ADVANCED LUNG CANCER DETECTION MODEL WITH U-NET-BASED SEGMENTATION, ARI-TFMOA FEATURE OPTIMIZATION AND MACHINE LEARNING CLASSIFICATION

SATISHKUMAR PATNALA[1], ANNEMNEEDI LAKSHMANARAO[2], SUJAN BABU VADDE[3], KESHETTI SREEKALA[4], MOUPALI ROY[5], Dr. ASMITA MANNA[6]

[1]Assistant Professor, Department of Computer Science & Engineering (AI&ML), GMR Institute of Technology, Rajam, AP, India
[2]Assistant Professor, Department of IT, Aditya University, Surampalem, India
[3]Assistant Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India
[4]Department of Computer Science Engineering, Mahatma Gandhi Institute of Technology (A), Hyderabad, India
[5]Assistant Professor, Department of Electronics and Communication Engineering, Narula Institute of Technology, Kolkata, India
[6]Assistant Professor, Department of Computer Engineering, Pimpri Chinchwad College of Engineering, Pune, India
E-mail: [1]srtsatishsrt@gmail.com, [2]laxman1216@gmail.com, [3]sujanbabuv@kluniversity.in, [4]ksrikala_cse@mgit.ac.in, [5]moupali.roy@nit.ac.in, [6]asmita.manna@pccoepune.org

ABSTRACT

Lung cancer is among the deadliest diseases globally, highlighting the importance of early and accurate detection to enhance survival chances. Traditional deep learning models face challenges such as high computational costs and lack of interpretability, while conventional machine learning classifiers require high-quality feature representations for optimal performance. To address these challenges, this study proposes an advanced lung cancer detection model integrating segmentation, feature optimization and classification using both machine learning and deep learning models. The framework employs AMTUnet++-ASPP for abnormality segmentation, ensuring accurate lung nodule detection from CT scan images. The segmented tumor regions undergo feature optimization using Adaptive Radiant Inertia Tuned Fuzzy Multi-Objective Algorithm (ARI-TFMOA), which refines extracted features and eliminates redundant information. The optimized feature vectors are then classified using multiple models, including ML algorithms and an artificial neural network (ANN) to distinguish between normal, benign, and malignant lung conditions. To evaluate the effectiveness of the proposed model, a comparative analysis is conducted by implementing three different approaches: a baseline CNN for direct classification, a U-Net-based segmentation followed by ML classification, and the proposed AMTUnet++-ASPP segmentation with ARI-TFMOA optimized classification using both ML models and ANN. The results demonstrate that the integration of advanced segmentation and optimized feature selection significantly enhances classification performance. In particular, the ANN model outperformed other classifiers, leading to superior accuracy, precision, and recall values. This hybrid approach, combining deep learning-based segmentation, feature optimization, and deep learning classification, provides a computationally efficient and clinically interpretable solution for AI-assisted lung cancer diagnosis, reinforcing its potential for early detection and decision support for radiologists.

Keywords: *Lung Cancer Detection, CT scan images, AMTUnet++, ARI-TFMOA, Machine Learning, Deep Learning.*

1. INTRODUCTION

Lung cancer continues to rank among the most common and deadliest forms of cancer globally, contributing significantly to the overall number of cancer-related deaths. Early and accurate detection plays a important role in enhancing patient outcomes, enabling timely medical intervention, and increasing survival rates. Traditional methods, such as manual investigation of CT scans by radiologists, are time-consuming, chance to human mistakes, and depends on expertise. As a result, artificial intelligence (AI)-driven techniques have gained attention for their ability to enhance the accuracy of lung cancer detection.

Deep learning, especially CNNs, has already proven its success in medical image analysis tasks, in which high-level features are automatically elicited. Unfortunately, CNN-based classification falls short in terms of interpretability, overfitting, and high computational cost. In addition, raw CT scans are full of complex anatomical structures, which give DL models a hard time to differentiate between cancerous regions and normal lung tissues. To overcome these drawbacks, segmentation-based methods, such as the U-Net and its advanced methods, have been proposed to facilitate the isolation of lung nodules from the CT images, suppression of some irrelevant background and enhancement of the classification performance.

To address these challenges and improve classification performance, in this study, a modern screening model for lung cancer detection is proposed, which combines DL-based segmentation with feature refinement and the hybrid-based classification of both ML and ANN models. The proposed method uses AMTUnet++-ASPP, a strong U-Net like architecture, for precise lung tumor segmentation, which is able to learn local and global contextual features at different scales. The segmented tumor regions are subjected to feature optimization using ARI-TFMOA, a fine-tuning process that further adjusts the selected features to discard irrelevant information.

Then, the optimized feature vectors are discriminated by utilizing the several models, including a five ML classifiers and ANN model, for separating normal, benign and malignant lung swellings. Finally, a comparative study is held with three methods, baseline CNN for direct classification, U-Net with ML-based classification (segmentation followed by ML classification), and our proposed AMTUnet++ - ASPP (segmentation), with ARI-TFMOA (classification) and ML-ANN (classification) modes. Experiment is performed on a lung cancer Kaggle dataset whose CT scan images are classified into normal, benign and malignant cases. Results showed that combining advanced segmentation, optimized feature selection and DL-based classifiers dramatically improved classification performance in term of accuracy, precision, recall, and overall robustness of the analysis.

This work aims to improve the accuracy of early-stage lung cancer detection by means of a fast and interpretable AI framework. Deep learning: Current deep learning models typically require massive computing resources and are sometimes with limited transparency, while traditional machine learning techniques are confronted with the extraction of the redundant or low-quality feature representations. In addition, most of the existing studies are only dedicated to segment or classification, not both and thus cannot be combined seamlessly into a unified diagnostic pipeline. To address this, the proposed work aims to collectively perform accurate lung tumor segmentation, feature enhancement, and robust classification in an end-to-end base model. This method aims not only at improving the diagnostic accuracy but also at reducing false alarms and a real-time clinical deployment- thus fulfilling an urgent requirement in AI-augmented cancer screening.

2. LITERATURE SURVEY

R. S. Raj et al. [1] designed a guided backpropagation CNN for classification of chronic lung disease. Through feature extraction by using the EfficientNetB2 architecture, they used GBP method to find discriminative regions of the input images and achieved a better model performance. A chest X-ray radiography (CXR) data set was used to validate their methodology with image denoising and normalization as preprocessing option. As a result, it found that the image classification approaches with the proposed model (GBPCNN) outperformed compared to the currently prevalent models (MobileLungNetV2, and CX-Ultranet) in the aspects of the accuracy. Md. Imaran Hossain et al. [2] developed a hybrid LeNet-LSTM model targeting on lung cancer classification, aiming at helping radiologists work with the large number of CT scan images produced while screening. They used a model which integrated LeNet for feature extraction through convolutional/maxpool layers, together with LSTM, which extracts sequential data while

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

eliminating the vanishing gradient problem associated with recurrent networks. J.Zapata-Paulini et al. [3] explored numerous ML algorithms for prediction of lung cancer. They analyzed random forest, decision tree, KNN, SVC etc. on a Kaggle dataset consisting of 309 records and 16 attributes. Data analysis and model evaluation were performed based on specificity, sensitivity, F1 score, accuracy, and precision. As for the best performing models, SVC, RF reached top values with 97% accuracy in predicting lung cancer. B. K. Sathiyamurthy et al. [4] developed a new enhanced U-Net based model for the development and staging of lung cancer with TNM system. They used segmentation and also used CLAHE and augmentation for preprocessing. Using the augmented U-net model with ARESNET and batch normalization, the specific tumor regions from CT images were correctly segmented.

A.L. Rao et al. [5] proposed a hybrid model that combined auto-encoders for features extraction, and it increased classification performance. The results indicated that the combined method was superior to individual ML classifiers, providing a more efficient and reliable approach for early detection of lung cancer. Odeh et al. [6] used XAI techniques to interpret DL models on COVID-19 flavored lung X-ray images. A voting mechanism was introduced to mitigate biases in both structure and training of the models, contributing toward interpretability and affect user trust towards AI based diagnosis and automated infection annotation. K. R. Sekhar et al. [7] used three transfer learning architectures. The other eight transfer learning models, however, produced their relative 91% accuracy, which subsequently demonstrated the potency of VER-network in classifying the lung cancer. A. Saha et al. [8] introduced the transfer learning model (VER-Net) by applying three TL architectures in order to detect lung cancer with CT scan images. We enhanced model performance through preprocessing, data augmentation, and hyperparameter optimization. On this analysis, VER-Net outperformed the rest of the models.

R. Javid et al. [9] performed systematic literature reviews on lung cancer diagnosis techniques utilizing DL methods incorporating X-Rays, WSI CT scans and MRI. Shah et al. This study by [10] proposes to use a deep ensemble of several 2D CNN Architectures and a suggested 2D Deep Ensemble CNN in LUNA 16 dataset for lung nodule detection in order to enhance the detection performance. Their approach achieved an accuracy of 95% in distinguishing these two classes of nodules, beating baseline approaches. A. R. Wahab Sait et al. [11] Proposed a lung cancer detection model with PET/CT images based on DL, combining DenseNet-121, deep autoencoders and MobileNet V3-Small classifier. Mohamed et al. [12] proposed a hybrid metaheuristic-CNN approach to lung cancer classification based on CT images. They created a CNN architecture that trained its weights and biases EOSA. The performance testing of their EOSA-CNN model was carried out on the IQ-OTH/NCCD dataset, and it could give 93.21% accuracy. Thanoon et al. [13] performed a review on DL techniques for CT scan-based lung cancer screening and diagnosis. They used classification and segmentation approaches and discussed pros and cons of several models. The analysis also explored the considerable promise of DL for lung cancer detection and advanced ideas for future research direction to improve computer-assisted diagnosis systems.

Kumari et al. [14] trained several ML models and introduced ranking-based techniques for enhanced accuracy prediction. The team tested their approach on two datasets, achieving high detection accuracy, showcasing the value of ML in cancer diagnosis. S. Bharathy et al. [15] discussed ML techniques for lung cancer detection involving multi-stage classification to enable segmentation and enhancement. Hence for segmentation they applied Thresholding, marker-controlled watershed and binary classifiers, and evaluated several ML models. L. Wang et al. [16] described the progress of DL-based medical imaging tools for early lung cancer detection. Different imaging modalities such as chest X-ray, MRI, PET and CT and their limitations regarding automatic classification were discussed. Pandain et al. [17] presented a CNN and GoogleNet-based DL model for lung cancer detection on CT scan images. They adopted two VGG-16 networks as the base architecture for the region proposal and classification networks. R. U. Osarogiagbon et al. [18] performed a prospective observational study of two strategies to facilitate early lung cancer detection: LDCT screening and program-based management of lung nodules. P. Chaturvedi et al. [19] reviewed different CAD techniques of cancer detection, including image segmentation and classification methods. In their study, they reviewed the approaches based on ML and DL, as well as their comparison studies on the detection of lung cancer. S. Lalitha et al. [20] applied an automated framework using CT scan images where ML algorithms were employed to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

classify benign, malignant and normal case. With an accuracy of 98.7%, the new model outperformed the previous models.

C. Jacobs et al. [21] studied the performances of DL algorithms in a crowd-sourcing competition to detect lung cancer by low-dose Scan CT. Analyzing 300 patient scans, they compared the model outputs with the assessments of 11 radiologists, and measured performance with receiver operating characteristic (ROC) analysis. A hybrid CNN-SVM model was introduced for CT-based lung cancer classification by A. Y. Saleh et al. [22]. They classify CNN for extracting the features and SVM for filtering the information (219) from irrelevant data and achieved 97.91% of cancer detection. A CNN-based model for tumor detection in lung cancer was proposed by Abidi et al. [23]. This method showed high specificity and efficiency and surpassed conventional methods based on neural networks in detecting lung cancer at an early stage. Y. Guo et al. [24] proposed a multi-class classification using CNN to detect skeletal metastasis due to lung cancer from the scintigraphy imaging data. By utilizing hierarchical feature extraction and classification, they had high classification performance and could be applicable as unsupervised automated metastasis detection in lung cancer patients.

Although substantial improvements have been made in deep learning and machine learning approaches for lung cancer diagnosis, existing models are also hindered by some limitations, e.g., inefficiently operation, inexplicability, or incomplete feature expressions, etc. These popular CNN-based classification models can't avoid overfitting problem and the poor localization of tumor focuses and ML classifiers must rely on high-quality feature inputs. Even with the advance of some density-based segmentation methods such as U-Net and combined ML models, there is still no effective feature selection and fusion throughout the pipeline. It is desirable to develop an integrated model that can synergistically use advanced segmentation, optimized feature representation, and robust classifier for better diagnostic accuracy and clinical usefulness. In this paper, the gap is filled by the introduction of a new hybrid model that extends, and mitigates the shortcomings found in other works. This paper hypothesizes that a hybrid AI model combining deep learning-based segmentation, fuzzy multi-objective feature optimization, and ML/ANN-based classification will significantly improve diagnostic accuracy and robustness over traditional methods

3. RESEARCH METHODOLOGY

The proposed methodology is shown in Figure.1. The proposed methodology integrates DL-based segmentation, feature optimization, and hybrid classification using ML and ANN models to enhance lung cancer detection accuracy. The process begins with preprocessing CT scan images from a publicly available Kaggle dataset, including grayscale conversion, contrast enhancement, resizing, and data augmentation. Lung tumor segmentation is performed using AMTUnet++-ASPP, which refines U-Net++ with ASPP to improve multi-scale feature extraction and enhance segmentation precision.

Following segmentation, key geometric, statistical, and texture-based features are extracted and optimized using Adaptive Radiant Inertia Tuned Fuzzy Multi-Objective Algorithm (ARI-TFMOA), which removes redundant information and retains the most discriminative attributes. The optimized feature set is then classified using five ML models and an ANN, including Random Forest, Extreme Gradient Boosting, Decision Tree, AdaBoost, Gradient Boosting Classifier, and ANN. The dataset is split into training and testing sets in an 80-20 ratio, and hyperparameter tuning is applied to maximize classification performance. Finally, classification results are assessed based on several metrices... Illustrating a comparison between a baseline CNN classifier, a U-Net-based segmentation and ML-based classification, and AMTUnet++-ASPP with ARI-TFMOA indubitably confirms that the hybrid pipeline greatly outperforms the standard approach. It was clear from the study that the combination of DL segmentation, feature optimization, and ML classification gives a robust AI-assisted diagnostic tool for the early lung cancer detection and better patient's clinical decision.

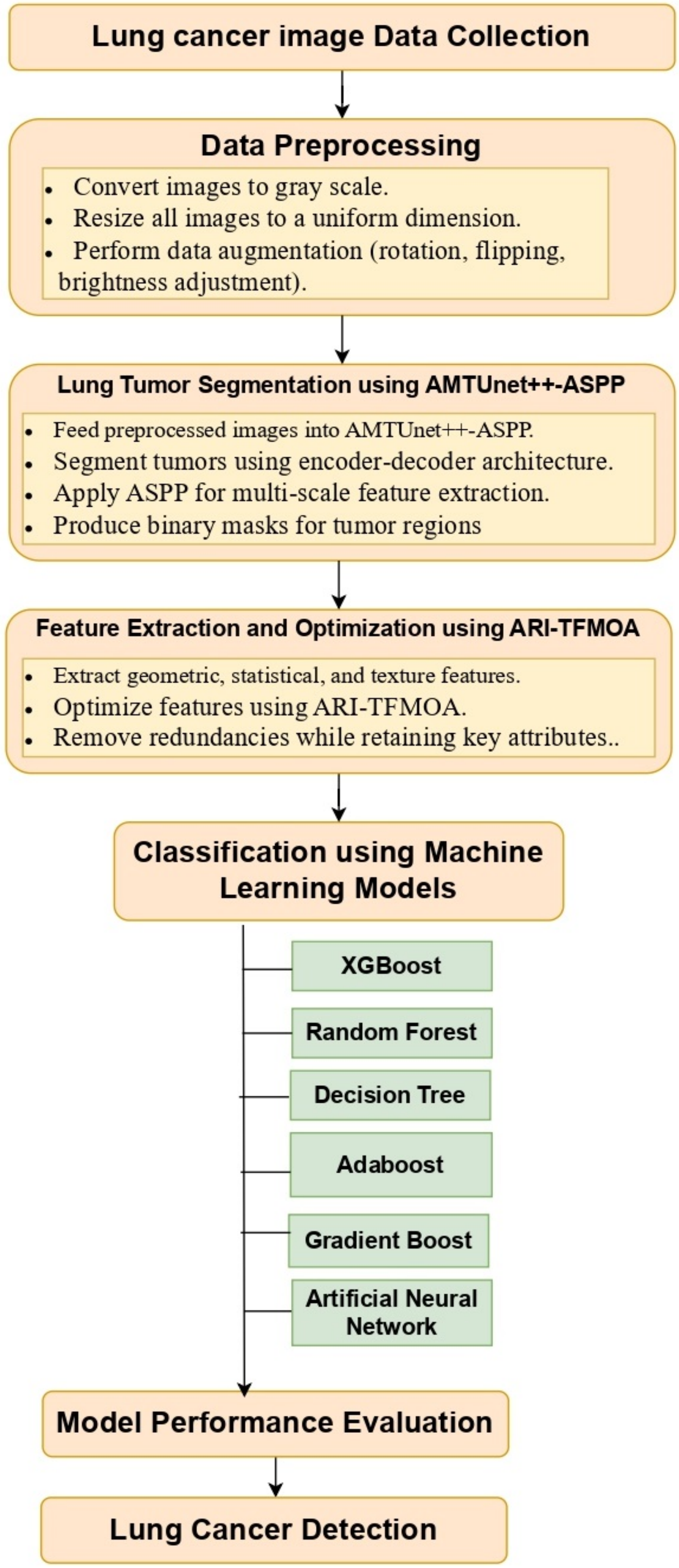

Figure 1: Proposed Method for lung cancer detection

Journal of Theoretical and Applied Information Technology
31st July 2025. Vol.103. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.1. Collection of Data

The dataset used in this study is extracted from the Kaggle lung cancer dataset [25] that comprises CT scan images classified into three classes: normal, benign, and malignant. The images for the dataset are selected to represent different patients so that different stages of lung cancer in various patients can be included in the dataset. By including these variations, a robust model used to identify cancer from normal lung issues can be developed

The CTScans were collected from specialized oncology centers and hospitals, where radiologists manually annotated the images to confirm the presence or not of lung cancer. A full scan is composed of many slices, each providing a different cross-sectional view of the lungs.

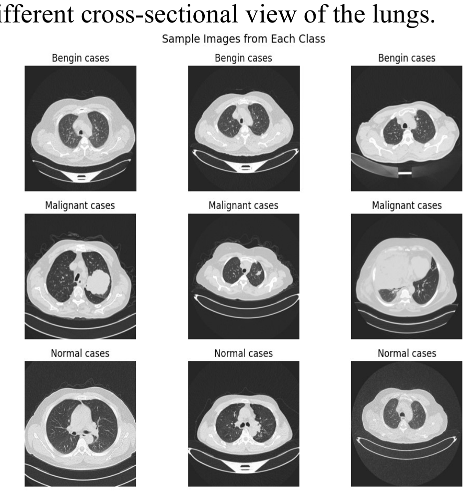

Figure 2: Sample images in the lung cancer dataset

It contains set of images stored in DICOM format which represents the ended images of human body with high resolution of detailed anatomical structures. In a realistic scenario, only a few CT scan slices could be holding the clear view of the lung nodules, hence only six slices are selected for this study. The sample images are shown in figure 2. Furthermore, the dataset is checked in terms of class balance, with the goal of having a balanced distribution of the three classes (normal, benign, malignant). Such a split is paramount to avoid classification bias and increase model generalization. To enhance the accuracy of the input images, the collected dataset is preprocessed before segmentation and classification.

3.2. Preprocessing

A complete preprocessing pipeline is implemented to improve the quality and robustness of the input CT scan images. All images are resized to a standard resolution of 224×224 pixels, confirming uniformity between the images processed in the dataset. Finally, the pixel intensity values are normalized between [0,1] so that model converges faster. Gaussian filtering is applied to eliminate noise and accentuate more of the object's structures, thus preventing the lung's nodules from getting damaged. Because CT images are often of low contrast quality, CLAHE is used to improve local contrast and thus make abnormalities more visible. Utilizing data augmentation techniques (e.g., rotation, Zooming) increases the variability of a training dataset, thus improving model generalization and reducing overfitting.

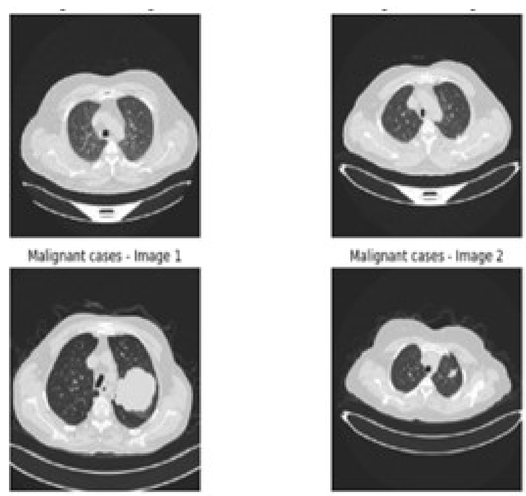

Figure 3: Sample preprocessed images

For example, we are generating segmentation ground truth masks based on radiologist annotations, which will help the AMTUnet++-ASPP model learn the accurate boundaries of the tumor. Figure 3 shows preprocessed images form three cases. Then, we split the dataset into training (80%) and test (20%) sets with an even distribution of normal, benign and malignant cases in each subset to ensure good evaluation of the model. Preprocessing steps were performed to prepare the dataset for segmentation and classification, improving the accuracy and robustness of proposed model.

4. EXPERIMENTS AND RESULTS

4.1 Lung Tumor Segmentation Using AMTUnet++-ASPP

Lung tumor segmentation is of the utmost importance as it separates the malignant and benign nodules from normal lung tissues. For high quality lung tumor segmentations, we adopt AMTUnet++-

ASPP, a state-of-the-art refinement of the original unet. AMTUnet++-ASPP applies Atrous Spatial Pyramid Pooling (ASPP) with multi-scale feature extraction ability to covers local and global contextual information, which allows it to overcome the limitation of the U-Net architecture applied in mesenchymal subgroup classification. This method enhances tumor-size versatility in detection while minimizing false positives.

The segmentation process of the AMTUnet++-ASPP begins when preprocessed CT scan images are introduced into the model; Inside its encoder network, multiple convolutional layers extract hierarchical spatial features. For high-resolution segmentation, deep skip connections are incorporated, allowing the model to preserve fine-grained information and alleviate information loss in the downsampling stage. Hence for some particular shapes of lung nodules as small, irregular shape which a typical U-shaped bottleneck layer cannot be recognized, therefore the ASPP module with dilation convolution operation has shown potential in segmentation process. The last decoder network reconstructs the segmented output to create an accurate tumor mask, allowing the isolation of the region of interest (ROI) from the background lung tissues.

Binary cross-entropy (BCE) loss is used in conjunction with Dice coefficient loss to jointly optimize the segmentation model ensuring both accuracy of pixel-wise classification and sharpness of tumor boundaries. Further morphological operations and connected component analysis, are applied to the segmented outputs to remove noise and artifacts. The segmented lung tumor masks are an important input for feature extraction and then are optimized to classify lung tumors.

4.2 Feature Optimization Using ARI-TFMOA

Optimizing features is an important step in enhancing classification performance by improving extracted features and discarding redundant information. Here we utilize ARI-TFMOA to extract effective tumor masksfeatures. In contrast to static feature selection approaches, the ARI-TFMOA method adapts the importance of features according to three objective functions, thus guaranteeing that the important features are the only ones influencing classification.A collection of deep features is extracted from the segmented images using descriptors related to statistics, texture, and form. The properties of the dataset are generalized geometric parameters like area, perimeter, eccentricity, and solidity as well as texture-based descriptors Gray Level Co-occurrence Matrix (GLCM) and Local Binary Pattern (LBP). High-dimensional feature sets pose a major challenge because they regularly include redundant or less relevant data, which results in overfitting and lowered classification accuracy.

To address this, we used a fuzzy decision-making mechanism in ARI-TFMOA to assign higher importance to those features which in turn can improve the classification accuracy. The dynamic radiated inertia adaptation mechanism of the algorithm guarantees the balance between exploration and exploitation, which provides a faster convergence velocity and higher stability of feature subspace. This increasingly repetitive optimization process assesses feature importance by maximizing classification accuracy subject to feature redundancy.Hence the reduced feature subset improves the performance of ML classifiers with a concise and useful feature presentation.setting. ARI-TFMOA effectively improves classification accuracy, precision and recall for brain tumor classification compared to models with raw extracted features. The optimized features chosen here are thus passed as inputs to the next phase of classification, thus minimizing the overhead of computational effort of ML classifiers whilst maximizing diagnostic accuracy.

4.3 Classification Using Machine Learning Models

Lung tumor segmentation is of utmost importance for accurate diagnosis. After extracting and optimizing features through AMTUnet++-ASPP segmentation and ARI-TFMOA optimization, the classification process is performed using six different models. These include five ML classifiers (Random Forest, XGBoost, Decision Tree, AdaBoost, GBC) and an advanced ANN model. These models were selected based on their ability to handle high-dimensional feature spaces and generalize effectively on test data.

The optimized feature set, derived from the segmented tumor regions, serves as the input to these classifiers. The classification task involves predicting whether each case is normal, benign, or malignant. Each classifier's hyperparameters were fine-tuned to achieve optimal performance. For example, Random Forest tuning focused on the number of trees, XGBoost and AdaBoost on learning rate and depth, and Decision Tree settings to reduce overfitting. The ANN was designed with appropriate hidden layers, activation functions, and

Journal of Theoretical and Applied Information Technology
31st July 2025. Vol.103. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

trained using backpropagation to maximize classification accuracy.

The results are shown in Table 1 and figure 4, demonstrate that the ANN outperformed all other models, achieving the highest values in all performance metrics. Specifically, ANN recorded an accuracy of 97%, precision of 97.5%, recall of 96.8%, and F1-score of 97.1%. XGBoost followed with an accuracy of 96%, while Random Forest achieved 95% accuracy. Decision Tree, AdaBoost, and Gradient Boosting also performed competitively but lower than ANN and XGBoost. It is clearly observed that ANN outperformed all ML models, achieving the highest accuracy of 97%, followed by XGBoost and Random Forest. The strong performance of ANN highlights its capability to capture complex patterns and relationships within the optimized features, thereby enhancing the classification performance significantly. In summary, the classification results confirm that the proposed hybrid model, combining DL-based segmentation, ARI-TFMOA feature optimization, and ANN-based classification, provides best accuracy. Compared to other classifiers, the ANN-based approach exhibited excellent precision and recall, making it highly suitable for critical tasks such as early lung cancer detection.

Table 1: Results with ML models

Model	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
Random Forest (RF)	95.0	95.2	94.8	95.0
XGBoost	96.0	96.3	95.8	96.0
Decision Tree (DT)	91.8	91.3	90.5	90.9
AdaBoost	94.0	93.8	93.2	93.5
Gradient Boosting (GBC)	95.0	94.5	94.2	94.3
ANN	97.0	97.5	96.8	97.1

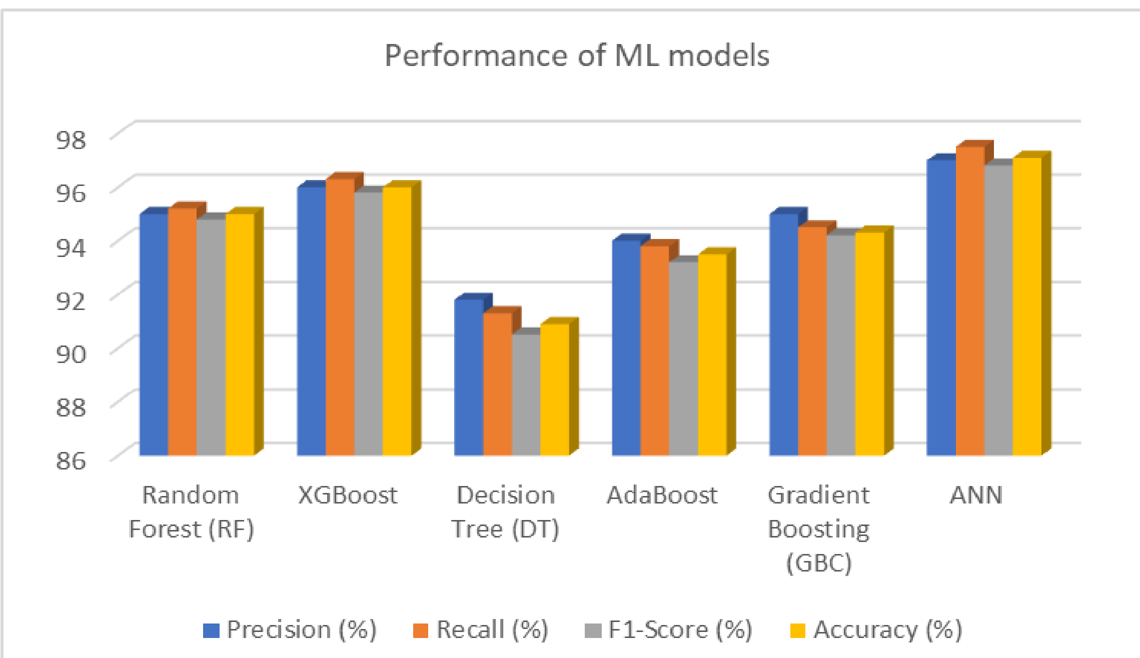

Figure 4: Performance of ML models

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2: Accuracy (%) of Optimization Methods across Batch Sizes

Batch Size	GWO-FS	PSO-FS	DE-FS	TFMOA-FS	ARI-TFMOA-FS
8	88.0	89.0	90.0	91.0	92.0
16	89.0	90.0	91.0	92.0	93.0
32	90.0	91.0	92.0	93.0	94.0
64	91.0	92.0	93.0	94.0	95.0
128	91.5	92.3	93.1	94.0	95.2

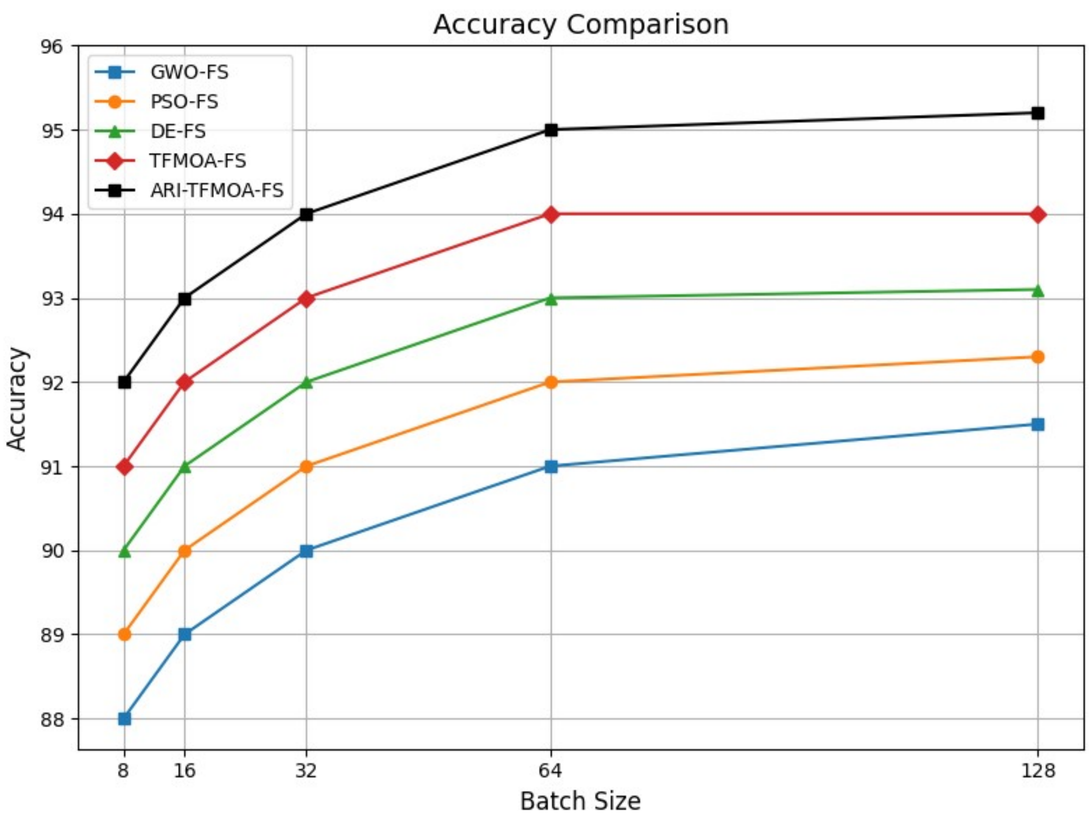

Figure 5: Accuracy trends of optimization methods across batch sizes

Table 3: Precision (%) of Optimization Methods across Batch Sizes

Batch Size	GWO-FS	PSO-FS	DE-FS	TFMOA-FS	ARI-TFMOA-FS
8	89.5	90.2	91.0	92.0	96.5
16	90.5	91.2	92.2	93.5	97.0
32	91.5	92.3	93.0	94.0	97.5
64	92.5	93.2	94.0	95.0	98.0
128	93.0	93.7	94.5	95.5	98.5

Table 4: Recall (%) of Optimization Methods across Batch Sizes

Batch Size	GWO-FS	PSO-FS	DE-FS	TFMOA-FS	ARI-TFMOA-FS
8	85.0	86.0	87.0	88.0	90.0
16	86.0	87.0	88.0	89.0	91.0
32	87.0	88.0	89.0	90.0	92.0
64	88.0	89.0	90.0	91.0	92.5
128	88.5	89.5	90.5	91.5	92.5

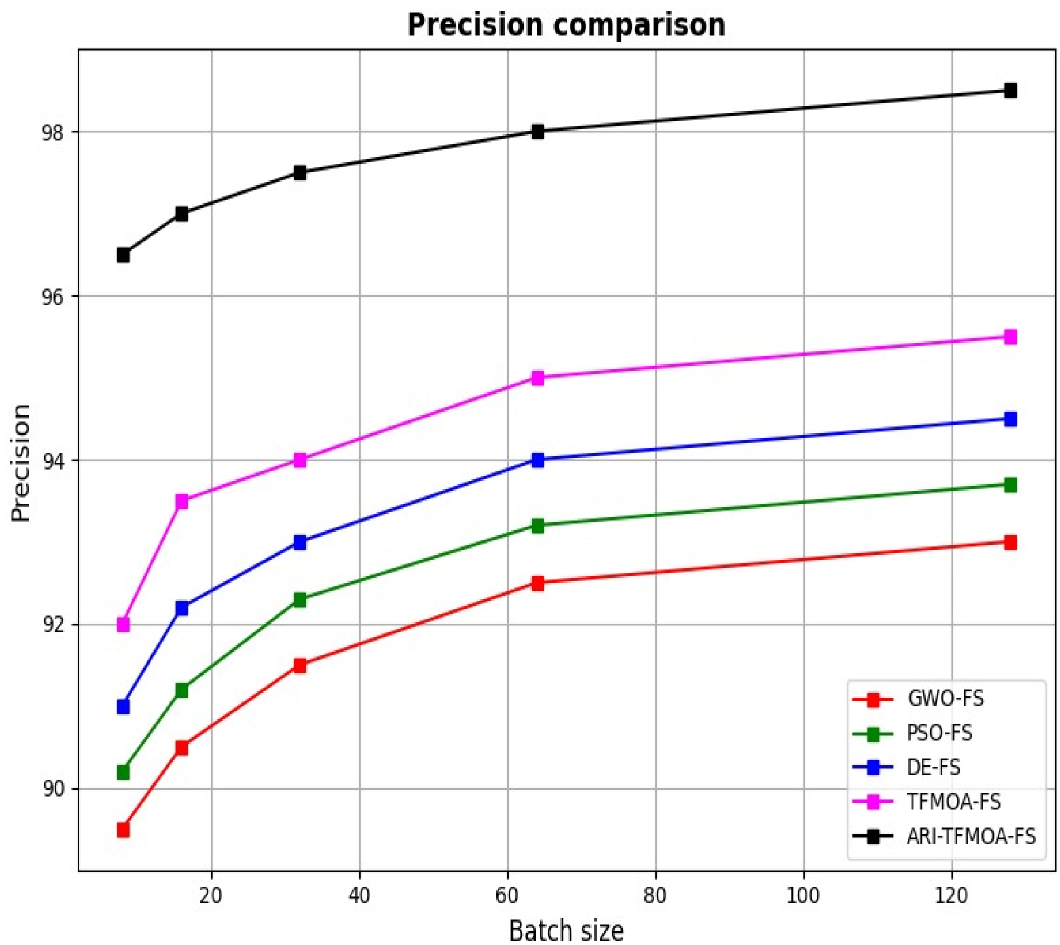

Figure 6: Precision trends of optimization methods across batch sizes

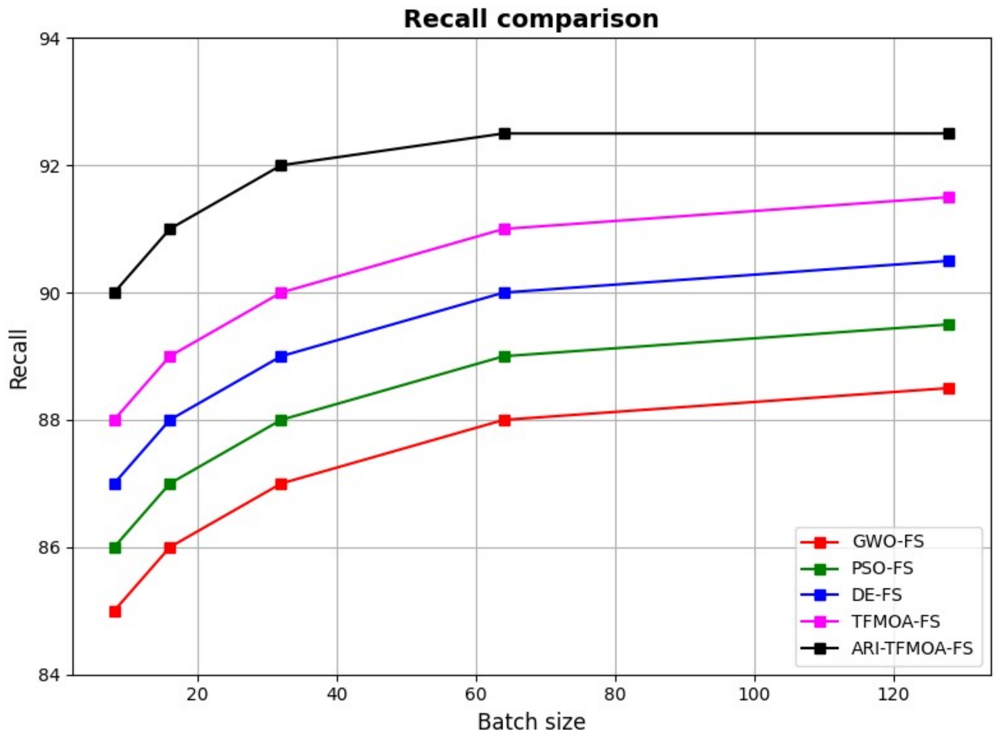

Figure 7: Recall trends of optimization methods across batch sizes.

4.4 Efficiency Evaluation of Optimization Techniques (Batch Size-Based Performance Trends)

To further assess the robustness and effectiveness of the proposed ARI-TFMOA-FS (Adaptive Radiant Inertia Tuned Fuzzy Multi-Objective Algorithm - Feature Selection) technique, a detailed comparative analysis was conducted using a batch size-based performance evaluation approach. This analysis was designed to investigate how different heuristic optimization methods influence the lung cancer detection model's classification performance under varying data processing scenarios. Although batch size change is typically associated with DL models, it is emulated here to simulate varying input volumes during the feature selection process within ML classification.

Five optimization techniques, namely GWO-FS (Grey Wolf Optimizer based Feature Selection), PSO-FS (Particle Swarm Optimization based Feature Selection), DE-FS (Differential Evolution based Feature Selection), TFMOA-FS (Tuned Fuzzy Multi-Objective Algorithm Feature Selection), and the proposed ARI-TFMOA-FS, were considered for this evaluation. The performance of these methods was analyzed across five batch sizes (8, 16, 32, 64, and 128) using three important metrics: Accuracy, Precision, and Recall.

Table 2 and Figure 5 show the accuracy (%) of optimization methods across batch sizes. The results indicate that ARI-TFMOA-FS consistently delivers higher accuracy at all batch sizes compared to the other methods. While all optimizers exhibit incremental improvements as batch size increases, ARI-TFMOA-FS achieves the highest accuracy of 95.2% at batch size 128, confirming its superior ability to select highly discriminative features.

Similarly, Table 3 and Figure 6 present the precision values (%) across batch sizes. The precision trends mirror those of accuracy. ARI-TFMOA-FS consistently outperforms other techniques, achieving a precision of 98.5% at batch size 128, which aligns well with the previously observed classification results. The significant precision advantage underscores its potential in reducing false positives, which is essential for accurate medical diagnostics.

The recall performance, illustrated in Table 4 and Figure 7, further validates the effectiveness of ARI-TFMOA-FS. At batch size 128, ARI-TFMOA-FS records a recall of 92.5%, demonstrating its strong capability to correctly identify positive cases and minimize the risk of missed diagnoses.

Overall, for all three-performance metrics ,ARI-TFMOA-FS consistently outperformed other heuristic optimization methods. While the alternative algorithms also showed performance improvements with increased batch sizes, the proposed ARI-TFMOA-FS technique maintained its position as the best-performing method throughout. These findings strongly suggest that ARI-TFMOA-FS not only ensures efficient feature selection but also significantly boosts classification performance, making it a highly promising and reliable optimization approach for lung cancer detection systems.

4.5 Comparison of proposed model with existing approaches

The comparison of proposed model with existing models is shown in Table 5. Proposed model achieved better accuracy 97% as compared to existing methods like UNET (94%) [4], Transfer Learning (95%) [7], Deep Ensemble 2D CNN (95%) [10]. It also outperformed both Transfer Learning (VER-Net) (91%) [8], and Ranking-Based Algorithms (95.5%) [14]. The results show that the performance is superior due to the combination of AMTUnet++-ASPP with ARI-TFMOA for lung tumor image segmentation and feature selection.

Table 5: Accuracy comparison

Method	Accuracy
UNET [4]	94%
Transfer Learning [7]	95%
Transfer Learning (VER-Net) [8]	91%
Deep Ensemble 2D CNN [10]	95%
Ranking based algorithms [14]	95.5%
Proposed Model	97%

In comparison to prior studies, the proposed model introduces a multi-stage hybrid framework that offers both depth and precision in lung cancer detection. Unlike conventional algorithms which only deal with either the segmentation or classification, advanced segmentation, optimization and classification (AMTUnet++-ASPP, ARI-TFMOA, and ML + ANN, respectively) were combined providing better interpretability and accuracy. The salient characteristics of this study are the well-picked features and outstanding classification performance. Nevertheless, such limitations, in particular, the dependence of the model on a single dataset and the relatively long training time required for ANN, are acknowledged, and may be solved in further research by validating cross-dataset and optimizing computations.

Different from the current state-of-the-art algorithms, our framework is the only approach in which segmentation, feature contraction and classification can be integrated throughout the whole procedure. In the previous works, the feature selectivity was either not designed optimally or the full end-to-end design was not realized. Having better precision and less feature redundancy, the study effectively achieves its goals and contributes to the development of AI-assisted lung cancer diagnosis.

5. CONCLUSION AND LIMITATIONS

Lung cancer remains one of the most challenging health concerns worldwide, necessitating early and precise detection to improve survival rates. This paper proposed an advanced lung cancer detection model integrating AMTUnet++-ASPP-based segmentation, ARI-TFMOA feature optimization, and hybrid classification using ML and artificial neural network (ANN) models to enhance diagnostic accuracy. The AMTUnet++-ASPP model successfully segmented lung tumors, preserving fine-grained structural details, while ARI-TFMOA refined feature selection, ensuring that only the most discriminative features were used for classification. The final classification was performed using six models, including five ML classifiers and ANN, among which ANN achieved the highest accuracy (97.0%), outperforming other models.

Through extensive evaluation—including accuracy, precision, recall, and F1-score—the proposed method demonstrated superior performance compared to traditional ML and DL-based approaches. These results validate the original research problem: that a hybrid AI model combining segmentation, feature optimization, and classification can address current gaps in accuracy, interpretability, and computational efficiency in lung cancer detection. The ANN's superior results particularly support the hypothesis that deep models, when coupled with refined features, can significantly enhance detection reliability. The

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

results indicate that the hybrid approach—combining deep learning-based segmentation, optimization-based feature selection, and ANN-driven classification—significantly enhances lung cancer detection accuracy. This makes the proposed model a viable and highly effective AI-driven solution for clinical decision support systems. However, there are certain limitations to be acknowledged. First, the model is trained and validated on a single publicly available dataset, which may limit its generalizability across diverse clinical settings. Second, the training time of the ANN model is relatively high compared to traditional ML models, which may impact real-time deployment in resource-constrained environments. Third, the segmentation process relies on ground truth masks derived from expert annotations, which may introduce subjectivity or variability if not standardized across datasets. These limitations highlight areas for future improvement, such as cross-dataset validation, computational optimization, and development of semi-supervised segmentation approaches. Future work may also explore the integration of transformer-based architectures to further enhance segmentation precision and feature representation.

REFERENCES:

[1] R. S. Raj, P. K. Madalu Palakshamurthy, and B. E. Rangaswamy, “Multi-class chronic lung disease classification based on guided backpropagation convolutional neural network using chest X-ray images,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 37, no. 2. Institute of Advanced Engineering and Science, p. 1328, Feb. 01, 2025. doi: 10.11591/ijeecs.v37.i2.pp1328-1338.

[2] Hossain, Md. I., Ahmad, S., Hasan, S., Rahim, Md. A., & Shin, J. (2024). Lung cancer detection and classification using LeNet-LSTM model on computed tomography images. In Multidisciplinary Science Journal (Vol. 7, Issue 4, p. 2025188). Malque Publishing. https://doi.org/10.31893/multiscience.2025188.

[3] Joselyn Zapata-Paulini et al., “Performance analysis of 10 machine learning models in lung cancer prediction,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 37, no. 2. Institute of Advanced Engineering and Science, p. 1352, Feb. 01, 2025. doi: http://doi.org/10.11591/ijeecs.v37.i2.pp1352-1364.

[4] B. K. Sathiyamurthy and V. K. Madhaiyan, “Automated lung cancer T-Stage detection and classification using improved U-Net model,” International Journal of Electrical and Computer Engineering (IJECE), vol. 14, no. 6. Institute of Advanced Engineering and Science, p. 6817, Dec. 01, 2024. doi: 10.11591/ijece.v14i6.pp6817-6826.

[5] A. Lakshmanarao, N. Gopal, N. Vullam, M. Sridhar, M. K. Kanth, and U. M. Rayudu, “Lung cancer detection using hybrid integration of autoencoder feature extraction and ML techniques,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 37, no. 1. Institute of Advanced Engineering and Science, p. 416, Jan. 01, 2025. doi: 10.11591/ijeecs.v37.i1.pp416-424.

[6] A. A.-R. Odeh and A. Mustafa, “Explaining transfer learning models for the detection of COVID-19 on X-ray lung images,” International Journal of Electrical and Computer Engineering (IJECE), vol. 14, no. 4. Institute of Advanced Engineering and Science, p. 4542, Aug. 01, 2024. doi: 10.11591/ijece.v14i4.pp4542-4550.

[7] K. R. Sekhar, G. R. L. M. Tayaru, A. K. Chakravarthy, B. Gopiraju, A. Lakshmanarao and T. V. S. Krishna, "An Efficient Lung Cancer Detection Model using Convnets and Residual Neural Networks," 2024 Fourth International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), Bhilai, India, 2024, pp. 1-5, doi: 10.1109/ICAECT60202.2024.10469187.

[8] A. Saha, S. M. Ganie, P. K. D. Pramanik, R. K. Yadav, S. Mallik, and Z. Zhao, “VER-Net: a hybrid transfer learning model for lung cancer detection using CT scan images,” BMC Medical Imaging, vol. 24, no. 1. Springer Science and Business Media LLC, May 24, 2024. doi: 10.1186/s12880-024-01238-z.

[9] R. Javed, T. Abbas, A. H. Khan, A. Daud, A. Bukhari, and R. Alharbey, “Deep learning for lungs cancer detection: a review,” Artificial Intelligence Review, vol. 57, no. 8. Springer Science and Business Media LLC, Jul. 08, 2024. doi: 10.1007/s10462-024-10807-1.

[10] A. A. Shah, H. A. M. Malik, A. Muhammad, A. Alourani, and Z. A. Butt, “Deep learning ensemble 2D CNN approach towards the detection of lung cancer,” Scientific Reports,

vol. 13, no. 1. Springer Science and Business Media LLC, Feb. 20, 2023. doi: 10.1038/s41598-023-29656-z.

[11] A. R. Wahab Sait, "Lung Cancer Detection Model Using Deep Learning Technique," Applied Sciences, vol. 13, no. 22. MDPI AG, p. 12510, Nov. 20, 2023. doi: 10.3390/app132212510.

[12] T. I. A. Mohamed, O. N. Oyelade, and A. E. Ezugwu, "Automatic detection and classification of lung cancer CT scans based on deep learning and ebola optimization search algorithm," PLOS ONE, vol. 18, no. 8. Public Library of Science (PLoS), p. e0285796, Aug. 17, 2023. doi: 10.1371/journal.pone.0285796.

[13] M. A. Thanoon, M. A. Zulkifley, M. A. A. Mohd Zainuri, and S. R. Abdani, "A Review of Deep Learning Techniques for Lung Cancer 13, no. 16. MDPI AG, p. 2617, Aug. 08, 2023. doi: 10.3390/diagnostics13162617.

[14] G. Kumari, U. V. Ramesh, P. Pattabhi Ramamohan, K. S. Pavani and A. Lakshmanarao, "Cancer Detection with Ensemble Learning Model from Novel Precedence based Algorithms," 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 2023, pp. 2187-2190, doi: 10.1109/IC3I59117.2023.10397629.

[15] Bharathy S et al,, "Lung Cancer Detection using Machine Learning," 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2022, pp. 539-543, doi: 10.1109/ICAAIC53929.2022.9793061.

[16] L. Wang, "Deep Learning Techniques to Diagnose Lung Cancer," Cancers, vol. 14, no. 22. MDPI AG, p. 5569, Nov. 13, 2022. doi: 10.3390/cancers14225569.

[17] R. Pandian, V. Vedanarayanan, D. N. S. Ravi Kumar, and R. Rajakumar, "Detection and classification of lung cancer using CNN and Google net," Measurement: Sensors, vol. 24. Elsevier BV, p. 100588, Dec. 2022. doi: 10.1016/j.measen.2022.100588.

[18] R. U. Osarogiagbon et al., "Lung Cancer Diagnosed Through Screening, Lung Nodule, and Neither Program: A Prospective Observational Study of the Detecting Early Lung Cancer (DELUGE) in the Mississippi Delta Cohort," Journal of Clinical Oncology, vol. 40, no. 19. American Society of Clinical Oncology (ASCO), pp. 2094–2105, Jul. 01, 2022. doi: 10.1200/jco.21.02496.

[19] P. Chaturvedi, A. Jhamb, M. Vanani, and V. Nemade, "Prediction and Classification of Lung Cancer Using Machine Learning Techniques," IOP Conference Series: Materials Science and Engineering, vol. 1099, no. 1. IOP Publishing, p. 012059, Mar. 01, 2021. doi: 10.1088/1757-899x/1099/1/012059.

[20] S. Lalitha, "An automated lung cancer detection system based Screening and Diagnosis Based on CT Images," Diagnostics, on machine learning algorithm," Journal of Intelligent & Fuzzy Systems, vol. 40, no. 4. SAGE Publications, pp. 6355–6364, Apr. 12, 2021. doi: 10.3233/jifs-189476.

[21] C. Jacobs et al., "Deep Learning for Lung Cancer Detection on Screening CT Scans: Results of a Large-Scale Public Competition and an Observer Study with 11 Radiologists," Radiology: Artificial Intelligence, vol. 3, no. 6. Radiological Society of North America (RSNA), Nov. 01, 2021. doi: 10.1148/ryai.2021210027.

[22] A. Y. Saleh, C. K. Chin, V. Penshie, and H. R. H. Al-Absi, "Lung cancer medical images classification using hybrid CNN-SVM," International Journal of Advances in Intelligent Informatics, vol. 7, no. 2. Universitas Ahmad Dahlan, Kampus 3, p. 151, Jul. 30, 2021. doi: 10.26555/ijain.v7i2.317..

[23] S. F. H. Abidi, Sumukhi T., Vinod Kumar H., and Santhosh B., "Lung Cancer Detection Using Deep Convolutional Neural Network," International Journal of Organizational and Collective Intelligence, vol. 11, no. 4. IGI Global, pp. 13–20, Oct. 2021. doi: 10.4018/ijoci.2021100102.

[24] Y. Guo et al., "Automated detection of lung cancer-caused metastasis by classifying scintigraphic images using convolutional neural network with residual connection and hybrid attention mechanism," Insights into Imaging, vol. 13, no. 1. Springer Science and Business Media LLC, Feb. 09, 2022. doi: 10.1186/s13244-022-01162-2.

[25] https://www.kaggle.com/datasets/hamdallak/the-iqothnccd-lung-cancer-dataset.