2024 27th International Conference on Computer and Information Technology (ICCIT),
20-22 December 2024, Cox's Bazar, Bangladesh

Custom Lightweight CNN and Data-Efficient Models for Efficient and Fast Convergence in Lung Cancer Classification

Mariam Binte Bashir
Electrical and Computer Engineering, NSU
Bashundhara R/A, Dhaka, Bangladesh
mariam.binte@northsouth.edu

Abu Mukaddim Rahi
Electrical and Computer Engineering, NSU
Bashundhara R/A, Dhaka, Bangladesh
abu.rahi@northsouth.edu

Md. Khurshid Jahan
Electrical and Computer Engineering, NSU
Bashundhara R/A, Dhaka, Bangladesh
khurshid.jahan@northsouth.edu

Abdullah Al Shafi
Electrical and Computer Engineering, NSU
Bashundhara R/A, Dhaka, Bangladesh
abdullah.shafi1@northsouth.edu

Abstract—**Considering that lung cancer is the primary cause of cancer-related death across the globe, there are significant differences in patient outcomes as well as the subsequent clinical management if early detection, prediction, and diagnosis are done correctly. Machine learning techniques, particularly some studies, have received significant attention due to their effectiveness and accuracy in predicting, diagnosing, or treating lung cancers. This research presents a new direction using the data-efficient image transformer (DeiT). The paper also evaluates a custom lightweight convolutional neural network (CNN) and several other pre-trained models.The custom lightweight CNN obtained testing accuracies of 99.71%. DeiT appeared to be very efficient as it successfully achieved 99.89% testing accuracy with almost 16 times faster convergence rate than other models, such as ResNet50, InceptionV3, VGG16 and many more.**

Index Terms—**Lung Cancer, DeiT,CNN, Machine Learning, Medical Imaging.**

I. INTRODUCTION

The most destructive cancer is lung cancer, which might occur in the lung tissues, normally in the cells lining the air passages. The primary cause of death universally is also caused by uncontrolled cell division in the lungs. Approximately 90 percent of lung cancers are caused by cigarette smoking. Other risk factors include chemicals like radon and asbestos exposure, asbestos, arsenic, chromium, and nickel, or a family history of lung cancer which can result in deadly issues [1]. According to the WHO, lung cancer remains the main cause of death globally, with around 1.80 million deaths, or 18.7% of all cancer deaths. In addition, in Bangladesh, lung cancer reached 12,174, or 1.70% of total deaths [2]. Many cancers seriously affect the lungs, among them two most common kinds are non-small cell lung cancer and, small cell lung cancer.

The most common symptoms include:

- Long-term cough that does not go away and worsens over time, mild chest pain, and shortness of breath (dyspnea).
- Loss of appetite with unexplained weight loss for no known cause, wheezing, and hoarseness.
- Swelling in the appearance of the neck, arms, or upper chest.

A lung cancer diagnosis often starts with an imaging test to examine the lungs like a computed tomography (CT) scan. Since the major factor for lung cancer is smoking, it can be prevented by quitting smoking, avoiding secondhand smoke, and steering clear of other substances that can harm the lungs [3].

In this work to identify lung cancer, we have used models that include CNN and various deep-learning approaches. We used the open-source dataset 'IQ-OTH/NCCD' for lung cancer patients.The main objective is to figure out the cancerous cells in the lung nodules by using CT-scan images. For this, lung cancer has been classified into three main types: benign, malignant, and normal, on the basis of CT-scan slices of nodules within the lung. In this work, we used several transformers namely DeiT Transformer and EfficientFormer Model. These models shows promising outcomes detecting computed tomography(CT). The DeiT Transformer is highly regarded for nodule detection, especially on small datasets. This method achieves high accuracy, comparable to larger and more complex models like ViT. Efficient former methods achieved similiar performance and highly effective to detect lung cancer. It is also used to attach a hook for extracting features of visualization of intermediate layers to improve performance and accuracy.

A. Novelty

- Developed a lightweight custom CNN architecture for lung cancer classification, achieving significant accuracy with a shorter convergence time compared to other pre-trained CNN models.
- DeiT demonstrated outstanding accuracy with faster convergence, showcasing its efficiency and effectiveness.

II. LITERATURE REVIEW

Several work done in this field. These study proposed different types of architectures and compared in different studies. Various pre-trained DL architectures were explored along with traditional machine learning models.

979-8-3315-1909-4/24/$31.00 ©2024 IEEE

Ruchita Tekade et al. developed a system to diagnose lung cancer using a deep-learning model. To improve malignancy prediction they combined a 3D multipath VGG-like network for classification. on the LUNA16 dataset, they used a hybrid deep learning model combining CNN and RNN for classifying lung nodules, achieving an accuracy of 94.2% [4].

Preeti Katiyar et al. demonstrated a comparison of methods for the identification and classification of lung cancer using CT scans. They have employed various classification techniques, SVM, ANN, CNN, and DNN. Notably, DNN achieved 97%, underscoring the effectiveness of these methods in Computer-Aided Diagnosis (CAD) systems [5].

Zeyu Ren et al. developed a combined framework, LCGAN, for the labeling of lung cancer with high performance. The LCGAN framework generates 10,000 synthetic images per class and integrates them with real datasets for a robust training set. The authors have used various deep learning models and their model VGG-DF which obtained the best results with an accuracy of 95.80% on the original dataset and 99.84% combining original and synthetic datasets [6].

M. Mohamed Musthafa et al. proposed a method of machine learning in medical imaging. particularly for lung cancer diagnosis. The authors used a public dataset named IQ-OTHNCCD. For the dataset, they have implemented the SMOTE technique and Label encoding. Among different classifiers, CNN gave the authors the best accuracy, approximately 99.64% [7].

Mamoona Humayun et al. focus on the diagnosis of lung cancer and recognition models in deep learning with an implementation of a transfer-learning approach. They include the screening of datasets from a hospital. The authors have used various deep learning models like VGG 16, VGG 19, and Xception with an accuracy of 97.4% [8].

Rehan Raza et al. proposed a novel transfer learning-based predictor known as Lung-EffNet for lung cancer classification. Using five variations of EfficientNet this predictor is evaluated, specifically B0–B4. The Lung-EffNet model achieved 99.10% accuracy on the test set [9].

Sharmila Nageswaran et al. demonstrated the precise classification and prediction of lung cancer with the utilization of machine learning and image processing technology. For improving image quality they employed geometric mean filters in image preprocessing. Machine learning techniques such as ANN, KNN, and RF are employed for classification purposes. Among these, the ANN model has been found to produce more accurate predictions for lung cancer [10].

Kadiyala Ramana et al. provided innovative capsule networks that utilize saliency-based techniques to improve segmentation. They also deploy optimized pre-trained transfer learning to enhance the accuracy of predicting lung tumors from CT scans. Furthermore, it has outperformed several well-known deep-learning models. The experimental findings show that the suggested framework has reached a peak performance with an accuracy of 98.5% [11].

Michaela Cellina et al. presented an approach that combines a modified deep transfer learning EfficientNet with a masked autoencoder for distribution estimation (MADE). This approach enables various benefits such as feature learning, dimensionality reduction, uncertainty estimation, etc. The expected outcomes indicate that the Mask-EffNet outperforms several CNNs in terms of both accuracy and effectiveness. The Mask-EffNet model achieved a test set accuracy of 98.98% [12].

III. PROPOSED SYSTEM

This project outlines the procedures and application of advanced machine learning models for the proposed lung cancer detection system. Initially, several image pre-processing steps were applied on CT scan slices. The dataset was then divided into training, validation, and test sets using the holdout validation technique. To identify the most suitable classification model, we conducted tests and comparisons across various models. Due to the challenges in acquiring large amounts of detailed data for training, particularly in the context of medical images, we decided to employ a data efficient transformer model for enhanced performance. This approach facilitated the multi-class identification of lung cancer, categorizing it into three distinct groups: benign, malignant, and normal. The following sections provide comprehensive details on the preprocessing steps and the recommended model. Figure 1 illustrates the workflow of the proposed lung cancer prediction system.

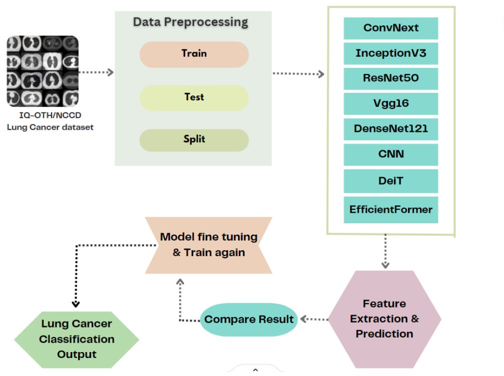

Fig. 1. The general workflow of the proposed system.

A. Dataset

We used the dataset from an augmented open-source dataset from Kaggle named IQ-OTH/NCCD for the CT scan image diagnosis. over three months This lung cancer dataset was obtained at the specialist National Center for Cancer Diseases hospitals in the fall of 2019. The dataset comprises 1,190 images representing CT scan slices from 110 cases. These are divided into three classes: normal, benign, and malignant (see “Fig. 2”). Specifically, as malignant, 40 cases are diagnosed, 15 cases as benign, and 55 cases as normal. The CT scans were initially collected in DICOM format. Data preprocessing techniques were employed to balance the dataset and enhance accuracy [13].

B. Data preprocessing

It is essential to preprocess the data before creating a model to remove unnecessary noise and outliers from the dataset that can hinder the model’s performance. We cleaned and prepared the data for building the model after gathering the necessary dataset. The code iterates through each image in the dataset, resizing it to 256x256 pixels. To ensure that the data is randomized before splitting into training and testing sets all

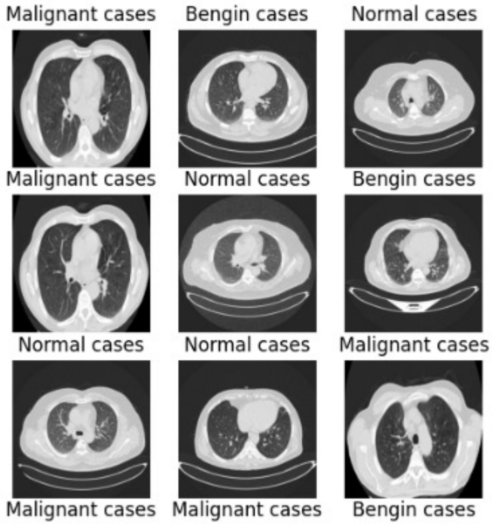

Fig. 2. Sample images in the dataset are classified into three categories.

the images are loaded and the dataset is shuffled. The image pixel values are normalized by dividing by 255. We took the pixel values to the range [0, 1] for faster convergence during training. To ensure the model has separate data for training, validating during training, and testing after training we split the folder library to split the dataset into training, validation, and test sets with an 80-10-10. To perform real-time data augmentation ImageDataGenerator is used. This helps prevent overfitting by providing slightly modified versions of images each epoch. We attach hooks to intermediate layers of the model to extract features for visualization. The array stores activation from these layers to determine whether the model is learning useful features which allows us to analyze how the model represents images at different stages. The einops library is used to rearrange tensor dimensions for visualization or further analysis.

C. Proposed Architecture

The proposed lung cancer detection systems procedures and the application of advanced machine learning models are detailed in this project. First, the CT scan slices are loaded from a Keras-compatible directory structure, followed by several image pre-processing steps. The dataset was then split into training, validation, and test sets using the holdout validation technique. The suitable classification model for this dataset was identified by testing and then comparing it with several different models. It is tough to collect large amounts of data with annotations. As it is based on medical images, it is difficult to collect a large number of annotated data to train the model. Hence we used a transformer model for better performance. Here, by employing a machine learning approach, we fine-tuned various models for the classification of lung cancer into three different categories, which are benign, malignant, and normal. In the following sections, the details of the preprocessing steps and all the proposed models are given.

1) ConvNext : It has been used for feature extraction on datasets in the learning process. Its design incorporates advanced techniques that enhance its feature extraction and representation capabilities. It helps the model generalize better and perform well after pre-training. By improving its ability to detect and classify abnormalities it helps the model to focus on relevant parts of the image. Transferring the learned features from general image data to medical imaging tasks, it enhances the performance to predict lung cancer accurately [14].

2) InceptionV3 : The InceptionV3 model is a transfer learning model that transfers the learned parametric weights of the neural network to a new task, thus it is a pre-trained model. Its application proved useful in extracting scepters at a few different scales. After passing through the various layers, it could identify both small and big anomalies in the lung scans. Its efficiency was so high that the model was able to achieve good results even on sparse medical datasets. Then, it successfully deals with the high-definition photographs by recognizing small features that say this is the cancerous lung tissue [15].

3) ResNet50 : ResNet50 is a convolutional neural network for image classification. It allows training very deep networks using residual learning. Its robust feature extraction enables it to perform well on diverse lung images. Its ability to handle high-dimensional and complex data serves effectively in predicting the capability of lung cancer. Besides, fine-tuning enhances its performance for tasks such as classifying between benign and malignant lung nodules. Due to the fact that it can manage high dimensional and complex data, the algorithm plays an effective role in lung cancer prognosis [16].

4) DenseNet12 : DenseNet is a type of CNN architecture, wherein all layers receive inputs from all other layers and allows for feature reuse while reducing the parameter count. Capturing intricate patterns and details in images plays an efficient role. It is utilized with pre-trained weights, and a custom classifier is added to tailor it for the lung cancer dataset [17].

5) Vgg16 : VGG16 is a model of deep learning with convolutional neural networks designed for image recognition tasks. This architecture is composed of 16 layers of artificial neurons that operate sequentially over the image data to improve the performance of the system. Less time is needed for training and optimization of the model. It also returns highly accurate image classification results because of its architecture and amount of layers [18].

6) Lightweight Custom CNN Model : It is a lightweight costumed model that is specifically designed with layers that progressively capture both high-level and fine-grained features from lung images. It consists of roughly seven convolutional layers, each followed by batch normalization, with several pooling layers in between. It uses the loss function categorical cross-entropy since this is a multi-class classification problem. SGD is chosen as an optimizer with a low learning rate (0.001) which ensures smooth and stable convergence. Despite its high accuracy, the model remains computationally efficient due to its moderate architecture size compared to heavy models. This lightweight model's performance shows its ability to effectively classify lung cancer images while balancing both speed and precision.

7) EfficientFormer: EfficientFormer is a family of vision transformer models designed for efficient deployment on mobile devices. EfficientFormer uses a more efficient attention mechanism compared to the standard dot-product attention in DeiT. EfficientFormer has a smaller size, which adds to the causes of high computational efficiency. It allows for fast inference speeds, suitable for real-time applications, yet achieves high accuracy comparable with larger, more complex models. With minimal latency, its latency-driven slimming ensures high performance [19].

8) DeiT : The original DeiT model was primarily designed for image classification. It is an advanced vision transformer model designed to achieve high performance in image classification tasks with significantly less data compared to traditional vision transformers. In lung cancer classification the patch tokens represent the image segments. [20].

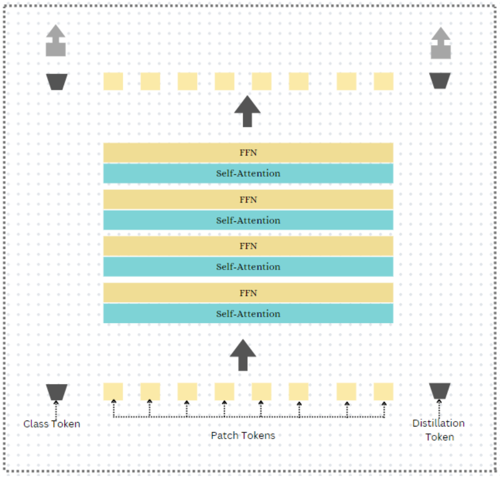

Fig. 3. Shows the architecture of DeiT.

In lung cancer classification these patch tokens represent the image segments. Different local regions of the medical image such as different parts of a CT scan or X-ray are captured by them. The class token used here is responsible for aggregating information from the patches and predicts the final class, such as the presence or absence of lung cancer. The Distillation Token is unique to DeiT, this token enhances the model's performance by learning from a teacher model during training, aiding the model in better generalizing to unseen lung cancer data.

It converges faster and more smoothly than other vision transformers due to optimized attention layers and training. We were able to achieve the highest test accuracy of 99.9% by running only 7 epochs.

IV. RESULT AND DISCUSSION

This section presents the results and discussion of the proposed system for the classification of lung cancer. A couple of deep learning methodologies have been discussed here with their efficiency. The results were considered on different metrics such as accuracy, precision, recall, F1-score, etc., which are enough to describe the model's performance in detail. The performance of the algorithms is shown in "Graph 1". This graph allows us to visualize how the system and the model we used performed.

A. Model Comparison

Table 1 highlights the machine learning models used in this study, including pre-trained deep learning models such as ConvNext, InceptionV3, ResNet50, VGG16, and DenseNet121.

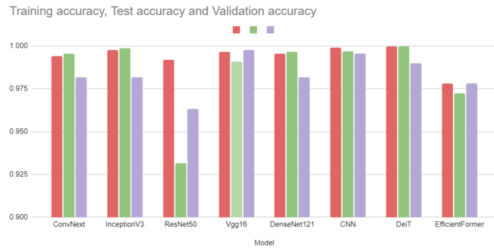

Fig. 4. Shows the accuracy of different models in the proposed lung cancer prediction system.

All the models were trained over 50 epochs. Among these, InceptionV3 performed best, with a training accuracy of 0.9954 and a test accuracy of 0.9966. Additionally, we developed a lightweight custom CNN model, specifically designed to capture both high-level and fine-grained features from lung images. This model achieved an impressive test accuracy of 0.9970. This model indicates near-perfect discrimination among normal, benign, and malignant cases. Moreover, we utilized two transformer models, DeiT and EfficientFormer. EfficientFormer achieved an accuracy of 0.9999, while DeiT matched this accuracy with even less time. Consequently, the DeiT transformer model outperformed all other models in our study.

TABLE I
PERFORMANCE OF DIFFERENT CLASSIFIERS IN LUNG CANCER PREDICTION

Model Name	Training Accuracy	Test Accuracy	Epoch
ConvNext	0.99429	0.99543	50
InceptionV3	0.99771	0.99885	50
ResNet50	0.99201	0.93151	50
Vgg16	0.99657	0.99772	50
DenseNet121	0.99543	0.99657	50
Lightweight CNN	0.99901	0.99712	50
EfficientFormer	0.97830	0.96430	50
DeiT	0.99999	0.99890	7

CONCLUSION

This analysis focuses on detecting and classifying lung cancer from CT scan images using various deep learning models. Early identification of lung malignancy is crucial due to rising mortality rates. Our goal is to attain high accuracy and prediction rates with quicker convergence, utilizing smaller datasets and fewer epochs. To this end, we employed a custom lightweight CNN, transformers, and other deep learning models to enhance accuracy and sensitivity. Specifically, we utilized two transformer models: 1) Data-Efficient Image Transformer (DeiT) and 2) EfficientFormer. Among these, DeiT outperformed the others, demonstrating that good accuracy can be achieved on a small dataset with fewer epochs compared to traditional deep learning techniques.

REFERENCES

[1] Ramana, K., Kumar, M. R., Sreenivasulu, K., Gadekallu, T. R., Bhatia, S., Agarwal, P., & Idrees, S. M. (2022c). Early prediction of lung cancers using Deep Saliency Capsule and Pre-Trained Deep Learning frameworks. *Frontiers in Oncology*, *12*.

[2] World Health Organization: WHO & World Health Organization: WHO. (2023, June 26). *Lung cancer*.

[3] Lung cancer. (2024, May 1). Cleveland Clinic.

[4] Tekade, R., & Rajeswari, K. . Lung cancer detection and classification using deep learning. In IEEE Conference Publication — IEEE Xplore 2018.

[5] Katiyar, P., & Singh, K. A Comparative study of Lung Cancer Detection and Classification approaches in CT images. In *IEEE Conference Publication — IEEE Xplore 2020*.

[6] Ren, Z., Zhang, Y., & Wang, S. (2022). A hybrid framework for lung cancer classification. *Electronics*, *11*(10), 1614.

[7] Musthafa, M.M., Manimozhi, I., Mahesh, T.R. *et al.* Optimizing double-layered convolutional neural networks for efficient lung cancer classification through hyperparameter optimization and advanced image pre-processing techniques. *BMC Med Inform Decis Mak* 24, 142 (2024).

[8] Humayun, M.; Sujatha, R.; Almuayqil, S.N.; Jhanjhi, N.Z. A Transfer Learning Approach with a Convolutional Neural Network for the Classification of Lung Carcinoma. Healthcare 2022, 10, 1058.

[9] Rehan Raza, Fatima Zulfiqar, Muhammad Owais Khan, Muhammad Arif, Atif Alvi, Muhammad Aksam Iftikhar, Tanvir Alam, Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images, Engineering Applications of Artificial Intelligence, Volume 126, Part B, 2023, 106902, ISSN 0952-1976.

[10] Medjahed, S. A., Saadi, T. A., Benyettou, A., & Ouali, M. (2017). Kernel-based learning and feature selection analysis for cancer diagnosis. *Applied Soft Computing*, *51*, 39–48.

[11] Ramana, K., Kumar, M. R., Sreenivasulu, K., Gadekallu, T. R., Bhatia, S., Agarwal, P., & Idrees, S. M. Early prediction of lung cancers using Deep Saliency Capsule and Pre-Trained Deep Learning frameworks. *Frontiers in Oncology*, *v-12*, 2022.

[12] Enhancing Lung Cancer Diagnosis through Advanced CT Scan Image Analysis: A Novel Approach using Mask-EffNet. *Nanotechnology Perceptions*, I.N.-S5, *v-20*, 2024.

[13] alyasriy, hamdalla; AL-Huseiny, Muayed (2021), "The IQ-OTHNCCD lung cancer dataset", Mendeley Data, V2, doi: 10.17632/bhmdr45bh2.2.

[14] Erdoğan, A. ConvNEXT — Next generation of convolutional networks. *Medium*. (April 1, 2023)

[15] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).

[16] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[17] *Densely connected convolutional networks*. (2017, July 1). IEEE Conference Publication — IEEE Xplore.

[18] Thakur, R. (2024, March 12). *Beginner's Guide to VGG16 implementation in Keras*. Built-In.

[19] Li, Y., Yuan, G., Wen, Y., Hu, J., Evangelidis, G., Tulyakov, S., ... & Ren, J. (2022). Efficientformer: Vision transformers at mobilenet speed. *Advances in Neural Information Processing Systems*, *35*, 12934-12949.

[20] Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., & Jégou, H. (2021, July). Training data-efficient image transformers & distillation through attention. In *International conference on machine learning* (pp. 10347-10357). PMLR.