**Lung Cancer Detection Using Deep Learning and Explainable Methods**

Ayah Alomar\*, Moayed Alazzam\*, Hala Mustafa\*, Ahmad Mustafa\*

\*Faculty of Computer and information technology, Jordan university of science and technology, Irbid, Jordan

{[afalomar20, maazzam20, htmustafa20} @cit.just.edu.jo](mailto:afalomar20,%20maazzam20,%20htmustafa20%7d%20@cit.just.edu.jo), ammustafa@just.edu.jo

***Abstract-* Lung cancer is one of the most prevalent deadly diseases and it can extend to the rest of the human body, one of the important ways to treat is through CT scan images by building a deep learning model that detects cancer and explain the models by XAI and radiologists to make the results trusted for medical field. The deep learning models inceptionV3 and ResNet50 were used to classify CT scans of the lungs for the presence of cancer. The models were trained on a Kaggle dataset that pre-processed and augmented as physicians determine the best features of images, it was able to accurately detect lung cancer in new patients. Additionally, an XAI model was used to explain the decision-making process of the deep learning model, providing insights into which features of the CT scan were most important for the model's diagnosis.**

**The ResNet50 achieved the highest performance with 100% accuracy for testing and training images, the InceptionV3 got an accuracy of 99.92% for training and testing datasets, and the LIME and GRAD-CAM explained the model's performance by highlighting the most important features for each model, on the other hand, the radiologists provide this research with a worth insight to explain the deep learning and XAI models by determine the cancer cells in each image and they give some explains about the model's misclassification, the combination of deep learning, XAI models, and the radiologist's diagnosis showed promise for improving the accuracy and interpretability of lung cancer diagnosis using CT imaging.**

***Keywords-Deep learning, Artificial Intelligence, Explainability, lung cancer, CT scan, radiologist's diagnosis*.**

1. **Introduction**

One type of cancer that develops in the lungs is lung cancer [1]it is more deadly than breast, colon, and prostate cancers combined, making it the primary cause of cancer deaths globally. Non-small cell lung cancer (NSCLC) and small cell lung cancer are the two main subtypes of lung cancer (SCLC). The risk factors for lung cancer include smoking, exposure to certain chemicals and pollutants, and a family history of the disease. NSCLC is the more prevalent variety and tends to grow and spread more slowly than SCLC. Avoiding tobacco use and limiting exposure to other recognized risk factors are the best ways to prevent lung cancer [2].

A CT (computed tomography) scan is a type of imaging testing that creates finely detailed images of the inside of the body using specialist X-ray equipment [3]. It can be used to find lung cancer and establish the tumor's stage (how advanced it is), Since a CT scan of the chest can give precise images of the lungs and adjacent tissues, it is frequently used to diagnose lung cancer.

Overall, the use of deep learning with CT scans has the potential to significantly improve the accuracy and efficiency of lung cancer detection, by automating the analysis of medical images and identifying patterns that may be missed by human interpretation [4].

Explainable artificial intelligence, or XAI, is a branch of AI that focuses on creating algorithms and systems that can transparently communicate their decision-making processes and how they came to a certain result [5]. In order to increase the accuracy of the results and detect any biases or flaws in the algorithm's decision-making process, this can be especially helpful for medical image analysis [6].

1. **Related work**

Due to the severity and effects it has on people's lives, a number of study articles have recently been published that address the problem of lung cancer [7]. Therefore, publications are examined in order to accomplish this study goal because the concepts for these studies based on image analysis come from a variety of sources, including transfer learning and deep learning [8].

In research [9], author's aim is to detect lung cancer by analyzing pathology images during the widespread of using them in treatment as a routine procedure, and they sought to give a general review of the current and future uses of AI techniques in lung cancer, they began by outlining the potential and problems that exist today for lung cancer pathology image, then went over recent advances in deep learning that may have an effect on lung cancer digital pathology, authors concluded by summarizing the current uses of deep learning models in the diagnosis and prognosis of lung cancer.

Regarding to the methods and material, they applied deep learning models especially the supervised learning CNNs were used because it proved their importance in the pathology images classification field, lung cancer, head cancer, and others diseases. the deep learning models have main two characteristics, it facilitates feature extraction from images because the model does it automatically and it has many layers and kernel which is mean it can deal with any complex function, for training and testing the dataset they used the traditional way a transfer learning method that derived from a similar problem, they applied the deep learning model on a dataset image with size 300X300 pixels, Their future work is to apply the interpretation machine learning models to interpret the black box model such as deep learning and transfer learning and to gain meaningful features.

Another research [10], discuss the computer-aided diagnosis can be extremely important in the early diagnosis of lung cancer, which is a fundamental step for the treatment of lung cancer. The majority of CAD approaches that have been described classify each lung nodule separately to diagnose lung cancer. however, does not correspond with actual practice, in which doctors identify patients by comparing a collection of nodule photos rather than focusing on a single one at a time. A significant obstacle to their acceptance is the limited interpretability of the results these approaches provide, the attention mechanism offers improved interpretability, the results indicate that the approach can produce accuracy of 0.807.

After discussing this research’s and looking at their most prominent strengths and weaknesses, this study's idea came to reduce the prevalence of lung cancer by using deep learning techniques, as noticed in the papers that use the augmentation technique, the researchers randomly focused on features on the images without taking into consideration that the images serve the medical sector and must focus on specific areas of the lung affected by cancer to get the highest results correctly because the first goal is to preserve human life and not only to obtain high accuracy, because of that, in this research the radiologists defined the important features from lungs images to augment it and to explain the results to be trusted.

1. **Methodology**

**3.1. Overview**

Using deep learning and XAI to detect lung cancer typically involves a combination of data collection, preprocessing, model development, evaluation, and interpretation. The first step is collecting a dataset of chest CT scans from Kaggle with and without lung cancer [12], the data is then preprocessed to make it suitable for deep learning, which includes resizing the images, normalizing, and applying data augmentation techniques [13]. The next step is to develop and train deep learning models ResNet50 and InceptionV3 on the preprocessed data to classify whether a given CT scan is from a patient with lung cancer or not. The model is then evaluated using various metrics such as accuracy, precision, recall, and F1 score, and XAI techniques are used to interpret the model's decision-making process and identify any potential biases or errors. Finally, the model is given to radiologists to determine the detection of lung cancer is accurate or not.

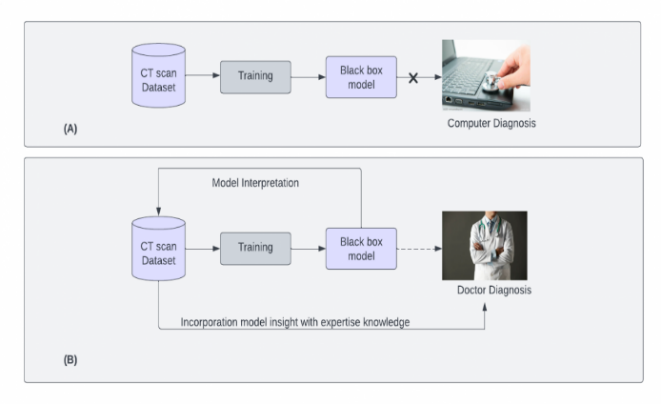


Figure 1.The difference between the usual workflow of deep learning and our approach

**3.2. Dataset preprocessing and splitting**

Preprocessing of images in a lung cancer dataset is a crucial step in making the data suitable for deep learning [11]. The first step is to resize the images to a consistent size, this is to ensure that the model's input is of the same size and shape, the second step is to convert all images to the same file extension. The consistent images format ensures that the images can be loaded and processed by the deep learning model without any compatibility issues [14]. The third step is data augmentation techniques that is applied to the dataset, this is to artificially increase the size of the dataset by applying various transformations to the images such as rotation, and zooming. These techniques can help the model to generalize better and improve it is performance, the dataset increased from 1000 images into 1688 images as the table [1] shown.

*Table 1. Dataset splitting*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Training set | Testing set | Validation set |
| Lung cancer | 591 | 84 | 169 |
| Normal | 591 | 84 | 169 |
| **Total** | **1182** | **168** | **338** |

**3.3. Transfer Learning**

ResNet50[15] and InceptionV3[16] are both convolutional neural networks (CNNs) that have been trained on the ImageNet dataset for image classification tasks. ResNet50 is a 50-layer deep network that uses residual connections to improve the flow of information through the network and reduce the vanishing gradient problem. InceptionV3, on the other hand, uses a modular architecture called Inception, which allows for the network to have a wider receptive field and capture more information from the input image [17]

**3.4. Explainable Machine Learning**

The predictions of a deep learning model on an image dataset of lung cancer CT scans were understood and interpreted using two well-known explainable machine learning approaches, Grad-CAM [19] and LIME [18]. Grad-CAM (Gradient-weighted Class Activation Mapping) is a method for producing heat maps that highlight the areas of an image that are most crucial for the model's prediction, the heat map that was produced reveals which parts of the image were most important for the model's prediction and can be used to determine which attributes the model is employing, this helps on spot any possible biases or inaccuracies in the model's predictions.

Another approach called LIME (Local Interpretable Model-agnostic Explanations) produces explanations for a model's predictions by approximating the model locally around the input. A simpler model is trained on the affected images by sampling the input space surrounding the input image. The explanation for the prediction made by the original model is then produced using the simpler model. LIME can offer a more thorough justification of how the model generates its prediction.

**3.5. Experiments and evaluation metrics**

There are many hyper parameters that modified to improve the performance of the model when experimenting in transfer learning for images datasets [20], the learning rate decay is a crucial hyper parameter that used to gradually lower the learning rate, preventing the model from overfitting by trying two values .001 or. 0001, the optimizer Adam is also one of the important hyper parameters and a regularization method called dropout removes certain neurons by random during training to avoid overfitting used with values .5 or .2. The amount of training samples used in each iteration of the training process is controlled by the batch size, the final factor controlling how many times the model is trained on the full dataset is the number of epochs and 100 epochs is applied with using the early stopping technique [23]. In order to evaluate the performance of the model, metrics such as highest validation accuracy or lowest validation loss as well as common evaluation metrics for image classification include accuracy, precision, recall, and F1-score is used [21].

In other words, by altering the parameters that are particular to the task, dataset, and computational resources [22], applying hyper parameters in transfer learning on image datasets allow fine-tune the model and make the most of the pre-trained model. As a result, the model may perform better and generalize more well to new data by finding the best combination of hyper parameters that work well for the specific dataset and task [24].

1. **Results**
   1. **Transfer Learning**

The accuracy of deep learning model is significantly impacted by the choice of hyper parameters. It is essential to compare the results in order to choose the best model and to select one model that has the best overall performance and is most suitable for the task of lung cancer detection. Factors like the number of layers in the network, the number of neurons in each layer, and the learning rate can all affect the model's performance as shown in the tables [3].

*Table 2. Models Best Results*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Training Accuracy** | **Training Loss** | **Validation Accuracy** | **Validation Loss** | **Testing Accuracy** | **Testing Loss** |
| **Inception0.001\_ND\_0.5\_64** | **99.92** | **.01** | **99.1** | **0.03** | **99.2** | **.03** |
| **ResNet50\_0.0001\_D\_0.2\_128** | **100** | **0.0** | **100** | **0.0** | **100** | **0.0** |

* 1. **Explainable Machine Learning**

LIME and Grad-CAM combined to provide a more explanation of how the model is classifying images as either cancerous or normal in a ResNet50 and InceptionV3 model for CT scan image analysis of lung cancer. In order to determine which features of the image are most crucial for making the diagnosis, LIME creates heat maps that highlight the areas of the image that have the most influence on the model's prediction. On the other hand, the Grad-CAM visualization technique creates a heat map that highlights the crucial areas of the input image that are used to make a prediction by emphasizing the areas of the image that the deep learning model is concentrating on, it is used to understand how the algorithm makes decisions. LIME and Grad-CAM can give more understanding of how the model generates its predictions, combining the data from the two ways makes it feasible to pinpoint not just the image's key elements but also the precise areas that the model is concentrating on. This can offer important insights into the model's underlying mechanics, support the creation of improved models for classifying CT scan images for lung cancer, and improve the comprehension of the diagnostic procedure.

* 1. **Radiologists Diagnosis**

A CT scan image for lung cancer is classified using a deep-learning model called ResNet50. The model is trained to identify characteristics of lung cancer in the CT scan image and classify it as either cancer or normal. by using ResNet50, the classification results are accurate without any misclassification, but in this part, it is important to know if the model trained the most important features or not by taking randomly images and applying the LIME and Grad-cam to compare their explanation with radiologists explain, in ‘Fig.8’ the original images with square mark determined by radiologists to focus on the cancer cells on CT images, in LIME image the model focus on different area that does not related to cancer cells.

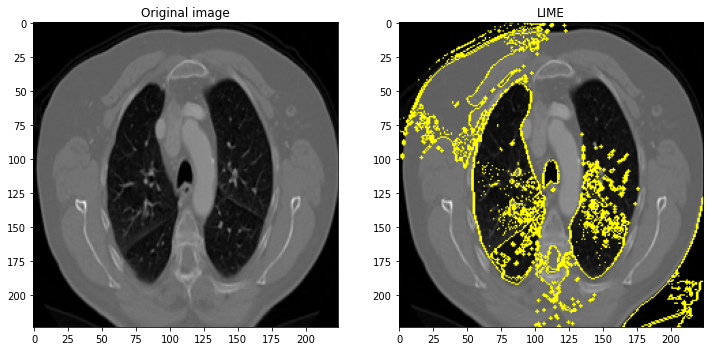
****

Figure2.The different between radiologist and LIME diagnosis

The models ResNet50 and InceptionV3 were give a high accuracy in classifying lung cancer images but were found focused on the wrong features in some images, it is mean that the model is not properly identifying the characteristics of cancer. This leads to false positives or false negatives, which can have serious consequences in a clinical setting.

In this scenario, it is important to re-evaluate the model's performance and identify the reasons for why it is focused on the wrong features. Additionally, radiologists play a crucial role in identifying the problem and providing feedback on the model's performance and which features are most important for accurate diagnosis. By addressing the problem of focusing on the wrong features, such as the bad image resolution can be a contributing factor to a deep learning model focusing on the wrong features when classifying lung cancer images. When an image has low resolution, it can be more difficult for the model to distinguish important features such as tumors or nodules, which can lead to the model focusing on other, less relevant features instead. also, low resolution images can result in a loss of important details and textures, which can further complicate the model's ability to accurately identify and classify cancerous regions.

1. **Discussion**

Using deep learning models such as ResNet50 and InceptionV3 to classify lung cancer images, along with interpretation methods LIME and Grad-CAM, helps to identify important features in the images that are indicative of cancer. However, it is important to evaluate the model's performance by radiologists to ensure that it is focusing on the correct features and not being influenced by irrelevant or confounding factors. By involving radiologists in the evaluation process, they provide valuable insights and feedback on the model's performance and help to identify any areas where it may be lacking. Additionally, interpretation methods such as LIME and Grad-CAM can be useful in understanding the model's decision-making process and identifying which features are most important in the classification of lung cancer images. This can help to improve the model's overall performance and ensure that it is focused on the correct features for accurate diagnosis.

It is worth to mention that the radiologists support this research by many advices that helps not to fall into the error of other research by using a dataset with just two classes, cancer and normal, because there is many research in lung cancer field classify the cancer types by CT scan that also known as a computed tomography scan, is a diagnostic imaging test that uses X-rays to create detailed images of the body.

1. **Conclusion and future works**

Deep learning models ResNet50 and Inceptionv3 are employed to classify lung CT images. These models are used to find patterns and features that are suggestive of various lung disorders after being trained on a dataset of CT scan images with satisfied results for ResNet50 achieved 100% testing accuracy and the InceptionV3 achieved 99.2% testing accuracy. The outcomes of these deep learning models are also interpreted using XAI models like Grad-CAM and LIME, which shed light on the model's prediction process. Grad-CAM generates a heat map that indicates the areas of the CT images that the model is concentrating on while generating its predictions, while LIME generates explanations for the model's predictions by determining the most crucial aspects of the image.

Together, these techniques can be used to improve the understanding and interpretability of the deep learning models used to classify lung CT images, also the physicians provide feedback on whether the model's classifications are in line with their own observations and clinical experience and identifying which regions of the image the model is focusing on when making it is predictions.

Overall, the collaboration between physicians, deep learning and XAI models in the interpretation of medical images can lead to more accurate and reliable predictions, and ultimately, better patient care, for future work it is worth to use more XAI models to interpret images classification to be more trusted for medical fields and keep the radiologists evaluation in mind to train the models on more images that failed to classify it.

**References**

[1] K. L. Kohsasih and B. H. Hayadi, “Classification SARS-CoV-2 Disease based on CT-Scan Image Using Convolutional Neural Network Classification SARS-CoV-2 Disease based on CT-Scan Image Using Convolutional Neural Network,” no. November, 2022, doi: 10.15294/sji.v9i2.36583.

[2] K. Ramana *et al.*, “Early Prediction of Lung Cancers Using Deep Saliency Capsule and Pre-Trained Deep Learning Frameworks,” *Front. Oncol.*, vol. 12, no. June, pp. 1–13, 2022, doi: 10.3389/fonc.2022.886739.

[3] G. Ren *et al.*, “A Transfer Learning Framework for Deep Learning-Based CT-to-Perfusion Mapping on Lung Cancer Patients,” *Front. Oncol.*, vol. 12, no. July, pp. 1–11, 2022, doi: 10.3389/fonc.2022.883516.

[4] S. Wang *et al.*, “Artificial intelligence in lung cancer pathology image analysis,” *Cancers (Basel).*, vol. 11, no. 11, pp. 1–16, 2019, doi: 10.3390/cancers11111673.

[5] J. U. Lim *et al.*, “Association between clinical outcomes and local treatment in stage IV non-small cell lung cancer patients with single extrathoracic metastasis,” *Thorac. Cancer*, vol. 13, no. 9, pp. 1349–1360, 2022, doi: 10.1111/1759-7714.14398.

[6] W. Jiang, G. Zeng, S. Wang, X. Wu, and C. Xu, “Application of Deep Learning in Lung Cancer Imaging Diagnosis,” *J. Healthc. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/6107940.

[7] D. S. Jeon *et al.*, “Sex differences in the characteristics and survival of patients with non-small-cell lung cancer: A retrospective analytical study based on real-world clinical data of the Korean population,” *Thorac. Cancer*, vol. 13, no. 18, pp. 2584–2591, 2022, doi: 10.1111/1759-7714.14594.

[8] Q. Wang *et al.*, “Cascaded-Recalibrated Multiple Instance Deep Model for Pathologic-Level Lung Cancer Prediction in CT Images,” *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/9469234.

[9] C. Patrício, J. C. Neves, and L. F. Teixeira, “Explainable Deep Learning Methods in Medical Imaging Diagnosis: A Survey,” vol. 1, no. 1, pp. 1–36, 2022, [Online]. Available: http://arxiv.org/abs/2205.04766

[10] S. Walia, K. Kumar, S. Agarwal, and H. Kim, “Using XAI for Deep Learning-Based Image Manipulation Detection with Shapley Additive Explanation,” *Symmetry (Basel).*, vol. 14, no. 8, 2022, doi: 10.3390/sym14081611.

2022.

[11] “Chest CT-Scan images Dataset | Kaggle.” https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images (accessed Jan. 20, 2023).

[12] Z. Hussain, F. Gimenez, D. Yi, and D. Rubin, “Differential Data Augmentation Techniques for Medical Imaging Classification Tasks,” *AMIA ... Annu. Symp. proceedings. AMIA Symp.*, vol. 2017, no. April 2018, pp. 979–984, 2017.

[13] D. Ray, O. Pinti, and A. A. Oberai, “Deep Learning and Computational Physics (Lecture Notes),” 2023, [Online]. Available: http://arxiv.org/abs/2301.00942

[14] D. Kaul, H. Raju, and B. K. Tripathy, “Deep Learning in Healthcare,” *Stud. Big Data*, vol. 91, no. November, pp. 97–115, 2022, doi: 10.1007/978-3-030-75855-4\_6.

[15] R. Zhang, Y. Zhu, Z. Ge, H. Mu, D. Qi, and H. Ni, “Transfer Learning for Leaf Small Dataset Using Improved ResNet50 Network with Mixed Activation Functions,” *Forests*, vol. 13, no. 12, 2022, doi: 10.3390/f13122072.

[16] N. S. Shadin, S. Sanjana, and N. J. Lisa, “COVID-19 Diagnosis from Chest X-ray Images Using Convolutional Neural Network(CNN) and InceptionV3,” *2021 Int. Conf. Inf. Technol. ICIT 2021 - Proc.*, no. July, pp. 799–804, 2021, doi: 10.1109/ICIT52682.2021.9491752.

[17] A. A. R. Odeh, A. Alomar, and S. Aljawarneh, “Detection of COVID-19 using deep learning on x-ray lung images,” *PeerJ Comput. Sci.*, vol. 8, 2022, doi: 10.7717/PEERJ-CS.1082.

[18] J. Dieber and S. Kirrane, “Why model why? Assessing the strengths and limitations of LIME,” no. iii, 2020, [Online]. Available: http://arxiv.org/abs/2012.00093

[19] M. Lerma and M. Lucas, “Grad-CAM++ is Equivalent to Grad-CAM With Positive Gradients,” no. August, pp. 113–120, 2022, doi: 10.56541/awjv6348.

[20] J. Plested and T. Gedeon, “Deep transfer learning for image classification: a survey,” 2022, [Online]. Available: http://arxiv.org/abs/2205.09904

[21] H. Wang, T. Li, Z. Zhuang, T. Chen, H. Liang, and J. Sun, “Early Stopping for Deep Image Prior,” 2021, [Online]. Available: http://arxiv.org/abs/2112.06074

[22] G. Shao, L. Tang, and H. Zhang, “Introducing image classification efficacies,” *IEEE Access*, vol. 9, pp. 134809–134816, 2021, doi: 10.1109/ACCESS.2021.3116526.

[23] R. Mohakud and R. Dash, “Skin cancer image segmentation utilizing a novel EN-GWO based hyper-parameter optimized FCEDN,” *J. King Saud Univ. - Comput. Inf. Sci.*, no. May, 2022, doi: 10.1016/j.jksuci.2021.12.018.

[24] C. Chen *et al.*, “Deep Learning on Computational-Resource-Limited Platforms: A Survey,” *Mob. Inf. Syst.*, vol. 2020, 2020, doi: 10.1155/2020/8454327.