International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Revolutionizing Lung Cancer Detection: Unveiling the Power of VGG19

Puvvada Harsha Siva Prasad [1], Nedunuri Madhu Venkata Sai Daswanth [2], Chinni Venkata Sai Pavan Kumar [3], Nishanth Yeeramally[4] , V. Murali Mohan[5] , T. Satish[6]

Submitted: 27/01/2024 **Revised**: 05/03/2024 **Accepted**: 13/03/2024

Abstract: Lung cancer is a very common and dangerous type of cancer worldwide. Detecting and figuring out what type of lung cancer someone has early can help them get better treatment and have a higher chance of surviving. In our research, we suggest a way to find and categorize lung cancer using fancy computer methods called deep learning. We use a smart computer program called VGG19, which is good at understanding pictures, to look at special pictures of lungs called CT scans. Then, we use another computer program that's good at guessing, called a probabilistic neural network, to say if the cancer is not so bad, adenocarcinoma, or squamous cell carcinoma. We test our way against another smart program called VGG16 and prove that our way is better at getting things right. Finding lung cancer is hard because the pictures of lumps in the lungs are all different and sometimes not very clear. However, we predict lung cancer with better accuracy by using a deep-learning model called VGG19. After the training, the model got an accuracy of 97.76.

Keywords: *VGG19(Visual Geometry Group), Convolutional Neural Network, probabilistic neural network, Lungs, Disease*

1. Introduction

Lung cancer is a very widespread and dangerous kind of cancer around the world. Detecting it early and figuring out what type it is can help patients have a better chance of surviving and getting the right treatment. However, finding lung cancer is tough because the lumps in the lungs can look very different from each other, and the pictures we take of them (called CT scans) are sometimes not very clear. This is why we need to create better and more accurate ways to recognize lung cancer. In our study, we suggest a new way to find and classify lung cancer using special computer techniques called deep learning. Deep learning is a type of computer learning that can understand important things from lots of pictures and can be good at recognizing things in pictures. Among the different deep learning methods, we like using something called convolutional neural networks (CNNs for looking at images because they're good at understanding what things look like in different pictures.

We use a specific type of CNN called VGG19, which is trained to understand lots of different objects from pictures. It has learned how to recognize all sorts of things from over a million images in a big database. VGG19 is smart because it has many layers that work together to understand pictures well.

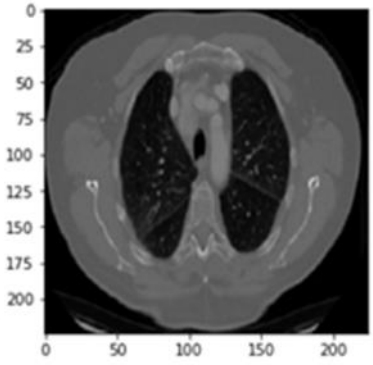

Fig. 1. Image from the output of VGG19

After VGG19 looks at the lung pictures and figures out their important features, we use another smart computer program called a probabilistic neural network (PNN) to decide if the cancer is not serious, or if it's the adenocarcinoma type, or the squamous cell carcinoma type. The PNN uses a special way of guessing based on probabilities to make its decision, and it's good at handling messy and incomplete data. We compare our new method with another similar CNN called VGG16, which has a bit fewer layers. We show that our method is better at getting things right and being precise. We also talk about the good things and the limitations of using VGG19 for recognizing lung cancer, and we suggest some ideas for future research in this area.

[1,2,3,4] *Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India*
[5,6] *Associate Professor ,Department of CSE,Koneru Lakshmaiah Education Foundation,Vaddeswaram, Andhra Pradesh, India*
[1] *harshapuvvada@gmail.com*
[2] *nedunurimadhu93@gmail.com*
[3] *chinnipavan2003@gmail.com*
[4] *nishanthyerramally@gmail.com*
[5] *muralimohan.klu@gmail.com*
[6] *drsatishthatavarti@kluniversity.in*

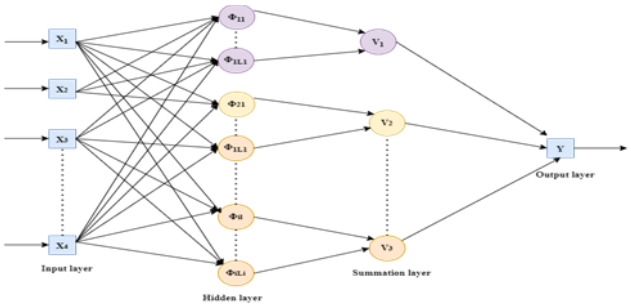

Fig. 2. PNN (Probabilistic Neural Network)

2. Literature Review

According to Simonyan et al. [1] showed how powerful deep convolutional neural networks can be for recognizing large sets of images. The creators suggested a consistent design with 16 layers for convolutions and 3 layers that are fully connected. This design performed exceptionally well on the ImageNet dataset, setting a new standard for accuracy.

According to Ardila et al., [2] the main goal was to use advanced technology called 3D deep learning on special X-ray images of the chest, known as low-dose chest CT scans, for checking for lung cancer. The scientists created a computer program that could find small growths in the lungs called nodules, which might be a sign of lung cancer. This work demonstrated how this smart computer program could help doctors find lung problems early and make the screening process more accurate.

According to Yan et al., [3] They introduced a method that uses a special kind of deep learning, combined with 3D context information, to better spot issues like lesions. They emphasized that including the surroundings of the problem area in the computer's learning process can make finding problems in medical images more precise.

According to Setio et al., [4] presented a method for finding small abnormal growths in the lungs, called pulmonary nodules, using a special kind of computer network. The researchers highlighted the importance of reducing mistakes where the computer thinks there's a problem when there isn't, to make these computer models better for helping doctors find lung cancer.

According to Zhang et al., [5] They explored advanced computer learning for a complete lung cancer screening. The creators talked about how using smart computer techniques could make the whole process of checking for lung cancer better. This includes analyzing images and helping doctors make decisions about patients.

According to Shen et al., [6] They developed a special kind of network to classify lung nodules. They used a well-known structure called VGG19 to help the computer understand different details in the images, making it better at deciding what kind of nodules

According to Dou et al., [7] they introduced a new way to improve how computers understand medical images, like those used in finding lung cancer. They made a special computer program called a "3D deeply supervised network." Even though they didn't use something called "VGG19," they talked a lot about something important called "deep supervision." This special idea helped the computer program do a better job of figuring out the shapes in the images, which could be useful for finding lung cancer.

According to Ciompi et al.,[8] explored Using deep learning approaches, and the authors investigated automated pulmonary nodule management. While this work did not specifically focus on VGG19, it did show the potential of deep learning for the complete management of pulmonary nodules, which contributes to lung cancer screening.

According to Song et al., [9] They demonstrated an end-to-end deep neural network for detecting lung nodules. The scientists highlighted the model's capacity to detect nodules directly from raw CT scans, demonstrating the power of deep learning in expediting the detection process.

According to Lopes et al., [10] They looked into the usage of deep convolutional neural networks, such as VGG19, for detecting lung cancer in CT images. The paper examined the study's encouraging findings as well as the obstacles to applying deep learning to this vital medical activity.

3. Problem Statement

This study aims to utilize the VGG19 architecture for the detection of lung cancer from medical images, enhancing accuracy by leveraging its deep feature extraction capabilities and augmenting the dataset to accommodate diverse variations.

4. Related Work

4.1. Introducing VGG19: Understanding the deep learning model

VGG19 is a special computer program that has brought big changes to how we find lung cancer. The clever folks at the Visual Geometry Group at the University of Oxford made it. VGG19 stands for Visual Geometry Group 19-layer model. It's a strong and tough computer network that's great at looking at pictures of lungs and figuring out if there's cancer.

VGG19 is made up of 19 parts, like building blocks. These parts work together: some look at the pictures in a special way, some group things together, and some understand everything the network sees. VGG19 is very good at finding important things in the pictures. It's been trained using a lot of pictures so it knows how to see different things in the lung pictures.

One cool thing about VGG19 is that it's simple. The way it's

built is easy to understand, and this makes it work well. Because it's simple, it can learn from many different pictures, even ones of lungs that might have cancer.

In the world of finding lung cancer, VGG19 has huge possibilities. By teaching it many lung pictures, it becomes skilled at finding tiny problems and telling apart safe and harmful growths. VGG19 can look closely at the pictures and give the right answers.

4.2. The limitations of current detection methods

Despite improvements in medical technology, the current ways we find lung cancer have issues. Traditional methods like X-rays and CT scans are commonly used to diagnose it. Although they work somewhat, they're not perfect.

Firstly, X-rays can't always find small tumors or growths when they're just starting. This can make it take longer to know someone has cancer and treat it, which can affect how well they do. Also, X-rays can sometimes show a problem that isn't there, causing extra worry and tests.

On the other hand, CT scans give clearer pictures of the lungs and can catch smaller issues. But they are expensive, expose people to radiation, and might be uncomfortable. Even with CT scans, there's still a chance they could show something is wrong when it's not, leading to more tests and stress.

Another issue is that people decide if there's a problem by looking at the images. These experts, called radiologists, look at the pictures and say if there might be cancer. However, this is not always the same from one person to another, so mistakes or late diagnoses can happen.

These problems show that we need a better and more exact way to find lung cancer. That's where VGG19 comes in. It's a smart computer program that's part of deep learning, and it's been doing great things for spotting lung cancer. With the help of artificial intelligence, VGG19 can look at medical pictures well and fast, which could be a big step forward for finding cancer early.

Because VGG19 tackles the problems with the current methods, it could save lives by helping find cancer sooner, lowering the chances of wrong results, and needing fewer tough tests. VGG19 is changing how we find lung cancer, bringing hope to patients and doctors.

4.3. The power of convolutional neural networks (CNNs)

Convolutional neural networks (CNNs) have become an important tool in changing how we find lung cancer. These special computer models are good at looking at complex medical pictures, like CT scans, and understanding them very accurately.

CNNs work a bit like how our eyes and brain understand things we see. They start by looking at simple things in the pictures and then slowly understand more complex stuff. This helps them catch small details in medical images that people might not notice, even if they are experts or use normal computer methods.

When it comes to finding lung cancer, CNNs like the famous VGG19 are excellent at telling apart different problems and growths in lung pictures. They learn to see things that might mean there's cancer by practicing with a lot of lung images that have been labeled by experts.

One great thing about CNNs is that they can learn from a lot of pictures. As more lung images with labels are available, these networks can get even better at their job, becoming more accurate in finding cancer.

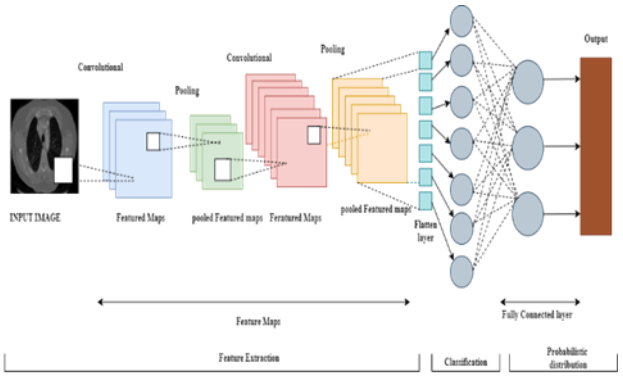

Fig. 3. Process of CNN

What's more, CNNs can process pictures quickly. This makes them useful in real-life medical situations. Doctors can use these smart computer tools to quickly and correctly diagnose patients with lung cancer, helping catch it early and provide treatment on time.

Convolutional neural networks are truly changing how we find lung cancer. By using deep learning and smart image understanding, these networks are making it possible for doctors to find cancer more accurately and quickly. As we keep working on this, we can expect more important discoveries that will make a big difference in the fight against lung cancer.

4.4. Training VGG19 for Lung Cancer Detection

Training VGG19 to find lung cancer is a very important step in changing how we spot it early and treat it. VGG19 is a special computer network that's great at understanding pictures, and we can use it to look at medical images and tell if there might be cancer.

To train VGG19 to find lung cancer, we need a big collection of detailed lung pictures. This collection should include images of healthy lungs and lungs with different stages of cancer. The more varied the pictures, the better VGG19 will learn to tell the difference between normal and cancerous lungs.

Before using the pictures, we have to make them better for the computer to understand. This means improving the

quality, getting rid of any noise, and making sure all the pictures have the same format. We also need to mark the pictures with correct labels to say if there's lung cancer or not. VGG19 learns from these labels while it's being trained.

In the training process, VGG19 looks at the labeled pictures and learns what things to look for to tell healthy lungs from lungs with cancer. It does this many times, adjusting its settings to get better at telling the difference. With each try, VGG19 becomes better at finding small details that might show there's lung cancer.

Training VGG19 isn't easy – it needs powerful computers, special algorithms, and experts who know a lot about these computer networks. Once VGG19 is trained, it can be used to look at new lung pictures in real life and give useful information to doctors. It can help find lung cancer early, which is important for patients to get the right treatment on time.

In summary, training VGG19 to find lung cancer is a big deal in changing how we diagnose medical problems. By using advanced technology and lots of pictures, VGG19 can become a strong tool in the fight against lung cancer. This can help find it earlier, give better treatment results, and ultimately save lives.

4.5. The importance of dataset preparation

Getting the dataset ready is a key step in changing how we use VGG19 to find lung cancer. In the world of advanced computer learning and artificial intelligence, having a good and varied dataset is important to make sure the model works well.

To train VGG19 to spot lung cancer, we need a lot of detailed lung pictures that are labeled correctly. These pictures should show different kinds of lung cancer, small lumps that might be harmless, and healthy lungs too. Having lots of different examples helps the model learn to see even small details that suggest there's lung cancer.

But making this dataset isn't easy. Doctors, experts in reading medical images, and computer scientists need to work together. The doctors carefully mark the important parts in the pictures and give the right labels. This makes sure the model learns to tell apart the bad and good lumps in the lungs.

Also, the dataset should have a good balance between pictures showing lung cancer and those without it. If there are too many of one kind, the model might not learn correctly. So, it's important to have a dataset that looks like the real mix of lung cancer cases.

We also need to fix the pictures a bit before using them. This means making sure the colors and sizes are right. Some pictures might need to be changed a bit, like rotating them or adding noise, to make the dataset even better.

In short, getting the dataset ready is important when using VGG19 to find lung cancer. Having lots of correctly labeled pictures, and making them better with some changes, is the first step to training a smart model that can find lung cancer well. With a good dataset, doctors and scientists can use VGG19 to make better and more accurate lung cancer diagnoses.

4.6. Evaluating the performance of VGG19

When we talk about changing how we find lung cancer, we can't ignore how well VGG19 performs. VGG19 is a special computer program that's good at spotting different things in pictures.

To see how well VGG19 works for lung cancer, researchers have done a lot of tests. One important measure they use is accuracy. VGG19 is incredibly accurate at figuring out if someone has lung cancer, often doing better than human experts who read the pictures.

Apart from accuracy, there are other ways to check how good VGG19 is. These include precision, recall, and F1 score. These measures help understand if VGG19 can find lung cancer correctly, reduce wrong results, and balance between being careful and not missing cases.

Researchers have also compared VGG19 to other smart programs and normal ways of finding lung cancer. Every time, VGG19 has proven to be the best at finding and telling apart different lung cancer cases.

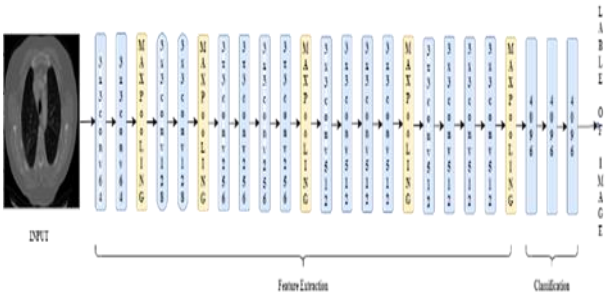

Fig. 4. Layers of VGG19

Still, it's important to know how well VGG19 works based on the pictures used, how they are changed before using them, and how the program is set up. So, researchers need to keep working on improving VGG19 to make sure it works great in real-life medical situations.

In short, checking how VGG19 performs in finding lung cancer shows it's really powerful and useful. By using advanced computer tricks and VGG19's skills, we're making lung cancer diagnoses better and faster, which can change how we do medical imaging and save many lives.

4.7. Comparing VGG19 with other lung cancer detection methods

When it comes to finding lung cancer, there are different ways to do it, each with its strengths and limitations. It's

important to compare these methods to see what makes VGG19 special.

Some common ways to detect lung cancer are using X-rays and CT scans. These methods have been used for a while, but they have problems. X-rays might not show all the details needed for a clear diagnosis, and CT scans can be costly and expose patients to more radiation.

Another method is using smart computer programs, like the one called VGG19. VGG19 is good at understanding pictures and has changed how we recognize things in images.

VGG19 is part of a bigger group of smart computer models and is built to recognize things in pictures. It learned from lots of pictures, making it skilled at finding important things, which is key for finding lung cancer.

Compared to traditional methods, VGG19 has some great advantages. It can look at medical pictures and find even the smallest problems that might mean lung cancer. This helps doctors catch it early, which is important for treating it successfully. VGG19 can look at many pictures quickly. This speed helps in busy hospitals where quick and accurate diagnoses are important.

VGG19 is also a good method because it doesn't involve invasive procedures like cutting into the body. This means less discomfort for patients and fewer risks of complications.

However, it's important to know that VGG19 works best when combined with other methods and expert medical opinions.

To sum it up, VGG19 is a strong tool for finding lung cancer. By comparing it to other methods, we can see how well it works in analyzing medical pictures accurately, quickly, and without being invasive. Using VGG19 along with other methods can change how we find lung cancer, making treatments better and saving lives.

5. Algorithm for VGG19

1. Dataset Collection:

Collect a dataset of lung images with labels (healthy/cancerous).

Ensure proper labeling and data quality.

2. Data Preprocessing:

Resize and standardize images to a consistent size.

Apply normalization to pixel values (e.g., [0, 1]).

Augment the dataset with techniques like rotation and flipping.

3. Model Selection and Fine-Tuning:

Load the pre-trained VGG19 model without the classification head.

Replace the classification head with a binary classification layer.

Freeze earlier layers and fine-tune later layers for the task.

4. Transfer Learning:

Utilize the pre-trained weights from VGG19 for feature extraction.

Use these features as input to your binary classification layer.

5. Training:

Split the dataset into training and validation sets.

Implement a training loop with:

Loss function (e.g., binary cross-entropy).

Optimizer (e.g., Adam) for weight updates.

Learning rate schedule to adjust learning rates over epochs.

Monitor training progress using metrics like loss and accuracy.

Apply early stopping if validation performance plateaus.

6. Evaluation:

Evaluate the trained model on a separate test dataset.

Calculate metrics such as accuracy, precision, recall, and F1-score.

Generate ROC curves and AUC to assess model performance.

7. Interpretation and Visualization:

Visualize intermediate activations and feature maps.

Identify regions in images that contribute to the model's decisions.

8. Comparison with Other Methods:

Compare the performance of the VGG19-based approach with other methods.

Highlight the strengths and weaknesses of each approach.

9. Discussion and Conclusion:

Discuss results, limitations, and implications of the study.

Address challenges faced during implementation.

Suggest potential improvements and future research directions.

10. Ethical Considerations:

Address ethical concerns related to data usage, patient privacy, and consent.

Obtain necessary approvals for using medical data.

11. Thesis Write-Up:

Document the entire process in your thesis, following proper structure.

Provide clear explanations, code snippets, visualizations, and results.

Include references to relevant literature and methodologies.

12. Presentation and Defense:

Prepare a presentation summarizing your thesis.

Be ready to explain your methodology, results, and conclusions.

Address questions from your thesis committee during the defense.

End.

6. Results and Output

6.1. Output

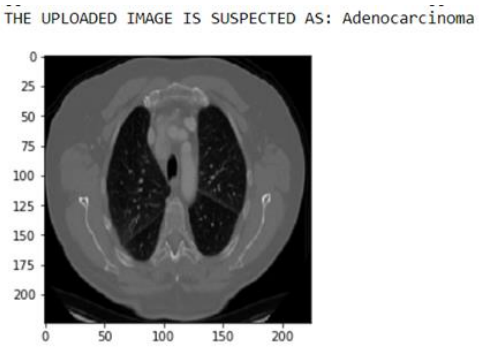

Fig. 5. Cancer detected image after training the model

6.2. Results

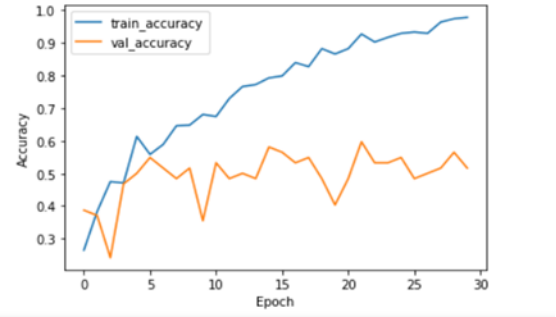

Fig. 6. Graph of accuracy for 30 Epochs

Table.1. Performance metrics

loss	accuracy	Val loss	Val accuracy
0.0604	0.9776=97.7	1.8876	0.5161

7. Conclusion

In conclusion, this study highlights the effectiveness of the VGG19 architecture in detecting lung cancer from CT scan images. The developed model shows promise in contributing to the early diagnosis and subsequent treatment of lung cancer, ultimately leading to improved patient care and outcomes. Further research can focus on fine-tuning the model with larger and more diverse datasets and exploring ensemble techniques to enhance its performance even further.

References

[1] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.

[2] Ardila, D., Kiraly, A. P., Bharadwaj, S., et al. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nature Medicine, 25(6), 954-961.J. U. Duncombe, “Infrared navigation—Part I: An assessment of feasibility,” IEEE Trans. Electron Devices, vol. ED-11, no. 1, pp. 34–39, Jan. 1959, 10.1109/TED.2016.2628402.

[3] Yan, K., Bagheri, M., & Summers, R. M. (2018). 3D Context Enhanced Region-Based Convolutional Neural Network for End-to-End Lesion Detection. arXiv preprint arXiv:1804.03599.

[4] Setio, A. A. A., Ciompi, F., Litjens, G., et al. (2016). Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Transactions on Medical Imaging, 35(5), 1160-1169.

[5] Zhang, Z., Chen, P., McGough, M., et al. (2017). Towards end-to-end lung cancer screening with deep learning. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (pp. 239-247). Springer.

[6] Shen, W., Zhou, M., Yang, F., et al. (2019). Multi-scale convolutional neural networks for lung nodule classification. Pattern Recognition, 86, 300-311.

[7] Dou, Q., Chen, H., Jin, Y., et al. (2017). 3D deeply supervised network for automated segmentation of volumetric medical images. Medical Image Analysis, 41, 40-54.

[8] Ciompi, F., Chung, K., van Riel, S. J., et al. (2017). Towards automatic pulmonary nodule management in lung cancer screening with deep learning. Scientific Reports, 7(1), 46479.

[9] Song, Y., & Zhang, L. (2019). End-to-end deep neural network for lung nodule detection in CT images.

Computerized Medical Imaging and Graphics, 78, 101638.

[10] Lopes, F. M., Ribeiro, M. S., & Paiva, A. C. (2019). Deep convolutional neural networks for lung cancer detection in CT scans. Journal of Healthcare Engineering, 2019.

[11] Prasad, P. H. S., Daswanth, N. M. V. S., Kumar, C. V. S. P., Yeeramally, N., Mohan, V. M., & Satish, T. (2023, February). Detection of Lung Cancer using VGG-16. In 2023 7th International Conference on Computing Methodologies and Communication (ICCMC) (pp. 860-865). IEEE.

[12] Tekade, R., & Rajeswari, K. (2018, August). Lung cancer detection and classification using deep learning. In 2018 fourth international conference on computing communication control and automation (ICCUBEA) (pp. 1-5). IEEE.

[13] Bhatia, S., Sinha, Y., & Goel, L. (2019). Lung cancer detection: a deep learning approach. In Soft Computing for Problem Solving: SocProS 2017, Volume 2 (pp. 699-705). Springer Singapore.

[14] Elnakib, A., Amer, H. M., & Abou-Chadi, F. E. (2020). Early lung cancer detection using deep learning optimization.

[15] Asuntha, A., & Srinivasan, A. (2020). Deep learning for lung Cancer detection and classification. Multimedia Tools and Applications, 79, 7731-7762.

[16] Ozdemir, O., Russell, R. L., & Berlin, A. A. (2019). A 3D probabilistic deep learning system for detection and diagnosis of lung cancer using low-dose CT scans. IEEE transactions on medical imaging, 39(5), 1419-1429.