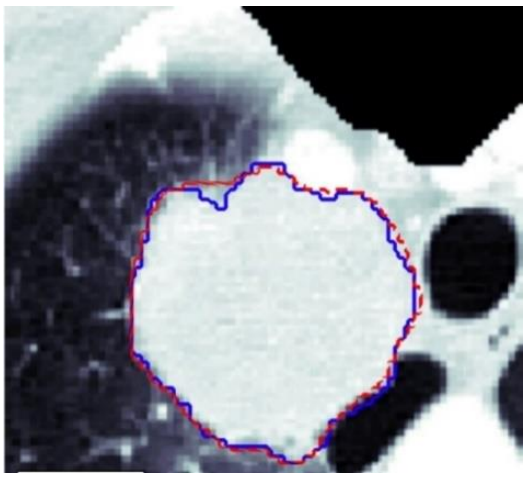

Figure 2. The enlargement of Fig.1. [1]

2.2. Semi-automated Measuring Techniques by deep learning model

Aside from that, the measurement of lung tumor size can also be accessed through CT in a semi-automated way. This method, deep learning (DL) model, generates and provides precise and accurate unidirectional measurements that are comparable the results drawn from human radiologists. This approach recently makes an impressive contribution to healthcare and diseases treatment fields, and

DL algorithm can deal with various types of images, and detecting the distance between patients and sensing machines [7]. To solve the problems caused by conversion of segmentation, scientists design semi-automated measuring device that applying DL learning model to classify the images. A new method, the measurement towards lung cancer lesions using deep learning algorithm that bases on semi-automated CT scanning [7].

A deep learning network consists of three neural networks that can label and classify the area of a targeted lung cancer lesion region, determining whether the size of images, which can be larger or smaller than 32 pixels. If the DL networks fails classifying and detecting images, the lesion size with 32 pixels should be responsible for this failure [3]. The first step of measuring begins with modifying the CT images into a 32 pixels size, and the images are magnified by $32/M_{px}$. (M_{px} is the measurement in pixels converted from unidirectional measurements in centimeters). Secondly, the training dataset for this DL algorithm is formed through zooming in and out, horizontal shifts and vertical shifts. The target lesion is located in a 128 by 128 pixel frame based on the center point of measurement machine, which is the center of the sensing device. Finally, human radiologists can access the cancer lesion CT images and relevant data [3].

Concluding results from various tests and experiments, the results concluded by DL algorithm and human radiologists are similar, which means the accuracy reaches 0.959. Many scientists and human radiologists use this method - semi-automated CT measurement by DL algorithm – achieve a mean pixel different and standard deviation of 2.85 and 2.51, which achieves similar performance in this study based on the Fig.3. From the Bland-Altman diagram (Fig.3), the DL network tends to overestimate the size of tumor compared with estimation made by human radiologists.

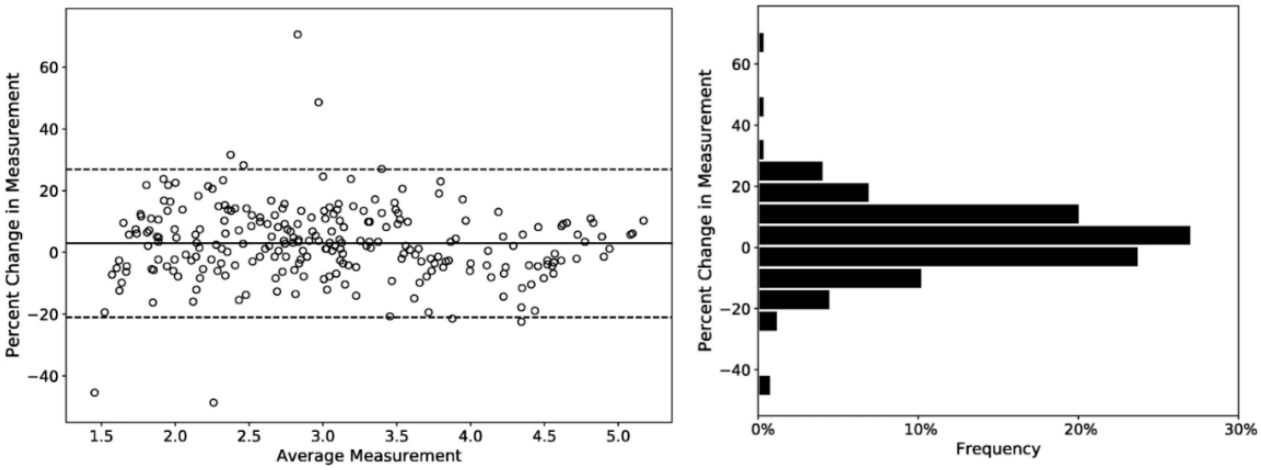

Figure 3. The Bland-Altman Diagram to show the data of semi-automated measuring

2.3. Lung Cancer Segmentation via Transfer Learning and GAN: Usefulness of a Pretrained Model

Datasets are frequently a significant problem in deep learning applications. To increase model accuracy, deep learning models often need enormous training datasets, while getting such datasets in the field of medicine is frequently difficult due to the high resolution, vast memory requirements, and privacy concerns of individual medical pictures. Therefore, many approaches have been investigated to solve this problem, including migration learning using pre-trained models, data augmentation and generation of datasets through GAN, etc. Mizuho Nishio et al [8] generates the dataset for segmentation by GAN [9] and uses the 3D graph cut method to generate labels.

First, the LUng Nodule Analysis 2016 (LUNA16) [10] dataset, the Decathlon lung dataset, and NSCLC radio genomics [11] were the three publicly available datasets with CT images that were used. For segmentation, the original LUNA16 dataset is inappropriate, since it contains 888 sets of 3D CT images for lung nodule detection and the large true nodules are unlabeled, so 165 of these small true nodules are selected to construct a synthetic dataset using GAN for segmentation. First, a pre-trained variant model of U-net was used for lung segmentation. Subsequently, using the GAN

model, 3D images of nodules were generated. And after nodules were generated, graph cut and Gaussian mixture were used to generate segmentation labels. Then, merging the generated nodal CT images with the original CT images. Finally, a total of 165 lung nodules were generated.

Table 1. Segmentation metrics of test set using Decathlon-small

Model	DSC mean	SD	JI mean	SD	SE mean	SD	SP mean	SD
W/o PM, epoch 100	0.5124	0.2857	0.3928	0.2555	0.4671	0.2838	0.99997	0.00007
W/$\boldsymbol{PM_{300}}$, epoch 100	0.5702	0.2513	0.4400	0.2388	0.5183	0.2524	0.99996	0.00007
W/$\boldsymbol{PM_{500}}$, epoch 100	0.5608	0.2681	0.4352	0.2491	0.5190	0.2761	0.99997	0.00006
W/o PM, epoch 300	0.5515	0.2742	0.4274	0.2510	0.4950	0.2811	0.99997	0.00007
W/$\boldsymbol{PM_{300}}$, epoch 300	0.6090	0.2412	0.4781	0.2361	0.5511	0.2510	0.99997	0.00007
W/$\boldsymbol{PM_{500}}$, epoch 300	0.5844	0.2607	0.4574	0.2466	0.5215	0.2719	0.99997	0.00006
W/o PM, epoch 100	0.5525	0.2855	0.4316	0.2573	0.5109	0.2851	0.99997	0.00006
W/$\boldsymbol{PM_{300}}$, epoch 500	0.5536	0.2856	0.4338	0.2629	0.4934	0.2878	0.99998	0.00005
W/$\boldsymbol{PM_{500}}$, epoch 500	0.5670	0.2731	0.4435	0.2557	0.5089	0.2786	0.99997	0.00006

The open-source model nnUnet [12] was used to perform deep learning segmentation, and since the original version of nnUnet did not have a migration learning feature, its code was modified. First, two pre-trained models were constructed using the LUNA16 dataset which are PM300 AND PM500, followed by migration learning on each of the three decathlon datasets. In addition, models without migration learning were constructed to evaluate the model effects.

The Dice Similarity Coefficient (DSC) was used to assess the segmentation models. As additional evaluation metrics, the Jaccard index (JI), sensitivity (SE), and specificity (SP) were computed. It has been experimentally demonstrated from Table 1 that the use of this pre-trained model increases lung cancer segmentation accuracy and outperforms large models in medium and small models, and also demonstrates that this approach may be more effective on small data sets since the DSC means of PM300 at epoch 300 comes to 0.609 which is 0.06 higher than the without pretrained model. Besides the pre-trained model requires fewer iterations compared to the model without pre-training which could reduce the training time. To sum up, the advantages of this experiment are that by using GAN to generate the dataset, the task of manual annotation by physicians is reduced, and also by using this dataset to build a pre-trained model for migration learning to achieve higher segmentation accuracy. It also performs well on small datasets.

3. Different categories of model for lung cancer detection

3.1. few-shot U-Net deep learning model

Machine learning has proved quite effective in applications that require a large amount of data, but it frequently struggles with tiny data sets as a result of overfitting [13]. Few-Shot Learning (FSL) is one of the solutions of this issue. Few-shot Learning anticipates that machine learning models who have been trained with a large amount of data from a certain category can quickly learn from a limited amounts of labeled samples when encounter with a new category. In contrast to conventional supervised learning, Few-Shot Learning enables the model to distinguish between two images with a minimal amount of training data instead of enabling the model to recognize images. Another distinction between supervised learning and Few-Shot Learning is that the sample categories in the test set in former appear in the training set, while the data in Query set in latter pertain to unknown categories.

N. E. Protonotarios et al [14] employed Few-shot learning to address the limitations of supervised learning, which requires a large amount of training data and the inability of the model to make performance improvements after the training. It is achieved by integrating Few-shot learning into the U-Net architecture and producing an online supervised learning solution that allows dynamic model fine-tuning. U-Nets is a symmetric architecture based on FCNs [15], which excels in various types of cancer detection. Its architecture consists of two main sections, an encoder, which uses multi-layer

convolution and maximum pooling to obtain the image's global features, and a decoder, which is utilized to turn the features of the picture into segmented maps. And, FSL was improved, since the model is prone to errors when classifying new categories, the fundamental notion is that the end user can select some samples with classification errors, correct their segmentation maps, and retrain them again. Utilizing a similar feedback method repeatedly, the model's performance is enhanced.

The researchers tested their suggested two-channel U-Net model with a single-channel U-Net in order to confirm the model's superiority, Table 2 shows an increase in the model's f1 score and IoU of 7.68% and 5.74%, respectively. Additionally, this model performs better than the CNN and co-learning U-Net models due to increases in its f1 score and IoU of 10.81% and 9.28%, respectively.

Table 2. Performance metrics of FSL U-Net against other models

Model	Accuracy ± 95%CI	Precision ± 95%CI	Recall ± 95%CI	F1 ± 95%CI	IOU ± 95%CI
U-Net CT	99.11 ± 0.04	59.37 ± 2.62	40.35 ± 2.19	44.31 ± 2.25	35.26 ± 1.95
U-Net PET	99.58 ± 0.07	46.22 ± 2.68	25.36 ± 2.12	26.90 ± 1.90	18.34 ± 1.49
CNN	98.79 ± 0.07	66.47 ± 3.14	46.21 ±2.58	48.00 ± 2.47	36.79 ± 2.05
Co-learning	98.91 ± 0.06	73.63 ± 3.56	42.46 ± 2.46	48.85 ± 2.66	38.20 ± 2.22
U-Net PET/CT	98.92 ± 0.06	76.92 ± 3.48	46.27 ± 2.53	52.29 ± 2.67	40.71 ± 2.23
U-Net FSL	99.38 ± 0.02	71.47 ± 1.98	62.70 ±1.91	63.78 ± 1.81	52.83 ± 1.66

3.2. Small Cell Lung Cancer Segmentation via Hybrid Segmentation Network

Table 3. Results of the testing set's evaluations of various methodologies

		2D CNN	M-net	3D CNN	HSN
DSC	median	0.751	0.840	0.844	0.898
	mean ± std	0.692 ± 0.190	0.7889 ± 0.123	0.840 ± 0.049	0.888 ± 0.033
Sensitivity	median	0.690	0.849	0.863	0.889
	mean ± std	0.690 ± 0.193	0.819 ± 0.125	0.830 ± 0.076	0.872 ± 0.059
Precision	median	0.845	0.845	0.863	0.935
	mean ± std	0.766 ± 0.201	0.781 ± 0.154	0.856 ± 0.060	0.909 ± 0.048

Table 4. Results of several ablation experiments evaluated on the testing set

		HSN	HSN_Dice	HSN-3D	HSN-N	HSN-L
DSC	median	0.898	0.893	0.889	0.882	0.885
	mean ± std	0.888 ± 0.033	0.881 ± 0.046	0.877 ± 0.055	0.872 ± 0.064	0.869 ± 0.068
Sensitivity	median	0.889	0.898	0.893	0.871	0.901
	mean ± std	0.872 ± 0.059	0.880 ± 0.068	0.874 ± 0.073	0.850 ± 0.083	0.878 ± 0.075
Precision	median	0.935	0.896	0.888	0.918	0.881
	mean ± std	0.909 ± 0.048	0.886 ± 0.059	0.884 ± 0.062	0.901 ± 0.069	0.866 ± 0.088

Convolutional neural networks (CNN), as the cornerstone of deep learning who achieved breakthrough results in computer vision in recent years, are widely used in image classification [15]. However, it has clear drawbacks when used for semantic segmentation because the model only considers local features while ignoring global features. To resolve the aforementioned issues, WEI CHEN, HAIFENG WEI et al [16] present in their research a hybrid segmentation network (abbreviated as HSN) based on CNN. The model is divided into four sections, first is Spatiotemporal-separable 3D (S3D) convolution whose basic principle is to decompose a 3D convolution into two subsequent convolutions: a 2D convolution for learning spatial characteristics and a 1D convolution for learning temporal features. It is, in essence, the decomposition of a $k \times k \times k$ convolution into a $k \times k \times 1$ convolution and a $1 \times 1 \times k$ convolution. Thus, substantial savings in computing costs are achieved. The Multiscale Separable Convolution (MSC) Block is the next structure that is suggested. It is based on the Inception-ResNet [17] architecture and S3D convolution, and it contains four cascaded branches with various-sized filters that are connected by residuals. Then, To capture fine-

grained characteristics, the Dilation Convolution is used to create a broad receiver field at the output while preserving a high spatial resolution. The inflated convolution introduces an inflated rate parameter in comparison to the regular convolution, which is used to insert r-1 zeros between two neighboring filter values before convolving them through the up-sampling filter. Finally, Hybrid features fusion Module (HFFM) fuses 3D and 2D features efficiently to enable the network to train both the 3D CNN and the 2D CNN simultaneously. In contrast to 2D CNN, which creates two-dimensional feature maps, 3D CNN creates volumetric feature maps. Consequently, these properties cannot be simply combined. They sample a stack of neighboring slices from a single CT volume along the z-axis for 2D CNN. Next, intercept some adjacent slices from 3D CNN that have the same index as 2D CNN. Then the dimensions are arranged and the two are stitched together to get the fused features.

The experimental results are shown in the Table 3 above. Generalized Dice loss (GDL) [18] function is employed because of its greater robustness. The author compares the results of his model with 2D CNN, M-NET, and 3D CNN. After 100 epochs, it was discovered that HSN has the highest Dice similarity score, sensitivity, and precision among the three models. The improved accuracy is attributed to HSN's ability to incorporate 2D and 3D features into a single model. To further verify the superiority of the model, the researchers conducted a series of ablation experiments, including HSN_DICE trained with Dice loss, HSN-3D with standard 3D convolution, HSN-N with normal convolution instead of dilation convolution, and HSN-L with greater dilation rate. As Table 4 clearly demonstrates, the original HSN has the best accuracy.

4. Discussion

Although automated CT imaging makes a huge contribution and benefits to human society, it is the duty of researchers to make improvements based on the performance of automated CT imaging. Additionally, automated CT imaging mainly focuses on the procedures before the treatment of lung cancer, it is a challenge for all scientists to think of methods dealing with results after the treatment of lung cancer.

In terms of the Fig 3., it also indicates that the proposed DL network and algorithm generally yield comparable measurements to a human radiologist with 240 (98.4%). However, the disadvantages of this algorithm should be admitted to a certain degree. For example, the size of the tumor might be overestimated by DL algorithm since the data training is not enough, and the semi-automated CT imaging techniques are not mature, and the selection of center point is also a challenging task for the DL network [3].

The GAN model has made significant progress in addressing the issue of small medical datasets, and the datasets generated through it help the model to learn and reduce the task of physicians, however, it also has some obvious disadvantage that it can only generate specific data. Additionally, comparing the two lung cancer detection models, they both have the advantages of improved segmentation accuracy and the application of small sample learning also reduces the training time. However, they also have some drawbacks, including the need to improve the model's generalization by training with more categories of datasets and the need to confirm whether segmentation is effective for other cancers.

5. Conclusion

This paper review a variety of methods for the detection and segmentation of lung cancer, including Automated CT Imaging, Semi-automated CT Imaging, Transfer Learning and Generative adversarial network, few-shot U-Net deep learning model and Hybrid Segmentation Network. Among them, fully automated detection is an innovative method, which performs better in image interpretation and tumor localization compared to the original method. And semi-automatic detection is a method based on deep learning that can process all kinds of images and can produce the same

results as human experts, and is widely used in the medical field. Besides, GAN is an effective method to solve the problem of too little data, and the data generated by it helps machine learning models to be trained. While the Few-Shot Learning (FSL) is another solution to this problem, which is characterized by the ability to accomplish classification tasks with very small data sets. Finally, another effective lung cancer detection model is Hybrid Segmentation Network (HSN), which combines the advantages of 2dcnn and 3dcnn, retains global features, and has better segmentation accuracy. However, there is still room for improvement in the current work. For instance, segmentation accuracy needs to be raised, which can be done by more effectively gathering pre-processed data and refining the model. Additionally, in the future, researchers believe that the introduction of credit that is used to evaluate the segmentation difficulty, size of the tumor, risky level, and location of the tumor can remind doctors that how much energy they should put in the medical operation.

References

[1] Primakov, *et al.* 2022. *Automated detection and segmentation of non-small cell lung cancer computed tomography images.*

[2] Lahans, T. 2007 *Integrating Conventional and Chinese Medicine in Cancer Care.*

[3] Woo, M., Devane, A. M., Lowe, S. C., Lowther, E. L., and Gimbel, R. W. 2021. *Deep learning for semi-automated unidirectional measurement of lung tumor size in CT.*

[4] Shaikh, F., and Rao, D. 2022 *Prediction of Cancer Disease using Machine learning.*

[5] Borrelli, P., Góngora, J. L. L., Kaboteh, R., Enqvist, O., and Edenbrandt, L. 2022 *Automated classification of PET-CT lesions in lung cancer: An independent validation study.*

[6] Albawi, S., Mohammed, T. A., and Al-Zawi, S. 2017 *International Conference on Engineering and Technology paper.*

[7] Li, M., Jirapatnakul, A., Biancardi, A., Riccio, M. L., Weiss, R. S., and Reeves, A. P. 2013 *Growth Pattern Analysis of Murine Lung Neoplasms by Advanced Semi-Automated Quantification of Micro-CT Images.*

[8] M. Nishio, K. Fujimoto, H. Matsuo, C. Muramatsu, R. Sakamoto, and H. Fujita. 2021 *Lung cancer segmentation with transfer learning: usefulness of a pretrained model constructed from an artificial dataset generated using a generative adversarial network.*

[9] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al 2014 *Generative Adversarial Nets.*

[10] Setio, A. A. A., Traverso, A., de Bel, T., Berens, M. S. N., Bogaard, C. v. d., Cerello, P., et al. 2017 *Validation, Comparison, and Combination of Algorithms forAutomatic Detection of Pulmonary Nodules in Computed Tomography Images: The LUNA16 challenge.*

[11] Bakr, S., Gevaert, O., Echegaray, S., Ayers, K., Zhou, M., Shafiq, M., et al. 2018 *A Radiogenomic Dataset of Non-small Cell Lung Cancer.*

[12] Isensee, F., Petersen, J., Klein, A., Zimmerer, D., Jaeger, P. F., Kohl, S., et al. 2018 *NNU-NET: Self-Adapting Framework for U-Net-Based Medi*cal Image Segmentation.

[13] J. Snell, K. Swersky and R. Zemel 2017 *Prototypical Networks for Few-shot Learning.*

[14] N. E. Protonotarios, I. Katsamenis, S. Sykiotis, N. Dikaios, G. A. Kastis, S. N. Chatziioannou, M. Metaxas, N. Doulamis and A. Doulamis 2022 *A few-shot U-Net deep learning model for lung cancer lesion segmentation via {PET}/{CT} imaging.*

[15] Long J, Shelhamer E and Darrell T 2015 *Fully convolutional networks for semantic segmentation*

[16] Chen, W., Wei, H., Peng, S., Sun, J., Qiao, X., & Liu, B. 2019 *HSN: hybrid segmentation network for small cell lung cancer segmentation.*

[17] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi 2017 *Inception-v4, inception-ResNet and the impact of residual connections on learning.*

[18] W. R. Crum, O. Camara, and D. L. G. Hill. *2006 Generalized overlap measures for evaluation and validation in medical image analysis.*