EAI Endorsed Transactions
on Pervasive Health and Technology
Research Article EAI.EU

# Application of Several Transfer Learning Approach for Early Classification of Lung Cancer

Janjhyam Venkata Naga Ramesh[1], Raghav Agarwal[2*], Polireddy Deekshita[3], Shaik Aashik Elahi[4], Saladi Hima Surya Bindu[5], Juluru Sai Pavani[6]

[1,3,4,5,6] Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur Dist., Andhra Pradesh - 522302, India
[2] School of Computer Science & Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh, India

## Abstract

INTRODUCTION: Lung cancer, a fatal disease characterized by abnormal cell growth, ranks as the second most lethal worldwide, as observed in recent research conducted in India and other regions. Early detection is crucial for effective treatment, and manual differentiation of nodule types in CT images poses challenges for radiologists.
OBJECTIVES: To enhance accuracy and efficiency, deep learning algorithms are proposed for early lung cancer detection. Transfer learning-based computer recognition algorithms have shown promise in providing radiologists with additional insights.
METHODS: The dataset used in this study comprises 1000 CT scan images representing lung large cell carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, and normal lung cases. A preprocessing phase, including picture rescaling and modification, is applied to the input CT scan images of the lungs, followed by the utilization of a specific transfer learning model to develop a lung cancer detection system.
RESULTS: The performance of various transfer learning strategies is evaluated using measures such as accuracy, precision, recall, specificity, area under the curve, and F1-score.
CONCLUSION: Comparative analysis indicates that VGG16 outperforms other models in accurately categorizing different types of lung cancer.

**Keywords:** Lung Cancer Classification, Deep Learning, Computed Tomography (CT), Transfer Learning, Clinical Decision Support System

Received on 13 December 2023, accepted on 09 March 2024, published on 15 March 2024



doi: 10.4108/eetpht.10.5434

*Corresponding author. Email: raghav.g2106@gmail.com

## 1. Introduction

The World Health Organization (WHO) predicts that lung cancer, characterized by uncontrolled cell growth in the lungs, will surpass other cancer types and become the second most prevalent disease worldwide by 2020. Due to the aggressive nature of cancer cells, prevention and early detection of lung cancer pose challenges. Early diagnosis is crucial for effective treatment, as the tumor size and spread determine the cancer stage [1]. Non-small cell lung cancer (NSCLC) is the most common type, with adenocarcinoma (ACA), squamous cell carcinoma (SCC), and large cell carcinoma (LCC) being examples of abnormal tissue [2]. This research focused on classifying NSCLC subtypes using a dataset of 1,000 histological images of lung cancer cases.

Small cell lung cancer is less common than non-small cell lung cancer and has a higher mortality rate [3]. Symptoms such as hoarse voice, difficulty breathing, weight loss, persistent cough, chest tightness, and breathing problems indicate lung cancer, often at advanced stages [4]. Diagnostic methods may be insufficient for early detection. Deep learning algorithms are proposed as a more effective approach for identifying lung cancer in its early stages.

Lung cancer can develop and spread undetected, affecting the lungs, trachea, and major airways. It is more prevalent in individuals with lung conditions and can affect both smokers and non-smokers [5]. Risk factors include cigarettes, radon gas, air pollution, and chemical exposure in the workplace.

Computed tomography (CT) is commonly used for cancer prediction, but visual interpretation by radiologists can lead to delays in diagnosis and treatment [6]. Computer-aided diagnosis (CAD) is employed to accurately classify cancer types [7]. Transfer learning techniques have been studied for their effectiveness in image classification [8], particularly in identifying different types of lung cancer. This article presents the top-performing model and demonstrates improved classification results by preprocessing and organizing the data. The research provides benchmark data and epoch curves [9] to facilitate further investigation.

## 2. Literature Survey

Artificial intelligence (AI) has made significant advancements in the accurate and rapid diagnosis of various diseases [10]. Lung cancer has been a major focus of medical research. The following researchers have conducted relevant studies and proposed alternative methods for diagnosing lung cancer.

[11] conducted a study involving categories such as pneumonia, COVID-19, X-ray, lung cancer, and chest CT scan. The VGG19 + CNN model outperformed other models, achieving an accuracy of 98.05%.

[12] developed a Python module based on deep learning for classifying different types of cancer images. They utilized the CNN model and SVM algorithm, achieving an overall accuracy of 94% with the support vector machine model.

[13] proposed a simplified deep-learning approach using a CNN architecture with four convolutional layers. The model achieved an accuracy of 97.9% and was considered suitable for real-time interpretation of CT scans due to its reduced computational requirements compared to state-of-the-art CNN architectures.

[14] employed the AlexNet, LeNet, and VGG16 DL models for lung cancer detection. The combination of feature extraction and an ANN classifier in AlexNet yielded a classification accuracy of 98.74%.

[15] suggested a rapid and automated hyperparameter-optimized random forest classification model, achieving 99.8% AUC values and 97.97% accuracy.

[16] identified lung cancer using LR, DT, NB, KNN, and SVM classifiers. Cure rates of 94.59% for LR, 81.08% for DT, 87.78% for NB, 89.18% for SVM, and 88.28% for ANN were achieved using lung cancer datasets.

[17] utilized NB, SVM, and RF classifiers on two different CT image sets. They achieved accuracies of 94.6% for NB, 90.0% for SVM, 92.3% for RF on the first dataset, and 93.6% for NB, 94.6% for SVM, and 94.2% for RF on a separate record.

In conclusion, most classification techniques focus on binary rather than multiclass classification. However, the ability to categorize lung nodules into multiple classes makes multiclass classification more favorable. Furthermore, there is a potential to discover more precise information about different forms of lung cancer.

## 3. Methodology

This research aims to compare results while using different transfer learning techniques to identify different types of lung cancer. Transfer learning is a popular strategy in neural networks where pre-trained models or models made for a particular task are used again as starting points for a model in another task to improve model performance and speed up the training. Transfer learning aids students in acquiring new abilities more rapidly by allowing them to tackle problems of a similar kind by using the features and weights from a previously trained model rather than creating a model from scratch. For the model to succeed at the task, it must learn how to effectively extract characteristics from photographs. Because the model must generate predictions for a considerable number of classes and the data will be trained on a large number of images, the transfer learning approach is thus effective.

### 3.1. Dataset

Data collection is a crucial task in the development of a detection and classification model. In this section, we present the dataset used in the current research on deep learning for lung cancer diagnosis. The CT scan images were obtained from Kaggle [18]. The dataset consists of 1000 images that need to be classified into four classes: squamous cell, adenocarcinoma, large cell carcinoma, and normal cell. The data is divided into train, test, and validation sets, with proportions of 70%, 20%, and 10%, respectively. Each of these datasets is further divided into four classes. The distribution of the dataset and the number of images in each class can be seen in Table 1.

### 3.2. Data preprocessing

Data preparation entails organizing and refining raw data, and it is an essential step before training the model. As part of the preprocessing stage, the training data undergo scaling, tagging, and normalization procedures. To comply with the input size requirements of the transfer learning approaches considered, all input images were resized to 224x224. To enhance accuracy and reduce loss values, the images are aggregated along with their corresponding labels for each batch.

Table 1. Dataset Distribution

| Dataset | ACA | SCC | LCC | Normal | Total Images (Per dataset) |
|---|---|---|---|---|---|
| Train | 195 | 155 | 115 | 148 | 613 |
| Test | 120 | 90 | 51 | 54 | 315 |
| Validation | 23 | 15 | 21 | 13 | 72 |
| Total Images (per class) | 338 | 260 | 187 | 215 | 1000 |

## 3.3. Classification Models

In the current study, the transfer learning techniques VGG, RESNET, DENSENET, MOBILENET, INCEPTION, and XCEPTION were employed. These techniques made use of pre-trained models trained on the widely recognized ImageNet dataset. The ImageNet dataset comprises 14 million images categorized into 1000 classes, serving as a benchmark for computer vision.

Due to privacy and security concerns, comprehensive medical data could not be utilized in this study. Consequently, transfer learning techniques were given primary consideration. Transfer learning played a significant role in this study, particularly due to the limited dataset of only 1000 images. To enhance model performance, different parameters such as batch size and epochs were adjusted during fine-tuning. The dataset was trained using a batch size of 32 and 40 epochs. The outcomes are then compared to find the suggested model in subsequent analysis.

# 4. Results and Discussion

This section introduces the context, certain metrics of performance, and the results of the transfer learning algorithms for multi-class classification.

## 4.1. System

The ongoing research was conducted using a computer system that satisfied specific specifications. Our system consisted of an AMD Ryzen 7 5800H processor paired with a Radeon graphics processor, operating at a speed of 3.20 GHz. It included 32 GB of RAM, a 64-bit operating system, and a 512 GB SSD. For the study, we utilized an NVIDIA RTX 3050 GPU. Moreover, the research was carried out using Python 3.8, along with Scikit-learn, OpenCV, and Matplotlib libraries.

## 4.2. Evaluation Criteria

Once the model is trained and images are classified, an evaluation is conducted to assess the performance of the model. According to the existing literature [15], multiple parameters are employed for evaluating the model, including Accuracy, Precision, Recall, F1 score, Specificity, and Area Under the Curve (AUC).

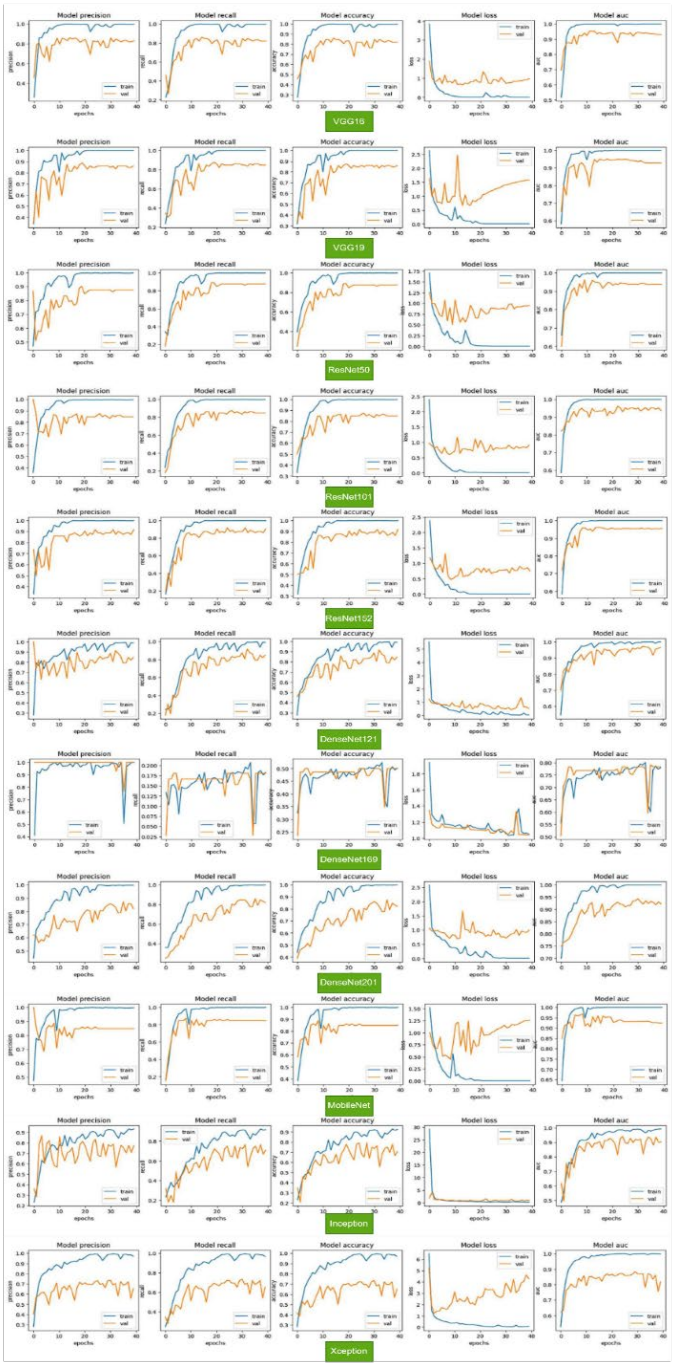


**Figure 1.** Metrics comparison of different deep learning models

**ACCURACY.** The neural network's ability to accurately differentiate between multiple categories, such as adenocarcinoma (ACA), squamous cell carcinoma (SCC), large cell carcinoma (LCC), and normal lung tissues, is referred to as its discriminative capacity. The outcome is determined by dividing the total number of

accurate predictions made by the classifier by the overall number of predictions.

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

**PRECISION.** The ratio of accurate positive forecasts to all positive forecasts can be computed using the following calculation.

$$Pre = \frac{TP}{TP+FP} \quad (2)$$

**RECALL OR SENSITIVITY.** Accurately identifying individuals with a specific ailment is crucial. The ratio of precise positive predictions to the actual positive samples can be employed to summarize this accuracy. The following formula can be used to calculate this ratio:

$$Rec = \frac{TP}{TP+FN} \quad (3)$$

**SPECIFICITY.** It is utilized to precisely identify individuals who are disease-free.

$$Spec = \frac{TN}{TN+FP} \quad (4)$$

**F1 SCORE.** The capability to accurately categorize data into appropriate classes, considering the multiple classifications, is known as the ability to precisely partition classes. Named after its computation as the harmonic mean of recall and accuracy, the F1 score determines the balance between false positive and false negative classifications. It is calculated using the following equation.

$$F1 - Score = \frac{2*(Pre+Rec)}{Pre+Rec} \quad (5)$$

$$F1 - Score = \frac{2*(TP)}{2TP+FP+FN} \quad (6)$$

**AUC-ROC.** Receiver Operating Characteristic (ROC) is a technique used for visualizing data and selecting the most suitable classifier for each classification criterion. It involves plotting the true-positive rates (TPR) against the false-positive rates (FPR). By varying the threshold levels, a probability curve is created that effectively differentiates the signal from the noise. The Area Under the Curve (AUC) is a summary of the ROC curve, which quantifies the classifier's ability to distinguish between classes. The AUC is calculated using the following method:

$$\int_{-\infty}^{\infty} TPR\ (T)\ FPR'(T)\ dT \quad (7)$$

Here, FPR' (T) = FPR is the first derivative of concerning T, where T stands for sample data.

As stated earlier, metrics such as true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) have been considered. Figure 1 illustrates the precision, recall, accuracy, loss, and AUC-score graphs for various applied transfer learning models over 40 epochs.

## 3.3. Results Comparison

The comparative evaluation of algorithm performance was conducted based on the parameters presented in Table 2. The analysis indicates that among the transfer learning algorithms utilized in this study, VGG16 achieved the highest accuracy (84.76%). Moreover, considering additional factors such as precision, recall, AUC, and F1-score, the results consistently demonstrate that the VGG16 algorithm outperforms the other algorithms utilized in this research.

Table 2. Results Comparison Table

| Models | Validation Accuracy(%) | Test Accuracy(%) | Precision(%) | Recall(%) | AUC(%) | F1-Score(%) |
|---|---|---|---|---|---|---|
| **VGG16** | 81.94 | **84.76** | **86.52** | **84.76** | **95.59** | **85.63** |
| VGG19 | 86.11 | 82.86 | 83.36 | 82.85 | 93.99 | 83.10 |
| ResNet50 | 87.50 | 77.46 | 78.66 | 77.46 | 90.03 | 78.05 |
| ResNet101 | 84.72 | 82.86 | 84.14 | 82.85 | 94.36 | 83.49 |
| ResNet152 | **91.67** | 81.27 | 81.47 | 81.26 | 91.84 | 81.36 |
| DenseNet121 | 84.72 | 40.32 | 64.26 | 40.31 | 70.18 | 49.54 |
| DenseNet169 | 50.00 | 53.33 | 62.90 | 53.33 | 80.82 | 57.72 |
| DenseNet201 | 81.94 | 69.52 | 72.80 | 69.52 | 87.80 | 71.12 |
| MobileNet | 84.72 | 55.56 | 72.26 | 55.55 | 74.45 | 62.81 |
| Inception | 70.83 | 46.98 | 47.43 | 46.98 | 74.34 | 47.20 |
| Xception | 65.28 | 45.08 | 66.29 | 45.07 | 71.61 | 53.65 |

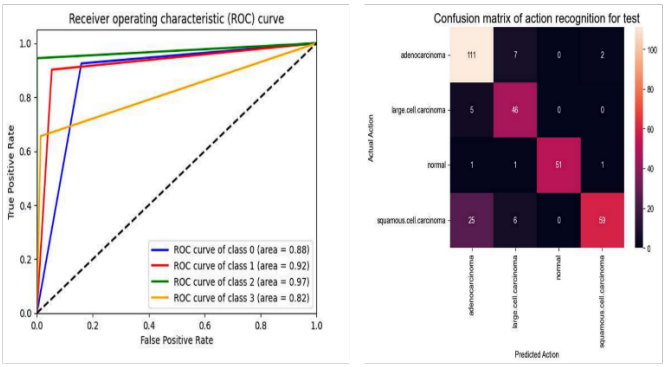


**Figure 2.** AUC-ROC curve and Confusion matrix of the best transfer learning model (VGG16)

Further, Table 3 provides a comprehensive overview of Precision, Recall, F-measure, Specificity, and AUC scores for each class, specifically about the applied transfer learning models. Additionally, Fig. 2. visually presents the AUC-ROC curve and Confusion matrix of the top-performing transfer learning model, which is VGG16.

Table 3. Comprehensive overview of metrics for each class

| Metrics/Models | VGG16 | VGG19 | ResNet 50 | ResNet 101 | ResNet 152 | DenseNet 121 | DenseNet 169 | DenseNet 201 | MobileNet | InceptionV3 | Xception |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision (adenocarcinoma) | 0.78 | 0.76 | 0.68 | 0.85 | 0.80 | 0.50 | 0.45 | 0.62 | 0.61 | 0.42 | 0.79 |
| Precision Score (large cell) | 0.77 | 0.86 | 0.76 | 0.90 | 0.77 | 0.22 | 0.00 | 0.62 | 0.31 | 0.40 | 0.26 |
| Precision Score (normal) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.96 | 1.00 | 0.80 | 1.00 | 0.85 | 0.91 |
| Precision Score (squamous cell) | 0.95 | 0.82 | 0.82 | 0.70 | 0.75 | 0.88 | 1.00 | 0.89 | 0.93 | 0.37 | 0.58 |
| Recall Score (adenocarcinoma) | 0.93 | 0.83 | 0.86 | 0.77 | 0.73 | 0.03 | 1.00 | 0.82 | 0.38 | 0.08 | 0.09 |
| Recall Score (large cell) | 0.90 | 0.73 | 0.49 | 0.75 | 0.86 | 1.00 | 0.00 | 0.59 | 0.98 | 0.45 | 0.96 |
| Recall Score (normal) | 0.94 | 0.94 | 0.96 | 0.98 | 0.98 | 0.94 | 0.87 | 0.94 | 0.94 | 0.94 | 0.93 |
| Recall Score (squamous cell) | 0.66 | 0.81 | 0.71 | 0.87 | 0.79 | 0.24 | 0.01 | 0.44 | 0.31 | 0.71 | 0.36 |
| Specificity Score (adenocarcinoma) | 0.84 | 0.83 | 0.74 | 0.91 | 0.88 | 0.98 | 0.25 | 0.69 | 0.85 | 0.92 | 0.98 |
| Specificity Score (large cell) | 0.94 | 0.97 | 0.96 | 0.98 | 0.95 | 0.31 | 0.99 | 0.93 | 0.58 | 0.86 | 0.46 |
| Specificity Score (normal) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 0.95 | 1.00 | 0.96 | 0.98 |
| Specificity Score (squamous cell) | 0.98 | 0.92 | 0.93 | 0.84 | 0.89 | 0.98 | 1.00 | 0.97 | 0.99 | 0.51 | 0.89 |
| F1-Score (adenocarcinoma) | 0.85 | 0.79 | 0.76 | 0.81 | 0.77 | 0.05 | 0.62 | 0.71 | 0.47 | 0.85 | 0.16 |
| F1-Score (large cell) | 0.83 | 0.79 | 0.60 | 0.82 | 0.81 | 0.36 | 0.00 | 0.61 | 0.48 | 0.83 | 0.40 |
| F1-Score (normal) | 0.97 | 0.97 | 0.98 | 0.99 | 0.99 | 0.95 | 0.93 | 0.86 | 0.97 | 0.97 | 0.92 |
| F1-Score (squamous cell) | 0.78 | 0.82 | 0.76 | 0.77 | 0.77 | 0.38 | 0.02 | 0.59 | 0.47 | 0.78 | 0.44 |
| ROC-AUC Score (adenocarcinoma) | 0.88 | 0.83 | 0.80 | 0.84 | 0.81 | 0.50 | 0.63 | 0.75 | 0.62 | 0.51 | 0.54 |
| ROC-AUC Score (large cell) | 0.92 | 0.85 | 0.73 | 0.86 | 0.91 | 0.66 | 0.50 | 0.76 | 0.78 | 0.66 | 0.71 |
| ROC-AUC Score (normal) | 0.97 | 0.97 | 0.98 | 0.99 | 0.99 | 0.97 | 0.94 | 0.95 | 0.97 | 0.95 | 0.95 |
| ROC-AUC Score (squamous cell) | 0.82 | 0.87 | 0.82 | 0.86 | 0.84 | 0.62 | 0.51 | 0.71 | 0.65 | 0.61 | 0.63 |

## 5. Conclusion

Lung cancer stands as the primary cause of global mortality. Early and accurate diagnosis plays a crucial role in improving therapy outcomes and survival rates for various cancer types. Given the significance of early detection, this study aims to develop an optimal multi-class classification model capable of accurately distinguishing between adenocarcinoma (ACA), squamous cell carcinoma (SCC), large cell carcinoma (LCC), and normal lung tissues.

To underscore the importance of deep transfer learning, this research builds upon previous studies focused on lung cancer categorization and detection. The foundations of transfer learning are initially explained, followed by an evaluation of various models employing deep transfer learning for lung cancer detection. This paper serves as a valuable resource for a comprehensive comparison of deep learning algorithms in the context of lung cancer detection and classification [20].

Furthermore, we believe that the proposed technique has the potential to accurately identify several diseases. Given the severity of lung cancer, it is crucial to employ all available tools to combat this ailment. The recommended VGG16 model proves beneficial in identifying lung cancer, reducing mortality rates, and enhancing the precision of tumor cell outcomes.

## 6. Future Remarks

To enhance the model's performance, we can employ strategies such as utilizing a large volume of data, adjusting various variables, and training on multiple models. However, when considering real-time AI scenarios, it's important to note that a CT image can contain nodules that are both benign and malignant. One limitation of the proposed approach is the lack of differentiation between normal and cancerous cell cases within a single CT scan. In the future, addressing this issue may involve developing a specialized multiple-instance learning model capable of distinguishing between different types of nodules within a single CT image.

Furthermore, this research extends beyond lung cancer detection and holds potential applications in industries and other healthcare domains. Additionally, certain federated learning and Internet of Things (IoT)-based algorithms will be utilized on various datasets as part of our study. These approaches aim to address the challenges posed by incomplete annotation of data in the field of medical imaging.

## References


[1] Smita Raut1, Shraddha Patil2, Gopichand Shelke", Lung Cancer Detection using Machine Learning Approach", International Journal of Advance Scientific Research and Engineering Trends (IJASRET), 2021.

[2] Oshima, Y., Shinzawa, H., Takenaka, T., Furihata, C., & Sato, H. (2010). Discrimination analysis of human lung cancer cells associated with histological type and malignancy using Raman spectroscopy. Journal of biomedical optics, 15(1), 017009.

[3] Cancer.org. 2021. Lung Cancer Statistics — How Common is Lung Cancer?. [online] Available at: https://www.cancer.org/cancer/lungcancer/about/key-statistics.html [Accessed 9 August 2021].

[4] V. Krishnaiah, G. Narsimha, and N. S. Chandra, "Diagnosis of lung cancer prediction system using data mining classification techniques", International Journal of Computer Science and Information Technologies, vol. 4, no. 1, 2013, pp. 39-45.

[5] Xing, PuĞYuan, et al. "What are the clinical symptoms and physical signs for nonĞsmall cell lung cancer before diagnosis is made? A nationĞwide multicenter 10Ğyear retrospective study in China." Cancer medicine 8.8 (2019): 4055-4069.

[6] A. Bhattacharjee, R. Murugan, and T. Goel, "A hybrid approach for lung cancer diagnosis using optimized random forest classification and k-means visualization algorithm," Health and Technology, pp. 1– 14, 2022.

[7] S. D. Pande and R. Agarwal, "Multi-Class Kidney Abnormalities Detecting Novel System Through Computed Tomography," in IEEE Access, doi: 10.1109/ACCESS.2024.3351181.

[8] R. Agarwal and D. Godavarthi, "Skin Disease Classification Using CNN Algorithms", *EAI Endorsed Trans Perv Health Tech*, vol. 9, Oct. 2023.

[9] R. Agarwal, A. S. Sathwik, D. Godavarthi, and J. V. Naga Ramesh, "Comparative Analysis of Deep Learning Models

for Multiclass Alzheimer's Disease Classification", *EAI Endorsed Trans Perv Health Tech*, vol. 9, Nov. 2023.

[10] R. Agarwal, S. D. Pande, S. N. Mohanty and S. K. Panda, "A Novel Hybrid System of Detecting Brain Tumors in MRI," in IEEE Access, vol. 11, pp. 118372-118385, 2023, doi: 10.1109/ACCESS.2023.3326447.

[11] Ibrahim, D. M., Elshennawy, N. M., & Sarhan, A. M. (2021). Deepchest: Multi-classification deep learning model for diagnosing COVID19, pneumonia, and lung cancer chest diseases. Computers in biology and medicine, 132, 104348.

[12] Wang, Y., Yang, L., Webb, G. I., Ge, Z., & Song, J. (2021). OCTID: a one-class learning-based Python package for tumor image detection. Bioinformatics, 37(21), 3986-3988.

[13] Masud, M., Muhammad, G., Hossain, M.S., Alhumyani, H., Alshamrani, S.S., Cheikhrouhou, O. and Ibrahim, S., 2020. Light deep model for pulmonary nodule detection from CT scan images for mobile devices. Wireless Communications and Mobile Computing, 2020.

[14] M. To˘gac ¸ar, B. Ergen, and Z. C¨omert, "Detection of lung cancer on chest ct images using minimum redundancy maximum relevance feature selection method with convolutional neural networks," Biocybernetics and Biomedical Engineering, vol. 40, no. 1, pp. 23– 39, 2020.

[15] R. Agarwal, J. Suthar, S. K. Panda, and S. N. Mohanty, "Fuzzy and Machine Learning based Multi-Criteria Decision Making for Selecting Electronics Product", *EAI Endorsed Scal Inf Syst*, vol. 10, no. 5, Jul. 2023.

[16] A. Bhattacharjee, R. Murugan, B. Soni, and T. Goel, "Ada-gridrf: A fast and automated adaptive boost based grid search optimized random forest ensemble model for lung cancer detection," Physical and Engineering Sciences in Medicine, pp. 1–14, 2022.

[17] S. Jaiswal, R. Deshmukh, N. Meshram, A. Deshpande, and B. Wakode, "A comparative study of lung cancer detection using machine learning algorithms", International Journal of Scientific Research in Engineering and Management, vol. 6, no. 6, 2022, pp. 1-6.

[18] A. K. Dutta, "Detecting lung cancer using machine learning techniques", Intelligent Automation & Soft Computing, vol. 31, no. 2, 2022, pp. 1007-1023.

[19] [2020]. Link: https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images?select=Data

[20] Lakshmanaprabu, S. K., Mohanty, S. N., Shankar, K., Arunkumar, N., & Ramirez, G. (2019). Optimal deep learning model for classification of lung cancer on CT images. *Future Generation Computer Systems*, *92*, 374-382.