

LUNG LESION DETECTION AND SEGMENTATION BASED ON CNN AND RESNET APPROACH WITH 3D IMAGES

Mrs A. Merry Ida ME, Mrs.M Suji M. Tech, Mrs.S. Antany Nimitha ME.

Assistant professor, Department of Computer Science and Engineering Loyola Institute of Technology & Science, Thovalai. Email: ida.cse@lites.edu.in

Assistant professor, Department of Computer Science and Engineering Loyola Institute of Technology & Science, Thovalai. Email: suji.cse@lites.edu.in

Assistant professor, Department of Computer Science and Engineering, Loyola Institute of Technology & Science Thovalai. Email: Nimitha99antany@gmail.com

Abstract— This project aims to develop a robust and efficient method for the detection and segmentation of lung lesions from medical imaging data, particularly focusing on Computed Tomography (CT) scans. Applied deep learning and CNN to detecting and classified lung disease using imagery data. We used the images that belong to 4 categories: healthy, covid-19, viral pneumonia, and bacterial pneumonia. Each category is consisting of 133 images and we used it to develop model that could detect and classify the images in less than 1 minutes. For the image dataset that is used on this project put it on google drive. I put label in each image's categories with numbers from 0 to 3 (0 = covid-19, 1 = normal, 2 = viral pneumonia, 3 = bacterial pneumonia) 1. Understand the Problem Statement and Business Case Deep learning has been proven to superior in detecting and classifying disease using imagery data. Skin cancer could be detected more accurately by Deep Learning than by dermatologist (2018).

Human dermatologist can detect skin cancer with 86.6% accuracy while deep learning can detect skin cancer with 95% accuracy. (Reference: "Computer learns to detect skin cancer more accurately than doctors". The Guardian. 29 May 2018) In this project, we try to develop a model that could detect and classify lung disease using 133 X-Ray images that belong to 4 classes: Healthy, Covid-19, Bacterial Pneumonia, and Viral Pneumonia. To import our data, we used image generator to generate tensor images data and normalize them. We used 428 images for training (80%) and 104 images for our validation (20%). Before we generate our data, we perform shuffling to prevents the model from learning the order of the training. We generate a batch of 40 images and labels. The following is label names for each classification: 0 = Covid- 19 1 = Normal 2 = Viral Pneumonia 3 = Bacterial Pneumonia.

Keywords—CNN, Resnet50, Lung Cancer

I. INTRODUCTION

Lung cancer is one of the malignant tumors that pose the greatest menace to human life and health. As reported in a statistical analysis of the global cancer burden, lung cancer accounts for nearly 11.4% and 18% of 19.3 million new cancer cases and 10.0 million cancer deaths. The five-year survival of early-stage lung cancers is significantly higher than that of advanced lung cancers, thus early detection and timely treatment are effective solutions for lung cancer. Chest computed tomography (CT) is a frequent way for non-invasive lung cancer screening, and it contributes to decreasing the mortality of high-risk individuals. Lung nodule is the principal clinical manifestation of early lung cancer, and CT-based nodule location detection is an indispensable procedure in lung cancer screening.

In clinical diagnosis, to conduct a thorough examination, radiologists are usually required to read dozens or even hundreds of CT slices for each patient in a slice-by-slice manner, while such work is labor-intensive and easy to cause operator bias. As a result, computerized lung nodule position detection is an active research topic in the medical imaging analysis field, which aims at assisting clinicians to improve diagnostic efficiency. As illustrated in lung nodules change greatly in scale, appearance and intensity, and they may occur anywhere in lungs and are often surrounded by complex background tissues. Therefore, it is crucial to extract 3D multi-scale discriminative features for achieving accurate nodule detection. Multimodal medical imaging has become incredibly common in biomedical imaging. From multimodality clinical visual information, meaningful information has been derived using clinical image classification. Computed tomography (CT) and Magnetic resonance imaging (MRI) are some imaging approaches. Different

imaging technologies provide different imaging information for the same part. Traditional ways of illness classification are effective, but in today's environment, 3D images are used to identify diseases. Compared to 1D and 2D images, 3D images have a very clear vision.

II 3-D PREPARATION OF LUNG NODULE

In general, a CNN consists of convolutional, pooling and fully-connected layers to extract multi-level learnable representations. They are

learned jointly, in an end-to-end manner, to solve a particular task. Unlike the conventional CNN, each channel in a 3D CNN is actually a 3D feature volume, rather than a 2D feature map. The convolutions and pooling of 3D CNN are operated in a cubic manner. We introduce some of the basic components of 3D CNN in the following paragraphs.

A. 3D CONVOLUTIONAL LAYER

We use $y = \text{conv}(x, w)$ to denote the convolutional function operated by the 3D convolutional layer, where x represents the original data or feature maps that the convolutional function operates on, w denotes the filters and y denotes the output of the convolutional layer.

B. ACTIVATION FUNCTION

The activation function, which is applied to each component of a feature map, introduces non-linearity in a CNN. We use the Rectified Linear Unit (ReLU) as the activation function in this paper. It works as follows using $y = \text{ReLU}(x)$ to represent it.

C. 3D POOLING LAYER

The pooling layer is another important operator in a CNN. A pooling operator runs on individual feature channels, coalescing nearby feature values into one via the application of a suitable operator. Common choices for this include max-pooling or average-pooling. We prefer to use max-pooling, just like several other researchers, which is defined as (3) where π denotes the pooling size. $y = \text{MaxP}(x)$ is used to represent it. As can be seen from the above equations, the output of each layer is also a 5D tensor, where the meaning of each dimension is the same as the input x .

III IMAGE PREPROCESSING

Image preprocessing organizes images before they are used in model preparation and induction. The goal of preprocessing is to improve the quality of the image so that it can be investigated more thoroughly. It

includes, but is not limited to, rectifications for resizing, arranging, and shading.

A. THRESHOLDING

Thresholding is a non-linear operation that changes a grayscale image into a binary image in which the two levels are allocated to pixels that are either below or above the set threshold value. It mainly converts an image from shading or grayscale into a twofold picture. Thresholding is used to convert a low-contrast lung scan to a high-contrast lung scan. Thresholding is also a very effective tool in image segmentation. Its purpose is to convert grayscale images to binary format. It takes the colorful or grayscale lung scans and turns them into binary scans. It diminishes the intricacy, works on acknowledgment and grouping, and changes the pixels to simplify the picture.

IV LUNG SEEK FRAMEWORK

The original CT image and preprocess it, unify the resolution, remove the noise, cavity and other interfering factors. And we extract the lung parenchyma to reduce the search space of the image. Then as shown in the nodule detection part, a 3D region proposal network with 3D SK-Resnet and a U-net shape structure was used to extract the features of lung nodules, which contain the three-dimensional coordinates, diameter and confidence score of these detected nodules performance of the overall detection system.

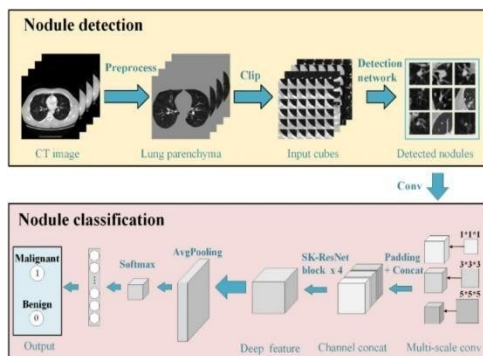


Figure: 1. A Nodule detection

3D SK-ResNet is made of 3D residual network and Selective Kernel module, it can effectively solve the eliminating gradient disappearance problem by using the quick connection of residual learning and obtain the recalibration feature by using SK block (Selective Kernel block). At the same time, the SK block can automatically acquire the importance of each feature channel through learning, so it can selectively stress useful features and inhibit less useful features.

Lung nodules possess great discrepancies in size, appearance and density, and exploring 3D multi-scale discriminative representations is a remarkable approach to boost detection performance. Given this fact, a multi-kernel driven 3D convolutional neural network is proposed to fulfil automated lung nodule detection, and the general structure of the MK-3DCNN model is displayed. As exhibited in the MK-3DCNN framework uses a UNet-like encoder-decoder structure as the backbone network to utilize the multi-layer features of the deep model, and introduces a region proposal network (RPN) as the output module to generate high-quality proposals. In the encoder part of the MK-3DCNN, a multi-kernel joint learning model is developed to capture multi-scale lung nodule information. Furthermore, a residual learning module combining a multi-model mixed pooling operation is designed to learn more comprehensive descriptions of nodule CT images, which could relieve the problem of information loss caused by the traditional single-model pooling manner. In addition, the decoder part mainly involves three components, including the deconvolution layer, residual learning unit, and concatenation operation.

V LUNG NODULE SEGMENTATION

Lung nodules possess great discrepancies in size, appearance and density, and exploring 3D multi-scale discriminative representations is a remarkable approach to boost detection performance. Given this fact, a multi-kernel driven 3D convolutional neural network (MK-3DCNN) is proposed to fulfil automated lung nodule detection, and the general structure of the MK-3DCNN model is displayed. As exhibited the MK-3DCNN

framework uses a UNet-like encoder- decoder structure as the backbone

network to utilize the multi-layer features of the deep model, and introduces a region proposal network (RPN) as the output module to generate high-quality proposals. In the encoder part of the MK- 3DCNN, a multi-kernel joint learning model is developed to capture multi-scale lung nodule information. Furthermore, a residual learning module combining a multi-model mixed pooling operation is designed to learn more comprehensive descriptions of nodule CT images, which could relieve the problem of information loss caused by the traditional single- model pooling manner. In addition, the decoder part mainly involves three components, including the deconvolution layer, residual learning unit, and concatenation operation.

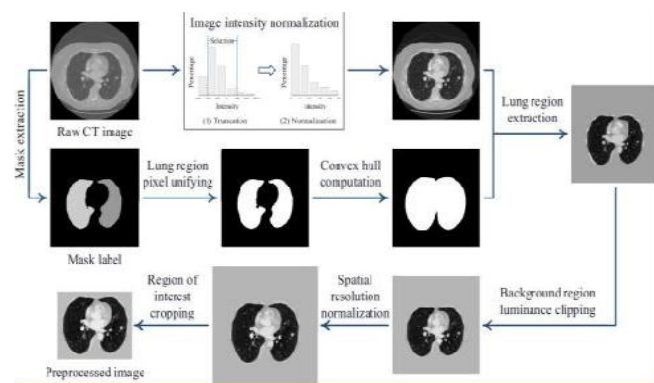


Figure 2. B General structure of the MK-3DCNN model

A LUNG NODULE DETECTION

Radiologists have long-standing methods for locating lung nodules in patients with lung cancer, such as computed tomography (CT) scans. Radiologists must manually review a significant amount of CT scan pictures, which makes the process time-consuming and prone to human error. Computer-aided diagnosis (CAD) systems have been created to help radiologists with their evaluations in order to overcome these difficulties. These systems make use of cutting-edge deep learning architectures. These CAD systems are designed to improve lung nodule diagnosis efficiency and accuracy. In this study, a bespoke convolutional neural network (CNN) with a dual attention

mechanism was created, which was especially crafted to concentrate on the most important elements in images of lung nodules. The CNN model extracts informative features from the images, while

the attention module incorporates both channel attention and spatial attention mechanisms to selectively highlight significant features. After the attention module, global average pooling is applied to summarize the spatial information. To evaluate the performance of the proposed model, extensive experiments were conducted using benchmark dataset of lung nodules. The results of these experiments demonstrated that our model surpasses recent models and achieves state-of-the-art accuracy in lung nodule detection and classification tasks.

VI CONCLUSION

In this paper, a multi-kernel driven 3D convolutional neural network (MK- 3DCNN) is developed for the automatic detection of lung nodules in thoracic CT scans. The MK-3DCNN method adopts a residual learning-based encoder-decoder structure as the backbone to exploit the multi-layer features of the deep network. Different from previous traditional convolutional networks with fixed kernel size, a multi-kernel joint learning block is designed to drive the detection model to capture 3D multi-scale spatial information from the nodule CT images with variable lesion sizes and shapes. In addition, a multi-mode mixed pooling strategy is

proposed to surrogate the conventional single-mode pooling way, the designed pooling method reasonably incorporates three different types of pooling operations, including max pooling, average pooling, and centre cropping pooling, and they can complement each other to attain more comprehensive nodule CT image representations. To fully evaluate the validity of the presented MK- 3DCNN, systematic experiments are performed on the public dataset LUNA16 and the clinical dataset CQUCH-LND, and experimental results indicate the MK- 3DCNN method outperforms some SOTA nodule detection approaches and possesses a good generalization ability in the clinical practice.

Different feature extraction models can be used in Faster R-CNN. In order to test which model is more suitable, VGG16, ResNet50 and ResNet101 are compared. We divide the training set into two parts. One part is used for training Faster R-CNN and the scan size is 68. The other part is used to get the validation result and the scan size is 29. It can be seen that ResNet 101 module gets the best performance compared with the other two modules.

VI CHALLENGES AND FUTURE RESEARCH DIRECTIONS

As with some previous related work, our study focuses on developing a one-stage end-to-end 3D model for automated detection of lung nodules in chest CT scans. To further evaluate the performance of the proposed MK-3DCNN method, and considering limited nodule samples, we introduce a 3D self-supervised transfer learning method to conduct an additional false positive reduction (FPR) experiment on the benchmark dataset LUNA16.

In the FPR process, a 3D encoder-decoder structure with residual connection is used to implement self-supervised pre-training to learn valuable representation information from large amounts of randomly cropped unlabeled data, which helps reduce the dependence on labelled samples. Then, the pre-trained encoder part is transferred as the feature extractor, and the global average pooling operation is exploited to convert the feature map generated from the last convolutional layer of the encoder into a 512-dimensional feature vector. Finally, a classifier consisting of two fully connected layers (the number of neurons is respectively set to 512 and 256) and a Sigmoid unit is constructed to achieve the FPR. In this experiment, five image perturbation strategies (nonlinear transformation, local pixel shuffling, local pixel swapping, inner pixel cutout, and outer pixel cutout) are integrated to enhance the image representation ability of the self-supervised learning network. Furthermore, conventional image rotation and image flipping approaches are used for data augmentation. The mean square error loss function and

the stochastic gradient descent (SGD) optimizer with an initial learning rate of 1.0 are selected for self-supervised pre-training, and the cross-entropy loss function and the adaptive moment estimation (Adam) optimizer with an initial learning rate of 0.001 are adopted for the FPR training. The learning rate will be halved when the model performance is not improved over 10 epochs, the input size is set to 64×64×32, the batch size is set to 32, and the early stop mechanism is employed to get a better model.

REFERENCES

- [1] Sung H., Ferlay J., Siegel R. L., et al., "Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *Cancer J. Clin.* 71, 209–249 (2021). 10.3322/caac.21660.
- [2] Wu G. X., Raz D. J., "Lung cancer screening," *Lung Cancer: Treat. Res.* 170, 1–23 (2016). 10.1007/978-3-319-40389-2_1.
- [3] Zhou Z., Gou F., Tan Y., et al., "A cascaded multi-stage framework for automatic detection and segmentation of pulmonary nodules in developing countries," *IEEE J. Biomed. Health Inform.* 26(11), 5619–5630(2022). 10.1109/JBHI.2022.3198509.
- [4] Huang H., Wu R., Li Y., et al., "Self-supervised transfer learning based on domain adaptation for benign-malignant lung nodule classification on thoracic ct," *IEEE J. Biomed. Health Inform.* 26(8), 3860–3871(2022).
- [5] Mkindu H., Wu L., Zhao Y., "Lung nodule detection of ct images based on combining 3d-cnn and squeeze-and-excitation networks," *Multimed. Tools Appl* 82(17), 25747–25760 (2023). 10.1007/s11042-023-14581-0.
- [6] Agnes S. A., Anitha J., Solomon A. A., "Two-stage lung nodule detection framework using enhanced unet and convolutional lstm networks in ct images," *Comput. Biol. Med.* 149, 106059 (2022). 10.1016/j.compbiomed.2022.106059.
- [7] Guo Z., Zhao L., Yuan J., et al., "Msanet: multiscale aggregation network integrating spatial and channel information for lung nodule detection," *IEEE J. Biomed. Health Inform.* 26(6), 2547–2558(2021). 10.1109/JBHI.2021.3131671.
- [8] Zhu L., Zhu H., Yang S., et al., "Pulmonary nodule detection based on hierarchical-split hrnet and feature

- pyramid network with atrous convolution,” *Biomed. Signal Process. Control*.85,105024(2023).
10.1016/j.bspc.2023.105024.
- [9] Zhao D., Liu Y., Yin H., et al., “An attentive and adaptive 3d cnnfor automatic pulmonary nodule detection in ct image,” *Expert Syst. with Appl.* 211, 118672 (2023).
10.1016/j.eswa.2022.118672.
- [10] MacMahon H., Naidich D. P., Goo J. M., et al., “Guidelines for management of incidental pulmonary nodules detected on ct images: from the fleischner society 2017,” *Radiology* 284(1), 228–243(2017).
10.1148/radiol.2017161659.
- [11] Chen Y., Hou X., Yang Y., et al., “A novel deep learning model based on multi-scale and multi-view for detection of pulmonary nodules,” *J Digit Imaging* 36(2), 688–699(2023). 10.1007/s10278-022-00749-x.
- [12] Wu R., Liang C., Li Y., et al., “Self-supervised transfer learning framework driven by visual attention for benign–malignant lung nodule classification on chest ct,” *Expert Syst. with Appl.* 215, 119339 (2023).
- [13] Setio A. A. A., Traverso A., De Bel T., et al., “Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge,” *Med. Image Anal.* 42, 1–13(2017). 10.1016/j.media.2017.06.015
- [14] Huang Y.-S., Chou P.-R., Chen H.-M., et al., “One-stage pulmonary nodule detection using 3-d dcnn with feature fusion and attention mechanism in ct image,” *Comput. Methods Programs Biomed.* 220, 106786 (2022).
10.1016/j.cmpb.2022.106786.
- [15] Han Y., Qi H., Wang L., et al., “Pulmonary nodules detection assistant platform: An effective computer aided system for early pulmonary nodules detection in physical examination,” *Comput. Methods Programs Biomed.* 217, 106680(2022).