End-to-End Deep Learning Strategies for Computer-Aided Lung Cancer Detection Systems

Bohdan Chapaliuk [1,*], Yuriy Zaychenko [1]

[1]Institute for Aplied System Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

* Corresponding Author: bohdan.chapaliuk@gmail.com (Bohdan Chapaliuk)

Abstract

Lung cancer is one of the most aggressive types of cancer, and the possibility to detect it at an early stage can save a lot of patient lives worldwide. Building an automated lung cancer detection system can help to speed up the process of cancer detection and save human lives. This article considers three different approaches used to design and build lung cancer detection systems. Mainly, these approaches use either the 2D convolutional neural networks (CNN) with a multi-instance learning task, 2D CNN with recurrent neural networks, or detection system pipelines with 3D CNNs, which can consist of single or multiple stages. The article presents the results of the experiments for each of the approaches. Finally, the obtained results and results from other recent papers are combined to compare all existing lung cancer detection system architectures and evaluate how different design decisions impact overall system accuracy.

Keywords: *deep learning, convolutional neural network, lung cancer detection system*

1. Introduction

This paper is an extension of the original work presented on the SAIC 2018 conference [1] and provides a more detailed review of the considered end-to-end models and adds more detailed results and comparison. The main purpose of the paper is to consider the architectures of a computer-aided lung cancer detection system and investigate how neural network architecture and system pipeline design impact overall classification accuracy.

The paper consists of four sections that give full information about lung cancer detection system specifics. The first section contains three subsections, which describe approaches used to build a lung cancer detection system. Also, the second section contains information about related works and convolution neural networks (CNN) trained in the scope of the paper research. In the second section, we describe the type of data usually used for lung cancer screening systems and core datasets used to train neural networks. The third section describes common preprocessing steps used to train CNN in this article as well as results comparison. Finally, we provide a conclusion with a comparison of the different lung cancer detection system architectures and discuss which of them provides the best accuracy. Also, we conclude which design decision impacts overall system accuracy.

2. 2. Related Works

Lung cancer screening uses CT scans of the human chest, which represent a high-resolution 3-dimensional image. This image used as input data for the lung cancer detection system, which is usually built using artificial neural networks. This fact imposes memory and computational constraints to the detection system since the whole CT scan cannot be fit into the video card memory for a large enough model, and 3D data computation and memory usage grows cubically. This section discusses strategies that are issued to work with such high-resolution data and describes models used to get experimental results presented at the end of the work as well as some of the state-of-the-art models from recent publications.

2.1. Two-dimensional convolution neural networks with a multi-instance learning task

A first and most straightforward strategy to build an end-to-end lung cancer detection system is to process each slice of the CT scan independently and then combine results using some predefined rules. In this case, the detection system works with 2D data, which consumes much less memory than an overall 3D patient image created from the CT scan slices. Also, such a strategy allows reusing the most popular pre-trained 2D convolution neural network (CNN) architectures like ResNet [2], DenseNet [3] or Xception [4] to speed up the training

process and improve the overall system accuracy. Each CT scan slice can be annotated, and result labels can be used to consider each slice independently.

However, Data Science Bowl 2017 (DSB 2017) dataset, which is extensively used in this research, has only one label per an entire patient's CT scan, which makes such a naive approach impractical. Therefore, the problem should be considered as a multi-instance problem, where the instance is one of the slices in the CT scan. The final decision about the cancer prediction can be made using a general multi-instance assumption, which says that if there exists an instance that is positive, the whole bag is positive. The whole bag is negative if all instances are negative [5].

In this paper, we follow the approach used for mammogram classification problem with sparse label assignment [6], except for the fact that trained model uses an entire slice from the CT scan as a single instance instead of the image patch, and final loss is calculated based on all slices in the patient scan that were passed through CNN. As a core model, we have used DenseNet [3], with modification to follow MIL task with sparse label assignment approach.

Let us assume that we have an input patient scan S, which contains k images. The goal of the model is to predict whether a patient's CT scan contains somewhere a malignant pulmonary nodule or not. Given that and assuming that the model considers an entire CT scan, we are solving a binary classification problem. For each CT scan slice, the output cancer prediction can be written as

$$r_n = \sigma(w \cdot CNN(k_m) + b) \qquad (1)$$

where σ is a sigmoid function, w is a logistic regression weight coefficient, CNN is a function which represents convolution neural network forward pass, k_m is the image in the patient scan at position m, and b is a bias.

Given cancer probability for each image in a patient scan, the overall cancer presence probability can be written as

$$p(y = 1|S,k) = max\{r_1, r_2, ..., r_n\} \qquad (2)$$

$$p(y = 0|S,k) = 1 - max\{r_1, r_2, ..., r_n\} \qquad (3)$$

where y is a true label from the dataset.

Finally, the loss function for the MIL model can be written as:

$$L = \frac{1}{N}\sum_{n=1}^{N}\left(-log\left(p(y_n|S,k)\right) + \mu\|\dot{r_n}\|_1\right) \qquad (4)$$

where N – is the number of patient scans, μ is a sparsity factor which represents a tradeoff between image importance and cancer sparsity assumption, and $\|\grave{r}_n\|_1$ is the L_1 norm of each image scan output vector.

The resulting loss is used to train a two-dimensional convolution neural network. In the scope of the paper 2D DenseNet [3] pre-trained on the ImageNet dataset were used to initialize the neural network at the beginning of the training process. During the training, the output of DenseNet for each CT scan slice was accumulated and passed to the loss function with a true label. Finally, this approach gives an accuracy level of only 0.62. Mainly, low accuracy value can be explained by the fact that CNN loses the 3D nature of information, which is an important part of building understanding about pulmonary nodule malignancy.

2.2. Recurrent neural network with attention and 2D CNN

To mitigate the problem of losing spatial information in the previous approach, we can take inspiration from the real-word radiologist, who decides whether a tumor or suspicious abnormal pulmonary nodules are present or not on the CT scan. Usually, a radiologist examines CT scan slices one by one, carefully combining information from neighborhood slices to understand pulmonary nodule malignancy. The same strategy can be achieved in a computer-aided lung cancer detection system by combining CNN and recurrent neural networks (RNN) in one model. In this case, we can still use transfer learning since we can reuse pre-trained 2D CNNs on ImageNet dataset like DenseNet [3]. Meanwhile, RNN can combine slice level features to retrieve features dependencies and spatial information. In the model, LSTM [7] or GRU [8] are commonly used as a recurrent neural network. An excellent example of applying such a strategy is shown by [9] work, which uses CNN, LSTM, and attention mechanism to achieve human-level accuracy on the brain tumor classification problem. As the work suggests, it leads to human-level performance, at least for the brain tumor classification problem. Nevertheless, the same approach can achieve great results in the lung cancer classification problem as well. In this paper, a similar neural network has been applied to the lung cancer detection problem. Trained model is shown on figure 1.
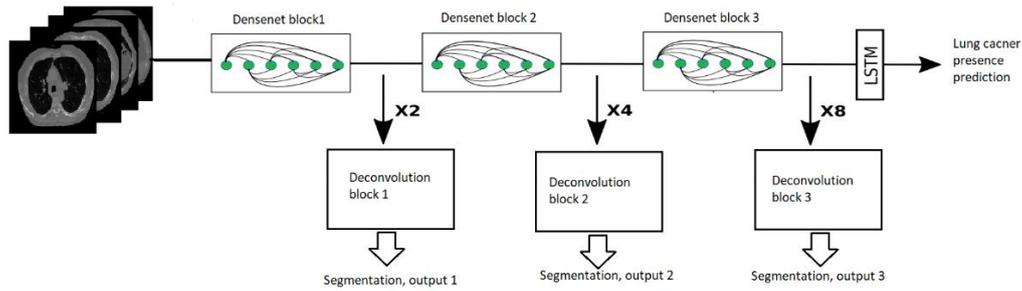
Figure 1. LSTM + DenseNet with attention mechanism neural network for lung cancer detection.

The main advantage of the method is that the neural network can learn features and their spatial dependencies in an end-to-end manner. Besides, the model allows for reusing radiologist knowledge about malignant nodule location. It helps speed the training process and improve overall model accuracy. Our experiments show that, without an attention mechanism, these models do not give better results than the end-to-end 3D convolution neural network discussed in the next section.

The main drawback of the discussed approach is a requirement for additional labeling to identify the location of the pulmonary nodules on the CT scan. Typically, additional labeling is an expensive part of the detection system development since labeling requires a massive amount of expert working time. Also, the model does not provide information about the location of the malignant pulmonary nodule, which can make a radiologist not to trust the detection system and thus avoid its usage.

2.3. Three-dimensional convolution neural network for lung cancer detection

The next available strategy and the most popular recently is to use the 3D nature of the CT scans and work with the data as with three-dimensional objects. For lung cancer detection problem, a CT scan represents a 3D view of the human chest, which contains an image of the patient's lungs. When all slices of the CT scan are stacked together, they form an image of the three-dimensional representation of the human chest. The resulting image can be used as an input to the neural network. Nevertheless, in addition to the lung tissue, the resulting image contains an image of vessels and bones, and therefore, typically lung cancer detection system uses additional preprocessing steps that clean them to simplify learning patterns. In case there is a massive amount of data available for the training stage, and there is enough computation power to build a large enough neural network, the stacked 3D might be used without vessels and bones cleaning steps which can allow CNN to learn more complex features.

To build a neural network that works with three-dimensional data, we can use 3D convolution operation and build 3D convolution neural networks like C3D [11] or 3D DenseNet [12] on an entire patient scan. In the scope of this article, these neural networks were trained and achieved 0.63 and 0.70% accuracy, respectively. 3D CNN networks were learned based on the resized 3D image since the original 3D view with a 3D CNN does not fit into the video card memory. An example of 3D DenseNet architecture can be found in figure 2.
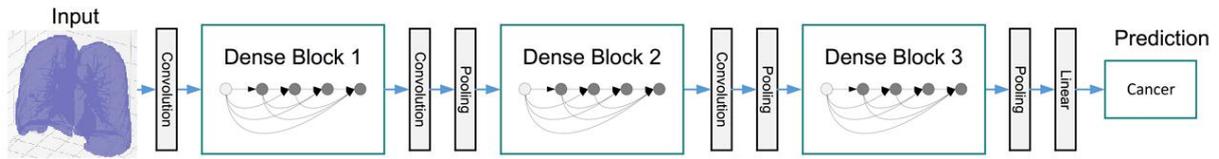

Figure 2. 3D DenseNet architecture trained on the entire 3D patient image created from CT scan slices

As results show, building and training model directly on the 3D image of lungs does not show good-enough results. One of the reasons for this is that there is a risk of losing some details during the resizing 3D scan to lower resolution, which might be useful to build stronger features. Also, there are not enough data in training set to allow the neural network to fetch more background about the nature of the malignant nodules. Therefore, to achieve better performance, a two-stage approach should be used, and as [13], [14] works show, it leads to the human-level performance of the lung cancer classification problem.

The main idea behind the two-stage approach is that the lung cancer detection task can be divided into two smaller tasks: segmentation and classification. In the scope of the segmentation task, the neural network finds out all pulmonary nodules in the CT scan and then passed them to the next classification step. On the classification step, the detection system chooses the most probable malignant pulmonary nodules and decides whether the cancer is present in the image using a multi-instance learning assumption, which is described in section 2.1. The benefit of introducing two stages into the lung cancer detection system is the ability to adapt any existing state-of-the-art segmentation models which can return a smaller patch of the original 3D image with pulmonary nodules instead of the entire lung image. It simplifies the development of the classification stage since smaller patches are much simpler to fit into the video card memory. Also, it allows a learning lung detection system with less amount of data. At the moment of writing the article, usually segmentation task is done by V-Net [15], Faster-RCNN [16] or Mask-RCNN [17] adapted to the 3D data [13]. Usually, the classification stage used simple CNN, which outputs a probability on whether the

pulmonary nodule is cancer or not. CNN on the classification stage shares weights with CNN from the segmentation stage to speed up the training process, improve classification and pulmonary nodule detection accuracy as well as reduce overfitting problem on both stages. As [13], [14] papers state, the two-stage lung cancer detection systems were able to achieve 0.87 and 0.92 accuracy level.

Even though the two-stage approach leads to human-level accuracy, it has several drawbacks. First, training a CNN for the segmentation stage requires additional labeling stuff to add information about positive and negative malignant pulmonary nodule candidates in the dataset. In the medical domain, labeling is an expensive and time-consuming work that requires a massive amount of radiologist working time. Second, on the stage of deciding cancer presence, the model loses global information about the lung and position of the pulmonary nodule candidate. Such information is helpful to understand pulmonary nodule malignancy and leads to a better understanding of cancer presence. Finally, the two-stage approach still ignores information on the pulmonary nodule growing process over time. The problem can be solved by adding even more stages into the pipeline.

A recent article in Nature Medicine magazine [10] describes a lung screening system that divides the lung cancer detection problem into four components or stages (see figure 3).
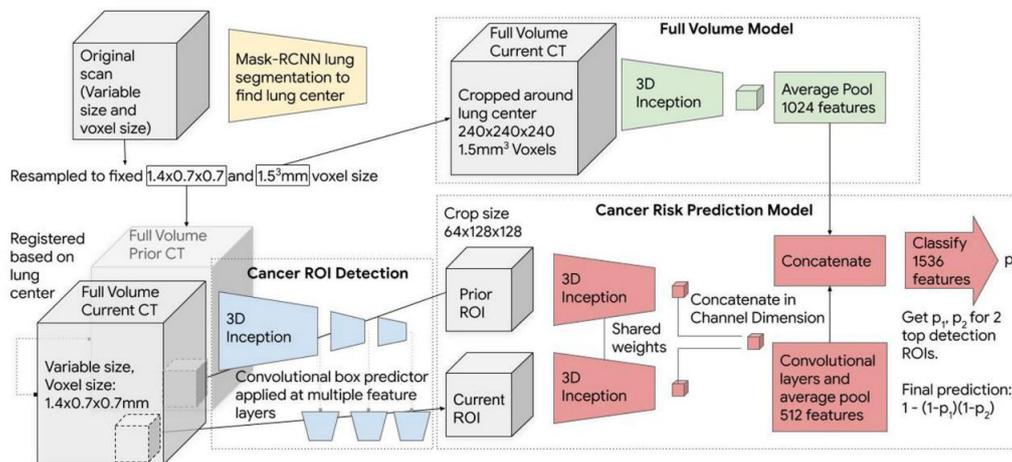

Figure 3. Recent [10] end-to-end lung cancer detection system that introduced even more stages to achieve better performance. (Image copied from the original paper)

First, a lung segmentation neural network like [17] is used to segment the lungs from the overall CT scan. The purpose of this stage is to clean all tissue things that are not related to lungs and leave only the lungs and related tissues on the image, which can be used further in the later stages. The second stage uses another CNN that generates pulmonary nodule region images in a similar way as on the first stage of the two-stage approach. The third stage uses an

end-to-end 3D convolution neural network that is trained based on the segmented lung image from the first stage in a similar manner as described at the beginning of this section. End-to-end 3D CNN generated features are used in the final stage to provide a more global context to the specific pulmonary nodule. Finally, a cancer risk prediction model uses generated features from the third stage and pulmonary nodule candidates from the second stage to predict the risk of cancer. Also, the last stage can use previous patient scans to improve prediction. As the diagram in figure 2 suggests, the detection system generates pulmonary nodule candidates for the previous CT scan and combines them with pulmonary nodule candidates for the current CT scan on the feature level. The resultant features were concatenated with features from the third stage and then used to perform a prediction about the lung cancer risk. The four-stage approach was able to achieve super-human accuracy on the used dataset in comparison to the average radiologist.

3. Data and Datasets

A standard tool to diagnose lung cancer is volumetric thoracic computed tomography (CT) [18]. CT scans consist of a series of X-ray images taken from different angles to produce cross-sectional (tomographic) images or slices of tissues, bones, or blood vessels in the human body. Combined CT slices form a 3D view, which is a high-resolution 3-dimensional representation of the human body part. Usually, a radiologist examines lung cancer by looking at scans one by one to find a tumor or abnormal areas with cancer characteristics. Such abnormal areas in lung CT scans are called pulmonary nodules. Most nodules seen in CT scans are not cancer, however pulmonary nodules with a diameter larger than 6 mm have a high risk to be dangerous. Often, a radiologist can investigate the previous scans of a patient or repeat CT scan to check whether some of the lung nodules are not growing over time. If the nodule has grown, it will mean that there is a high risk of lung cancer and related treatment should be arranged for the patient.

In the scope of the paper's experiments, LUNA [19] and Data Science Bowl 2017 (DSB 2017) [20] datasets have been used. DSB 2017 dataset contains CT scans for more than 1000 patients with "lung cancer presence" labels. One row of the dataset represents patient chest CT scan with an associated label. The label shows whether the CT scan contains cancer somewhere, which is determined after a year of the CT scanning. DSB dataset contains 1397 patient scans in the training dataset, 198 patient scans in the validation set, and 506 patient scans in the test set. The training set is highly unbalanced and contains 1035 samples of

patients who do not have lung cancer and 362 samples of patients with cancer. The LUNA dataset includes 1186 lung nodule labels in 888 patients that were annotated by radiologists. Lung nodule candidate labels are used to train CNN for segmentation task.

4. Experiments and Results

This section describes the experiments we performed for each of the strategies discussed in section 2. Experiments were performed using LUNA and DSB 2017 datasets, which was described in section 3. As hardware, we used a preemptible compute instance in the Google Cloud with one Nvidia Tesla P100 card, 8 GB RAM, and Intel CPU for a standard machine type. As a machine learning framework, Keras with Tensorflow as a backend has been used. Even though the models are different, there are standard preprocessing and data augmentation steps applied to them. Therefore, they are discussed in a separate section (section 4.1). Other sections discuss the training process specifics for the concrete strategy. Finally, the section provides the results summary to give an intuition about strategies accuracy level. Results summary includes results from other papers as well.

4.1. Data preprocessing steps

The unit of measurement in CT scans is the Hounsfield unit (HU), which is a standard quantitative scale for describing radiodensity. All CT scanners are carefully calibrated to measure this accurately. By default, the data in DSB 2017 dataset is not in the HU value, and therefore, they should be converted to the HU by multiplying the pixels with the rescale slope and adding the intercept. Fortunately, these values are conveniently stored in the metadata of the CT scans. Luna dataset does not require this step as CT scan pixel values are already in HU.

In HU, each tissue, bone, and blood has its values. This fact can be used to filter out from the image those that not related to the lung tissue substance. To do this, we filtered out all the values that are less than or equal to -600. After that, we found the biggest connected regions [21] in the 3D lung image. Normally, the selected region represents the lungs; thus, we can remove anything outside of this region to keep only tissues that are related to the lungs' image. The last preprocessing steps are to resize the image to the appropriate size and normalize the image pixel values. The new image size is described individually for each model in the following sections because different models used various image sizes to fit memory requirements. The normalization implemented by using the following equation:

$$new\_pixel = \frac{pixel - image\_min\_bound}{image\_max\_bound - image\_min\_bound} \qquad (5)$$

where pixel is the current pixel value; $image\_upper\_bound$ is equal to 400, since the things that are higher than 400 are simply bones with different radiodensity; $image\_min\_bound$ is equal to -1000 since it is the value of air radiodensity and everything less than that is not interesting for us.

4.2. DenseNet and multi-instance learning (MIL) task with sparse label assignment

This model corresponds to the approach explained in section 2.1. As a 2D convolution neural network, we chose DenseNet [3] since this network substantially reduces the number of parameters and shows state-of-the-art results on the ImageNet dataset. To train the model, we reused the pre-trained DenseNet on the ImageNet with a depth of 121 to initialize the neural network weights at the beginning of the training process. Each CT scan slice was resized to 200×200. The number of slices was not changed for this model even though the number of slices is different for each CT scan. During the training, the output of DenseNet for each CT scan slice was accumulated and passed to the loss function calculated using equation (4). We used the sparsity factor that was equal to 0.0001. When loss function was calculated for all slices in the CT scan, the backpropagation algorithm was applied to calculate gradients. In order to optimize loss function, Adam optimization algorithm was used with manually adjusted learning rate during the training process and an initial value 0.00005. The training process was stopped when the loss is not decreasing for several epochs.

When the training was finished, the trained model showed only a 0.62 accuracy level on the validation set. The main reason why a trained neural network shows a low accuracy level is that it loses 3D nature information, which is an important part of building understanding about lung cancer presence.

4.3. RADNET for lung cancer detection system

The implementation of this model follows the approach described in section 2.2. Mainly, it is an implementation based on the RADNET neural network described in [9] except for the fact that we used a CT scan of the human chest instead of a CT scan of the human brain (see figure 1). As input data, we used slices that we resized to the 200×200 size. Also, we created a mask that represents the place where a malignant pulmonary nodule is located on the current CT scan. To create a mask, we used information about the nodule candidate annotations for the Luna dataset. DSB 2017 does not contain information about tumor location. However, we reused malignant pulmonary nodule annotations created by [13] for DSB 2017 dataset in the

scope of the Kaggle competition. The mask contains information only about one pulmonary nodule that has a malignant mass characteristic.

As a 2D convolution neural network, a DenseNet [3] with depth 121 is used. As a recurrent neural network, we used LSTM [7]. For initial weights initialization pre-trained DenseNet on the ImageNet dataset was used. After the last concatenation layer in each dense block in DenseNet, we used one deconvolution layer, which upsamples the feature maps to the original image size. Added deconvolution layers compute the binary segmentation task, which explicitly focuses its attention on the region where the malignant mass is located. The final loss function defined as the weighted sum of the cross-entropy loss for both the classification task and segmentation task that is done using three additional deconvolutional layers. The training was done with the use of stochastic gradient descent with a manually adjusted learning rate during the training process and an initial value of 0.0005. The final trained model shows accuracy equal to 0.81 on the validation set.

4.4. C3D and 3D DenseNet

C3D [11] and 3D DenseNet [12] are the 3D convolution neural networks, which we trained in an end-to-end fashion following the approach described at the beginning of section 2.3. To use the 3D view of the patient's lungs as input for the 3D CNN, we stacked all CT scan slices together. The number of slices for each CT scan is different since various CT scanners with distinct resolutions are used to gather the patient's CT scans. However, all the CT scans are resized to the same resolution of 120×120×120 in order to have an input with static image size. The loss functions for both networks are defined as a weighted cross-entropy function so that we can mitigate the unbalanced dataset problem. Training is done using Adam optimization algorithm with a learning rate that equals to 0.0001 and 0.00004 for C3D and 3D DenseNet, respectively.

Trained C3D network shows accuracy equal to 0.63 and trained 3D DenseNet shows accuracy equal to 0.70. As the final accuracy results show, trained networks do not show good results. It happens because the features to identify lung cancer is very complex, and thus, the dataset should have a massive amount of the annotated data to learn them in an end-to-end way. The problem can be solved by dividing the learning task into several stages, as discussed in section 2.3.

4.5. Results summary

In addition to the described experiments, there are several results added from other publications. First, the results from the [13] work, which wins the Data Science Bowl competition in 2017, were added. This work trained a two-stage lung cancer detection system and achieved human-level accuracy on the DSB 2017 dataset. DeepLung [14] neural networks extend an idea of the previous work and improve the detection system accuracy by optimizing of the neural network architecture. Another work [10] introduced a four-stage approach (see figure 2) with a possibility to check malignant pulmonary nodules changes between previous and current CT scan. This work achieved super-human accuracy in comparison to the average radiologist.

All experiments in the article and results from other publications give a better and more complete understanding of the modern tendency in building the lung cancer detection system. The summary results are presented in table 1.

Table 1. Font Sizes for Papers

Models	Dataset	Accuracy
DenseNet + MIL task with sparse label assignment	Luna + DSB 2017	0.62
C3D	Luna + DSB 2017	0.65
3D DenseNet	Luna + DSB 2017	0.71
DenseNet + LSTM	Luna + DSB 2017	0.72
RADNET for lung screening	Luna + DSB 2017	0.81
DSB 2017 winner model [13]	Luna + DSB 2017	0.87
DeepLung [14]	Luna	0.92
D. Ardila et al. lung screening model [10]	Luna + NLST (DSB 2017 dataset based on this dataset)	0.95

5. Conclusion

Recent research advances in lung cancer detection systems show significant progress. At the moment this article was written, the best lung cancer detection systems can achieve a human-level and super-human accuracy on the given dataset. The best models can do it even in small datasets, which is a typical case for the medical domain. As experiment results have been shown to achieve a high accuracy level, lung cancer detection should use an approach that is aware of the 3D nature of the CT scans and can extract useful spatial information about the malignant mass in the image. The 2D CNN with RNN approach shows good results and gives a possibility to reuse a large amount of the 2D CNNs that are pre-trained on the ImageNet dataset. However, they have not got results as high as the results shown by the detection system with a multi-stage pipeline based on 3D CNNs. The reason for this is that 2D CCN with RNN approach is less flexible than systems with 3D CNNs pipelines and cannot be split for smaller tasks. This causes a requirement for a larger annotated dataset, which is a problem in the medical domain.

The state-of-the-art lung cancer detection system uses 3D CNNs to build a detection system pipeline. Such a system works with CT scans as a 3D representation of the lungs to build features that can recognize spatial information about malignant mass in the lung image for a specific pulmonary nodule. Usually, the lung cancer detection task is split into smaller tasks: segmentation and classification. However, the best model also adds additional components that help to generate features from the global or time context for the specific pulmonary nodule. The flexibility of adding new components, which can add useful information to help decide about cancer presence, make such detection systems more robust, reduce requirements for the dataset size, and lead to human-level performance.

References

[1] B. Chapaliuk and Y. Zaychenko, "Deep learning approach in computer-aided detection system for lung cancer," IEEE 1st International Conference on System Analysis and Intelligent Computing, SAIC 2018 - Proceedings, pp. 165-169, 2018. https://www.doi.org/10.1109/SAIC.2018.8516856

[2] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016. https://www.doi.org/10.1007/s00330-012-2627-7

[3]   G. Huang, Z. Liu, v. d. L. Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition   (CVPR),   pp.   2261-2269,   2017. https://www.doi.org/10.1109/CVPR.2017.243

[4]   F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1800-1807, 2017. https://www.doi.org/10.1109/CVPR.2017.195

[5]   T. G. Dietterich, R. H. Lathrop and T. Lozano-Pérez, "Solving the multiple instance problem with axis-parallel rectangles," Artificial Intelligence, vol. 89, no. 1-2, p. 31–71, 2002. https://www.doi.org/10.1016/S0004-3702(96)00034-3

[6]   W. Zhu, Q. Lou, Y. S. Vang and X. Xie, "Deep multi-instance networks with sparse label assignment for whole mammogram classification," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 89, no. 1-2, p. 31–71, 2017. https://www.doi.org/10.1007/978-3-319-66179-7_69

[7]   K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink and J. Schmidhuber, "LSTM: A Search Space Odyssey," IEEE Transactions on Neural Networks and Learning Systems,   vol.   28,   no.   10,   pp.   2222-2232,   2017. https://www.doi.org/10.1109/TNNLS.2016.2582924

[8]   R. Dey and F. M. Salemt, "Gate-variants of Gated Recurrent Unit (GRU) neural networks," 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS),   pp.   1597-1600,   2017. https://www.doi.org/10.1109/MWSCAS.2017.8053243

[9]   M. Grewal, M. M. Srivastava, K. Pulkit and V. Srikrishna, "RADNET: Radiologist Level Accuracy using Deep Learning for HEMORRHAGE detection in CT Scans," Arxiv, 2017. [Online]. Available: https://arxiv.org/abs/1710.04934. [Accessed 08 04 2018].

[10]   D. Ardila, A. P. Kiraly, S. Bharadwaj, B. Choi, J. J. Reicher, L. Peng, D. Tse, M. Etemadi, W. Ye, G. Corrado, D. P. Naidich and S. Shetty, "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography," Nature Medicine, 2019. https://www.nature.com/articles/s41591-019-0447-x

[11] D. Tran, L. Bourdev, R. Fergus, L. Torresani and M. Paluri, "Learning Spatiotemporal Features with 3D Convolutional Networks," Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), pp. 4489-4497, 2015. https://www.doi.org/10.1109/ICCV.2015.510

[12] T. D. Bui, J. Shin and T. Moon, "3D Densely Convolutional Networks for Volumetric Segmentation," Arxiv, 2017. [Online]. Available: https://arxiv.org/abs/1709.03199. [Accessed 08 04 2018].

[13] F. Liao, M. Liang, Z. Li, X. Hu and S. Song, "Evaluate the Malignancy of Pulmonary Nodules Using the 3D Deep Leaky Noisy-or Network," IEEE transactions on neural networks and learning systems, 2017. https://www.doi.org/10.1109/TNNLS.2019.2892409

[14] W. Zhu, C. Liu, W. Fan and X. Xie, "DeepLung: Deep 3D dual path nets for automated pulmonary nodule detection and classification," Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018, Vols. 2018-Janua, p. 673–681, 2018. https://www.doi.org/10.1109/WACV.2018.00079

[15] F. Milletari, N. Navab and S.-A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," 3D Vision (3DV), 2016 Fourth International Conference on, pp. 565-571, 2016. https://www.doi.org/10.1109/3DV.2016.79

[16] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.," IEEE transactions on pattern analysis and machine intelligence, vol. 39, no. 6, p. 1137–1149, 2017. https://www.doi.org/0.1109/TPAMI.2016.2577031

[17] K. He, G. Gkioxari, P. Dollar and R. Girshick, "Mask R-CNN," Proceedings of the IEEE International Conference on Computer Vision, Vols. 2017-October, p. 2980–2988, 2017. https://www.doi.org/10.1109/ICCV.2017.322

[18] A. Neroladaki, D. Botsikas, S. Boudabbous, C. D. Becker and X. Montet, "Computed tomography of the chest with model-based iterative reconstruction using a radiation exposure similar to chest X-ray examination: Preliminary observations," European Radiology, vol. 23, no. 2, p. 360–366, 2013. https://www.doi.org/10.1007/s00330-012-2627-7

[19] "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge," Medical Image Analysis, vol. 42, pp. 1-13, 2017. https://www.doi.org/10.1016/j.media.2017.06.015

[20] Kaggle, "Data Science Bowl 2017," [Online]. Available: https://www.kaggle.com/c/data-science-bowl-2017. [Accessed 02 04 2018].

[21] C. Fiorio and J. Gustedt, "Two linear time Union-Find strategies for image processing," Theoretical Computer Science, vol. 154, pp. 165-181, 1996. https://doi.org/10.1016/0304-3975(94)00262-2