



Research Article

Lung Cancer Classification via Entropy-Driven Feature Selection and Deep Learning Architectures on Multi-Modal Imaging Data

Uzair Ishtiaq¹, Erma Rahayu Mohd Faizal Abdullah¹, Zubair Ishtiaque², Fatima Khan Nayer³, Salma Idris³, Amjad Rehman^{3,*}

¹Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, Universiti Malaya, Kuala Lumpur 50603, Malaysia

²Department of Analytical, Biopharmaceutical and Medical Sciences, Atlantic Technological University, H91 T8NW Galway, Ireland

³Artificial Intelligence & Data Analytics Lab (AIDA) CCIS Prince Sultan University, Riyadh, 11586, Saudi Arabia

*Corresponding author: arkhan@psu.edu.sa

Abstract: Lung cancer is a hazardous form of cancer found in humans caused by abnormal growth of cells in lung tissue. This type of cancer can be categorized into three stages, namely benign, malignant, and normal. Early diagnosis of the disease is essential as lung provide oxygen to the human body. Therefore, this study aimed to classify lung cancer using a hybrid deep learning (DL) model Improved by entropy-based feature selection. The model proposed a DL architecture, namely LungCFEx24, and a Transfer Learning (TL) based Visual Geometry Group-16 (VGG16) framework. Images captured during the analysis were pre-processed and feature extraction was performed. Following the process, discriminative features were selected using Fuzzy and Shannon entropy. The selected features were then fused to improve classification toughness. These fused features were fed into various kernel flavors of Support Vector Machine (SVM) classifiers. Among these, the Quadratic SVM achieved the highest accuracy of 95.92%, 98.60% sensitivity, and 99.84% specificity on a publicly available Iraq-Oncology Teaching Hospital/National Center for Cancer Diseases (IQ-OTH/NCCD) Lung Cancer Dataset containing Computed Tomography (CT) images. In addition to the discussion, the validation accuracy produced by the model was 97.05%. Rigorous testing on standardized datasets and comparisons with baseline methods showed the efficiency and effectiveness of the proposed hybrid model in lung cancer classification.

Keywords: Convolutional neural network; Feature selection method; Health risks; Lung cancer detection; Medical images classification

1. Introduction

Cancer is caused by the unusual growth of human body tissues, which eventually leads to death when not controlled. These tissues can grow abnormally in different parts of the human body causing different kinds of cancer, including, oral, oesophagus, stomach, breast, pancreas, and lung cancer (Fribert et al., 2013). Lung plays an essential role in the human respiratory system by expanding and then relaxing to provide oxygen to the human body. According to a survey by the Forum of International Respiratory Societies, around 1.4 million people died out of 10.4 million who were suffering from mild or severe respiratory diseases (Wang, 2022). Following the discussion, lung cancer is a common disease among other respiratory diseases specifically in elderly people. It is due

“This work was supported by Prince Sultan University, Riyadh Saudi Arabia, Year 2025”

<https://doi.org/10.14716/ijtech.v16i4.7646>

Received February 2025; Revised March 2025; Accepted May 2025

to several factors including, age, lifestyle, genetics, and gender. The instances of the disease are increasing at a rapid growth rate and are expected to reach 11.6% (Siegel et al., 2024). Therefore, timely and accurate diagnosis of lung cancer is essential for the health and the treatment plan of patients accordingly.

Automatic detection of lung cancer is a domain of interest among associates nowadays since it is related to healthcare. Systematized diagnosis of lung cancer has several benefits over manual detection by the clinician. These include reduced burden on the medical experts, less chance of human error, and detection of minute signs of lung cancer that are complex for the medical expert to diagnose. Recently, several associates have conducted promising studies in the automated lung cancer detection domain (Güraksin and Kayadibi, 2025). These studies usually collect data on the disease first and then apply some pre-processing methods, including normalization of image (Wankhade and Vigneshwari, 2023), contrast enhancement (Bhaskar et al., 2023), median filtering (Pradhan and Chawla, 2024), data augmentation (Jasmine et al., 2023), and resizing (Wahab Sait, 2023). After the pre-processing phase, the discriminative features are extracted from lung cancer images. These features are extracted manually through machine learning-based (ML) algorithms (Pham et al., 2025) or automatically through deep learning-based (DL) procedures (Wang et al., 2025). In ML, the manual features may include statistical features (Al-Absi et al., 2014), texture features (Bebas et al., 2021), color features (Rahane et al., 2018), shape and structure features (Hussain et al., 2024), textual features (Tey et al., 2023), as well as intensity features (Asuntha and Srinivasan, 2020). These extracted features are then fed to ML classifiers (Naghipour, 2024) for the final classification of lung cancer images into respective classes as normal, benign, or malignant cases. However, in DL process, features are automatically extracted by the model (novel or transfer learning) and lung cancer images are classified into different classes. The major challenge in lung cancer diagnosis is the extraction of salient features from the training images and then the grouping of the test images. This can be achieved by extracting features through DL models and classifying through ML algorithms. DL-based classification of lung cancer using CT (computed tomography) data addresses the critical clinical challenge as it improves the diagnostic accuracy and consistency. Moreover, it contributes to earlier detection of the disease and more effective, personalized treatment planning for patients.

In this study, a hybrid model is proposed for lung cancer classification. Before classification, some dataset pre-processing methods are performed to refine the available data. For instance, images are resized for suitable use for the model and then segmented for the region of interest. Images enhancement method is used to improve the quality of images. After that process, images are augmented to remove bias in the classes with fewer instances. Once the pre-processing step is complete, feature engineering on the processed images is performed. Feature engineering is conducted in three phases including, feature extraction, selection, and fusion. During the study, a novel DL model LungCFEx24 is designed and TL-based method is used by applying VGG16 model for feature extraction. These two models are used on IQ-OTH/NCCD - Lung Cancer Dataset (Kareem et al., 2021) which is a publicly available lung cancer dataset. The extracted features are selected and then fused to construct a Master Feature Vector (MFV), as it is input into ML classifiers for classification. The major contributions of this study are as follows:

- A hybrid method is proposed to improve automated lung cancer classification accuracy by leveraging a novel DL-based model, ML classifiers using CT images.
- The study introduces LungCFEx24 for CNN-based feature extraction and uses a TL-based VGG16 for extracting features from lung cancer images.
- A two-phased feature fusion and feature selection method are proposed as follows. (i) The extracted features from LungCFEx24 and VGG16 are selected using Fuzzy and Shannon Entropy. (ii) The selected features are fused to form the MFV, which is then fed to classification algorithms.
- This study is the first in lung cancer classification domain fusing automated novel DL-based features extracted through the proposed LungCFEx24 and TL-based features from a pre-trained VGG16, and classifying the fused features using ML algorithms, including five variants of SVM classifiers.

The remainder of this study is structured in the following way. Section 2 presents a review existing literature on lung cancer classification. Section 3 describes the proposed method, including details of the publicly available dataset used for the experiments, data pre-processing methods applied, and the feature engineering procedures developed for this study. Feature engineering process includes the novel CNN-based and TL-based feature extraction, followed by feature selection, and the fusion of features. Additionally, Section 4 describes the results obtained from different experiments in terms of several evaluation matrices, and the conclusion of the results is presented in Section 5.

2. Literature Review

During the past decade, associates from the medical imaging community have punctually participated in the domain of medical images diagnostics for the detection and classification of different medical abnormalities (Ishtiaq et al., 2020). The use of ML and DL models has produced promising results in this area of study. Moreover, the major areas of study where the associates are rapidly contributing the advanced machine vision and images processing algorithms in medical imaging include skin lesion detection (Akram et al., 2024), breast cancer detection (Sushanki et al., 2024), and brain tumor detection (Mathivanan et al., 2024). Other major areas also include diabetic retinopathy detection (Ishtiaq et al., 2023), acute ischemic stroke classification (Nurfirdausi et al., 2022), oral cancer detection (Kavyashree et al., 2024), external root resorption identification (Reduwan et al., 2024), epileptic seizure detection (Manjupriya and Leema, 2025), and lung cancer detection (Gayap and Akhloufi, 2024). Studies have proposed different methods for the detection and classification of lung cancer using its images.

Results by (Al-Yasriy et al., 2020) used a Transfer Learning (TL) based method by applying a pre-trained AlexNet architecture for classification of lung cancer images. The experiments were performed on the IQ-OTH/NCCD - Lung Cancer Dataset. During the analysis, dataset was split into 70% training as well as 30% testing images, and the best-achieved accuracy after 86 epochs was 93.548%. In (SR and Rajaguru, 2019), the study presented a probabilistic neural network method based on an improved crew-search algorithm for classification of the disease. The symptoms of lung cancer are typically diagnosed at its advanced phase. Therefore, predicting the presence of cancer using medical imaging methods at its early stage is essential to prevent the severity of the disease. In this study, an improved crew search algorithm-based feature selection for the early detection and classification of lung cancer was presented. The result showed that accuracy obtained by the improved crew-search algorithm was 90%.

In a similar result, (Yan and Razmjoo, 2023) normalized dataset by applying some initial pre-processing steps. After pre-processing, median filtering was applied to remove noise and make images smooth. The contrast of images was then improved by applying gamma correction. The number of instances in different classes was equalized using different data augmentation methods. Moreover, the study used CNN with an improved snake optimization method (Hashim and Hussien, 2022) for the training and testing of the proposed model. The accuracy obtained by this method with and without pre-processing steps were 96.58% and 89.67%. In another study, (Toğaçar et al., 2020) applied CNN equipped with Minimum Redundancy Maximum Relevance (MRMR) feature selection method for the detection and classification of lung cancer using chest CT images. During this process, a number of TL-based models were used, including LeNet, AlexNet, and VGG16 to detect lung nodules. These models were applied in the experimental setups for the salient features extraction and then classification of CT images. The outcome of the process signified that AlexNet model achieved a 98.74% maximum accuracy.

In another study (Mohamed et al., 2023), some pre-processing steps were applied to the IQ-OTH/NCCD - Lung Cancer Dataset, including, images resizing, grayscale conversion, gaussian filtering, segmentation, and normalization of images, erosion, noise removal, as well as wavelet transformation. The pre-processed images were then fed to the CNN which was later optimized using Ebola search optimization algorithm. This method produced promising results and the best

accuracy obtained was 93.21%. (Kareem et al., 2021) evaluated the performance of SVM classifier on the IQ-OTH/NCCD - Lung Cancer Dataset. First, the finding applied some pre-processing methods, including, bit-plane slicing, gaussian filtering, erosion operation, lung borders extraction, and defining lung areas. following the process, the result segmented cancer nodules using Otsu's threshold method (Xu et al., 2011). Gabor filter (Mehrotra et al., 1992) and Gray Level Co-occurrence Matrix (GLCM) based features (Mohanaiah et al., 2013) were extracted. Finally, images were classified using SVM classifier, and the accuracy achieved by the used model was 89.88%.

3. Study Protocol

This study assessed the effect of feature extraction and selection based on a novel CNN to classify lung cancer images on a publicly available dataset, namely, "IQ-OTH/NCCD - Lung Cancer Dataset" to investigate the results found in the literature (Kareem et al., 2021). In the first stage of this analysis, pre-processing of images were performed using different methods for better extraction of the salient features from images by CNN including resizing of images. The proposed images were resized into 256×256 for the network to process conveniently compared to the original size of 512×512 . Segmentation was then performed to discover the Region of Interest (ROI) from images. Since data in different classes was imbalanced, data augmentation methods were adopted to avoid bias in the network. This included rotation, random flipping, mirroring, scaling, and zooming of images. After data augmentation, images were improved using enhancement methods. Images were typically improved for CNN to better learn and extract the discriminative features. After data pre-processing, salient features were extracted using two DL-based models. The extracted features from these models were further selected using two entropy-based selection methods. These two selected feature vectors were then fused to construct an MFV, which was input into ML-based classifiers for classification. The block diagram of the proposed model as discussed in the study protocol was shown in Figure 1.

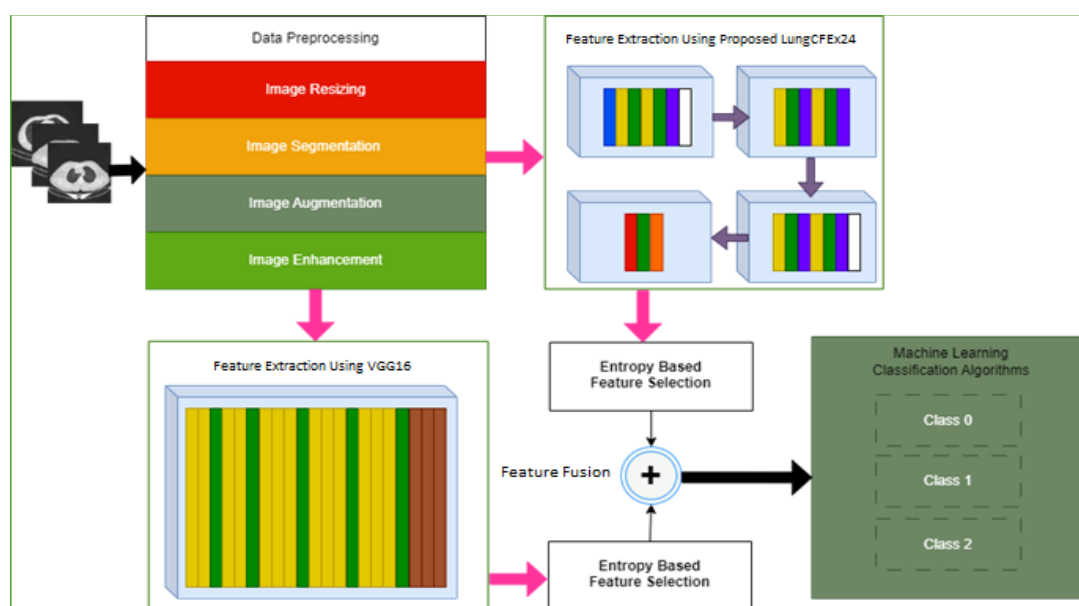


Figure 1 Proposed Model Block Diagram

The proposed method used during the analysis was described in the following phases.

Dataset: An online public dataset, IQ-OTH/NCCD - Lung Cancer Dataset was used for classification of lung cancer images into respective classes.

Data Pre-processing: The collected images of the disease were then preprocessed as pre-processing was an essential phase in lung cancer classification. The pre-processing methods used in this study were resizing of image, segmentation, augmentation, and enhancement.

Features Engineering: Discriminating features were then extracted and selected from the preprocessed images. Two methods were used for feature extraction, including, the novel LungCFEx24 and VGG16 for CNN-based features.

Features Selection and Fusion: After feature extraction, distinguishing features from LungCFEx24 and VGG16 were selected using Entropy algorithms and then fused to form MFV.

Classification and Evaluation: MFV was then fed to five kernel flavors of SVM for classification of lung cancer images into respective classes. Finally, these algorithms were evaluated on different evaluation matrices, namely, Sensitivity (SEN), Specificity (SPE), and Accuracy (ACC).

Figure 2 showed the schematic phases of the proposed model. Moreover, the details of the aforementioned phases were discussed in the subsequent sections.

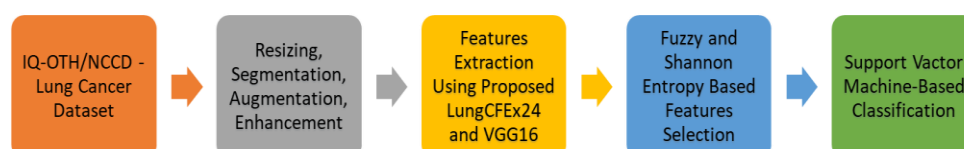


Figure 2 Schematic Phases of the Proposed Model

3.1. Dataset

The Iraq-Oncology Teaching Hospital/National Center for Cancer Diseases (IQ-OTH/NCCD) Lung Cancer dataset ([Kareem et al., 2021](#)) was initially collected in two specialist hospitals, namely, Iraq-Oncology Teaching Hospital and National Center for Cancer Diseases in 2019. A total of 977 CT scan images were collected from this dataset. Only the ground truths of the train class of dataset were available and not the same for the test class. Therefore, only images from the test class were used for the conduction of the experiments in this study. All images were available in Digital Imaging and Communication in Medicine (DICOM) format. The CT protocol included, 120 kV, and a slice thickness of 1 mm. The width of the window used during the analysis ranged from 350 to 1200 HU. In addition, window center from 50 to 600 HU were used for reading purposes. Sample images from benign, malignant, and normal cases were shown in Figure 3. The distribution of images in different classes of IQ-OTH/NCCD Lung Cancer Dataset was shown in Table 1. This study used the publicly available IQ-OTH/NCCD Lung Cancer Dataset, containing annotated CT scans including 977 images diagnosed as benign, malignant, and normal cases. Dataset was selected due to its comprehensive annotation, and suitability for evaluating diagnostic accuracy in lung cancer classification of images.

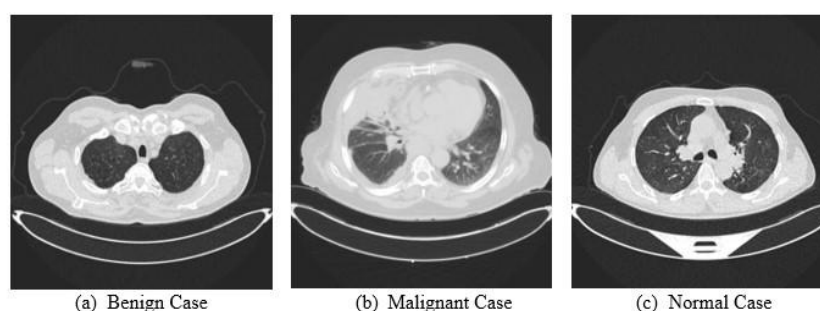


Figure 3 Sample Lung Cancer Images (a) Benign Case (b) Malignant Case (c) Normal Case

Table 1 Distribution of Images in IQ-OTH/NCCD Lung Cancer Dataset

Class Label	No. of Images	Percentage
0 – Benign Cases	120	12.29%
1 – Malignant Cases	561	57.43%
2 – Normal Cases	416	42.58%
Total	977	100%

3.2. Data Pre-processing

Data pre-processing was an essential phase in medical imaging as the unwanted data was not suitable for the network. However, clean lung images allowed the network to better learn the discriminative features, improving classification performance. The first phase in this study after data collection was data pre-processing. During the analysis, four pre-processing steps were incorporated into the model to improve the quality of lung cancer images. The original images in the IQ-OTH/NCCD Lung Cancer Dataset had different dimensions. Therefore, images were resized to 256×256 for standardization before use. Segmentation was performed to show the Region of Interest (ROI). Multiple data augmentation methods were applied to balance dataset as the IQ-OTH/NCCD Lung Cancer dataset was an imbalanced dataset and there might be a chance that the results could be biased. To remove this unfairness, rotation, random flipping, mirroring, scaling, as well as zooming in and zooming out of images were used as augmentation methods in this study. Images were enlarged to increase model strength and prevent overfitting. After augmentation, a total of 12000 images (4000 in each class) were collected as shown in Table 2.

During the process, dataset used was divided into training as well as test data, and the validation was performed using 5-fold cross-validation (CV). This division of the augmented dataset into training and test datasets took place by randomly selecting images to remove bias. Finally, images were improved for CNN to better learn and extract the discriminative features. The used enhancement methods in this study were the conversion of images into negative and thresholding.

Table 2 Distribution of Augmented Images in IQ-OTH/NCCD Lung Cancer Dataset

Class Label	Actual Images	Augmented Dataset	Training Dataset	Test Dataset
0 – Benign Cases	120	4000	2800	1200
1 – Malignant Cases	561	4000	2800	1200
2 – Normal Cases	416	4000	2800	1200
Total	977	12000	8400	3600

During the analysis, images negative was applied to convert the white into black and vice versa. This helped locate any salient feature present in any particular class. Moreover, adaptive thresholding was applied to obtain the higher components, such as details or edges present in images.

3.3. Feature Engineering

Features engineering was the most critical phase of the proposed model as this phase affected the performance of the recommended method. Moreover, DL-based features were extracted in this study to classify lung cancer images. The subsequent sections described features engineering phase of the proposed model.

3.3.1. DL-based Features

DL-based features were extracted through the novel CNN model, namely, LungCFEx24 and a TL-based pre-trained VGG16 model. This proposed method was named LungCFEx24 as there were 24 layers in this model for lung cancer feature extraction. It was trained on the 60% images of the augmented IQ-OTH/NCCD - Lung Cancer Dataset. The proposed LungCFEx24 used during the analysis was shown in Figure 4.

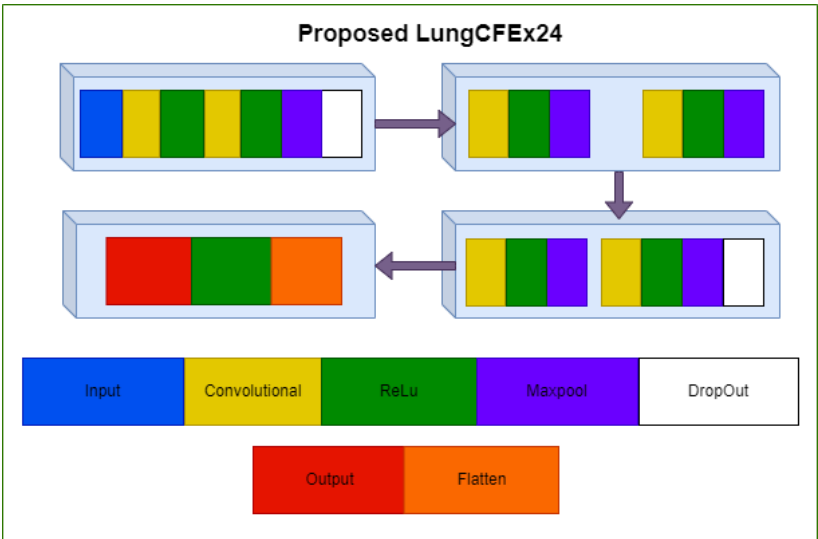


Figure 4 Architecture of Proposed Lung CFEx24 CNN Model

Two distinct types of deep CNN architectures were used for feature extraction. Consequently, two feature vectors of dimensions 8400×4096 , and 8400×2048 were obtained from LungCFEx24 as well as VGG16 CNN models. During the process, the parameters of deep neural networks were modified for optimal performance. LungCFEx24 model was trained using Adam optimizer, and the maximum number of epochs was 50 for training with a batch size of 96. In addition, L2 regularization of 0.001 was used and images were shuffled after every epoch. The learning rate was dropped by a factor of 0.1 during the process. In the proposed feature extraction LungCFEx24 model, the model extracted the discriminative features from the preprocessed lung images as shown in Figure 5. Additionally, the performance of the model was evaluated when it was increasing in five consecutive epochs. When the model was improving, then the discriminative features were extracted and MFV was constructed. However, when the model was not improving in five consecutive epochs, the hyperparameters of LungCFEx24 were modified to learn the discriminative features and MFV later was constructed.

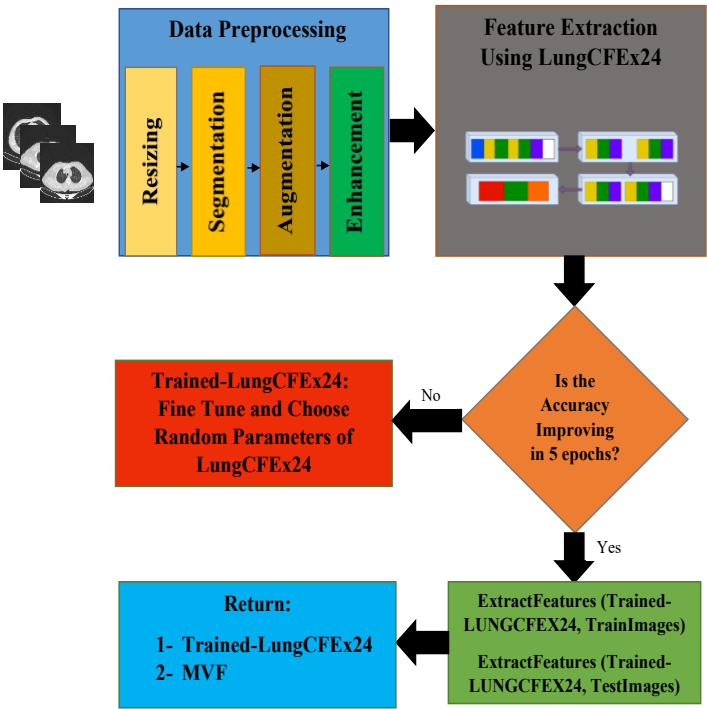


Figure 5 Flowchart of the Proposed LungCFEx24

3.3.2. Features Selection and Features Fusion

Fuzzy Entropy (Avci and Avci, 2009) and Shannon Entropy (Aljanabi et al., 2018) algorithms were used for feature selection. Both extracted vectors from LungCFEx24 and VGG16 CNN models were separately used for the calculation of Fuzzy as well as Shannon Entropy. In medical imaging, entropy played a crucial role in pattern recognition while analyzing the pixel arrangement (Versaci and Morabito, 2021). In the context of this study, the target function was defined using the average value obtained from the original vectors. The purpose of feature selection was to feed the ML algorithms with strong features, which were better than the mean features. Additionally, the process continued till the error rate of the ML algorithms was less than 0.1. The mathematical formulation for Shannon Entropy was shown using the following equation.

Where $occurrence_i^k$ represented the total number of occurrences of r_i in the class $Class_k$, and $r\#frequency_i^k$ signified the frequency of the r_i in the class $Class_k$:

$$r\#frequency_i^k = \frac{occurrence_i^k}{\sum_k occurrence_i^k} \quad (1)$$

The Shannon Entropy $SE(r_i)$ was mathematically established by the following equation.

$$SE(r_i) = - \sum_{k=1}^t (r\#frequency_i^k) \times (\log_2(r\#frequency_i^k)) \quad (2)$$

The mathematical formulation for Fuzzy Entropy was shown by the following equation.

$$FE(r_i) = - \sum_{i=0}^{N-1} \frac{r_i \times \mu(r_i)}{\sum_{i=0}^{N-1} r_i \times \mu(r_i)} \times \ln \left[\frac{r_i \times \mu(r_i)}{\sum_{i=0}^{N-1} r_i \times \mu(r_i)} \right] \quad (3)$$

Where N was the total gray levels, r_i represented the occurrences of gray levels at a given i , and μ signified the mean gray-level associated with r_i .

After selection of features, feature vectors $\mathfrak{F}_{LungCFEx24}$, and $\mathfrak{F}_{VGG16}$ of dimensions $8400 \times l$, as well as $8400 \times v$ for LungCFEx24 and VGG16 were obtained. Moreover, 'l' and 'v' represented the selected features obtained from $\mathfrak{F}_{LungCFEx24}$, as well as $\mathfrak{F}_{VGG16}$ for all the train images of dataset. After features selection step using the Shannon and Fuzzy Entropy algorithms, the selected attributes were fused to construct the final MFV, represented by $\mathcal{E}Final = \mathfrak{F}_{LungCFEx24} \mathfrak{F}_{VGG16}$, having the dimensions of size $[8400 \times (l + v)]$ was the final MFV supplied to the ML algorithms. Additionally, algorithm 1 showed the pseudocode of the proposed LungCFEx24 model.

Once the MFV was constructed, the final feature vectors $\mathcal{E}Final_{SE} = \mathfrak{F}_{LungCFEx24} \mathfrak{F}_{VGG16}$ and $\mathcal{E}Final_{FE} = \mathfrak{F}_{LungCFEx24} \mathfrak{F}_{VGG16}$ were constructed using Shannon Entropy and Fuzzy Entropy feature selection methods obtained from LungCFEx24 (with the dimension, $8400 \times l$) fused with the MFV through VGG16 (with the dimension, $8400 \times v$). Finally, $\mathcal{E}Final_{SE}$ and $\mathcal{E}Final_{FE}$ were input into the five kernel flavors of SVM which was an ML-based classification algorithm for final lung images classification. The kernels of SVM used in this study were linear, quadratic, fine, medium, and coarse Gaussian. The Linear SVM was a simple kernel flavor of SVM that drew a straight line or hyperplane in the original feature space to separate data. However, Quadratic SVM separated data with curved boundaries using a two-degree polynomial kernel. Fine Gaussian SVM was another kernel of SVM that used a narrow kernel function for flexible decision boundaries. Consequently, Medium Gaussian SVM used a moderately wide kernel function to balance flexibility. Coarse Gaussian SVM applied a wider kernel function that was used to form more generalized decision boundaries.

```

1  Input: DR_TrPath, DR_TsPath
2  Output: Trained-LungCFEx24, MFV
3  Procedure: TrainLungCFEx24 (path)
4    [TrainImage, TrainLabels] ← Load Training Images from Lung_TrPath
5    [TestImage, TestLabels] ← Load Testing Images from Lung_TsPath
6    TrainImages ← Resize(TrainImages, [256×256])
7    TestImages ← Resize(TestImages, [256×256])
8    TrainImages ← Segment(TrainImages, RoI)
9    TestImages ← Segment(TestImages, RoI)
10   TrainImages ← Augment(TrainImages, [Rotation, Flip, Mirror, Scale, Zoom-in, Zoom-out])
11   TrainImages ← ImageNegative(TrainImages)
12   TestImages ← ImageNegative (TestImages)
13   TrainImages ← AdaptiveThresholding(TrainImages)
14   TestImages ← AdaptiveThresholding (TestImages)
15   Repeat
16     Trained-LUNGCFEX24 ← Fine tune and choose random parameters
17     Trained-LUNGCFEX24 ← Train Trained-LUNGCFEX24 using TrainImages
18     Stop training if accuracy is not improving in consecutive five epochs
19     PredictedLabels ← Predict(Trained-LUNGCFEX24, TestImages, TestLabels)
20     ConfMat ← Confusion_Matrix(TestLabels, PredictedLabels)
21     Calculate accuracy
22   Until accuracy is maximum // (>=0 && <=100)
23   TrFeatures ← ExtractFeatures(Trained-LUNGCFEX24, TrainImages)
24   TsFeatures ← ExtractFeatures(Trained-LUNGCFEX24, TestImages)
25   MFV ← Save(TrFeatures, TrLabels, TsFeatures, TsLabels)
26   Return Trained-LUNGCFEX24
27   Return MFV
28 END Procedure

```

A total of five flavors of ML-based method, namely SVM was used in the proposed hybrid model to classify lung cancer images into respective classes.. The subsequent sections contained a comprehensive discussion of different experimental setups, dataset used, several performance measures applied to conduct these experiments, and a detailed analysis of the results achieved through the tests.

During the process of this study, classification of lung cancer was performed. The “IQ-OTH/NCCD - Lung Cancer Dataset” (Kareem et al., 2021) was used to conduct the experiments. This dataset contained 12000 augmented lung cancer images divided into three classes. For instance, the classes contained images for Benign (4000), Malignant (4000), and Normal Class (4000), which were marked as 0, 1, and 2, respectively. A total of 8400 images from a total of 12000 images were used to train the hybrid model consisting of the novel LungCFEx24 model and VGG16 model. The rest of the 3600 images were then used to validate the proposed hybrid model (containing 1200 augmented images per class). In the literature, different ratios of the training set and testing set were found. For example, different results used 80:20, 70:30, 60:40, and 50:50 to train and test

images. In line with this process, 70:30 was considered to be an optimum ratio and was often adopted by DL models [34]. Based on the literature, the same ratio was exploited to conduct the experiments, where 70% (8400) of the total images (12000) were used for training and the rest of the 30% (3600) were applied to test the proposed hybrid model.

4.2. Performance Evaluation

The final MFV $\mathcal{E}_{Final} = \mathfrak{F}_{LungCFEx24} \mathfrak{F}_{VGG16}$, was used to evaluate the performance of the proposed model. The performance metrics applied during the analysis were SPE, ACC, and SEN. Following the discussion, these metrics were shown mathematically as follows.

$$SEN = TP / (TP + FN) \quad (4)$$

$$SPE = TN / (FP + TN) \quad (5)$$

$$ACC = (TP + TN) / (TP + TN + FP + FN) \quad (6)$$

Where, TP = True Positive, FN = False Negative, TN = True Negative, and FP = False Positive.

4.3. Experimental Setup

Different experimental setups for the proposed model were discussed in this section. The study was categorized depending on selection of entropy for the final fused master features vector. In the first experimental setup, the final features vector consisted of features selection based on Shannon Entropy $\mathcal{E}_{Final_{SE}} = \mathfrak{F}_{LungCFEx24} \mathfrak{F}_{VGG16}$ after feature extraction using the proposed LungCFEx24 and VGG16 from the pre-processed images. However, the rest of the settings of the model are intact, and the final features vector consisted of features selection based on Fuzzy Entropy $\mathcal{E}_{Final_{FE}} = \mathfrak{F}_{LungCFEx24} \mathfrak{F}_{VGG16}$ in the second experimental setup. These experimental setups for classification of lung cancer images were performed using a system having 3.40 GHz (gigahertz) processor and 16 GB RAM (Random Access Memory). The software used for the conduction of the experiments was Python 3.8.0. Furthermore, the subsequent sections discussed the experimental setups for classification of lung cancer images.

4.3.1. Experimental Setup 1: Classification Results using $\mathcal{E}_{Final_{SE}} = \mathfrak{F}_{LungCFEx24} \mathfrak{F}_{VGG16}$

In this experiment, 12000 images from the IQ-OTH/NCCD - Lung Cancer Dataset were used to form the fused feature vector using Shannon Entropy method. A feature vector of 8400×1 was obtained from the proposed LungCFEx24 model. Another feature vector of $8400 \times v$ was extracted from the VGG16 model. After extracting these two feature vectors, Shannon Entropy method was used for the final feature selection. Following features selection, $\mathfrak{F}_{LungCFEx24}$, and $\mathfrak{F}_{VGG16}$ of dimensions 8400×1 , as well as $8400 \times v$ were obtained. A total of five kernels of SVM were the ML-based classification algorithms that were used for grouping lung cancer images.

During the analysis, the fused feature vector $\mathcal{E}(8400 \times [1 + v])$ was used for final classification. It was obtained after fusing the selected features represented by $\mathcal{E}_{Final_{SE}} = \mathfrak{F}_{LungCFEx24} \mathfrak{F}_{VGG16}$. Moreover, the proposed model algorithms used in the experiment were trained with 50 epochs. These selected feature vectors were then fed to the five kernel flavors of SVM classifiers for the evaluation of the proposed lung cancer images classification model. The results obtained for the grouping of images using SVM classifiers were shown in Table 1. Relating to this, the training and validation accuracies were shown in Figure 6.

4.3.2. Experimental Setup 2: Classification Results using $\mathcal{E}_{Final_{FE}} = \mathfrak{F}_{LungCFEx24} \mathfrak{F}_{VGG16}$

In the second experimental setup, 12000 images from the IQ-OTH/NCCD - Lung Cancer Dataset were applied to form the fused feature vector using Fuzzy Entropy method. A feature vector of 8400×1 was extracted from LungCFEx24 model. Additionally, feature vector of $8400 \times v$ was obtained from the VGG16 model. During the experiment, Fuzzy Entropy method was used for the final feature selection. After selection of features, the obtained feature vectors $\mathfrak{F}_{LungCFEx24}$, and $\mathfrak{F}_{VGG16}$ of

dimensions $8400 \times l$, as well as $8400 \times v$ were then fed to five kernels of SVM for classification of lung cancer images.

Table 1 Results of Experimental Setup 1 using Five Kernels of SVM Classifiers

	CLASSIFIER	CLASS	Valid Acc (%)	Train Acc (%)	SEN (%)	SPE (%)
SVM	Linear	0	93.21	93.49	93.11	99.36
		1			92.42	98.83
		2			97.63	98.17
	Quadratic	0	95.92	97.05	96.65	99.93
		1			99.35	99.80
		2			99.81	99.81
	Fine Gaussian	0	88.54	95.03	84.32	85.18
		1			86.63	85.47
		2			89.83	86.62
	Medium Gaussian	0	91.15	96.74	92.87	92.34
		1			95.35	93.45
		2			97.26	97.89
	Coarse Gaussian	0	93.92	95.65	92.77	95.32
		1			93.83	95.91
		2			96.86	97.29

The fused feature vector $\mathfrak{E}(8400 \times [l + v])$ was obtained after fusing the selected features represented by $\mathfrak{E}Final_{FE} = \mathfrak{F}_{LungCFEx24} \mathfrak{F}_{VGG16}$. Moreover, the proposed model algorithms were once more trained with 50 epochs. These selected final feature vectors with Fuzzy Entropy were input into the kernel flavors of SVM grouping algorithms for the evaluation of classification model. The results of SVM classifiers during the analysis were shown in Table 2.

4.3.3. Novelty of Work

This study proposed a hybrid method to improve automated lung cancer classification accuracy by using a novel DL-based model with ML classifiers. The proposed LungCFEx24 and VGG16 extracted CNN-based feature extraction from the pre-processed lung cancer images. Moreover, a two-phased feature fusion and feature selection method were also proposed. The extracted features were selected using Fuzzy as well as Shannon Entropy and then were fused to form MFV, which was later fed to SVM classification algorithms. Following this discussion, the proposed hybrid model was the first in lung cancer classification domain that focussed on fusing automated novel DL-based features extracted through the proposed LungCFEx24 as well as VGG16 and grouped the fused features using SVM classification algorithms.

Table 2 Results of Experimental Setup 2 using Five Kernels of SVM Classifiers

	CLASSIFIER	CLASS	Valid Acc (%)	Train Acc (%)	SEN (%)	SPE (%)
SVM	Linear	0	93.16	94.35	93.26	98.67
		1			92.35	98.13
		2			96.41	97.04
	Quadratic	0	94.45	96.21	97.21	98.87
		1			97.27	98.70
		2			98.88	98.04
	Fine Gaussian	0	83.67	92.88	82.24	84.22
		1			84.75	83.15
		2			87.93	86.55
	Medium Gaussian	0	90.07	95.91	92.12	92.82
		1			94.42	94.27
		2			96.85	96.53
	Coarse Gaussian	0	92.91	93.18	91.32	94.41
		1			93.59	95.65
		2			95.75	97.69

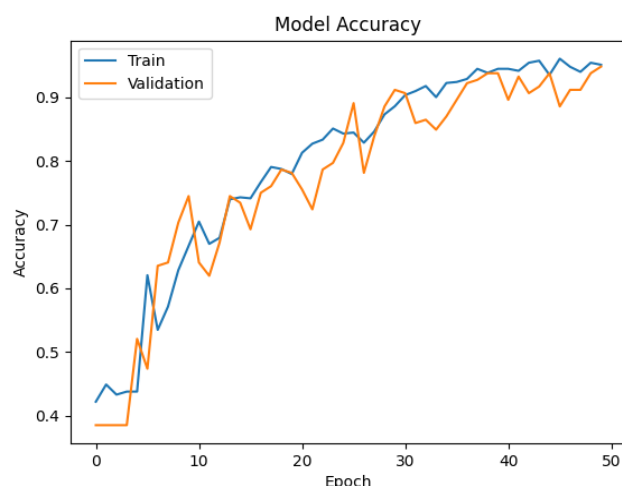


Figure 6 Training and Validation Accuracies of the Proposed Model

5. Comparison of The Proposed Model and Its Quantitative Analysis

The comparison of the proposed hybrid model with current state-of-the-art models in terms of several performance evaluation matrices was discussed in the following segments. Additionally, the quantitative analysis of the experiments conducted in this study was presented in the subsequent sections.

5.1. Comparison of Proposed Model with State-of-the-Art Methods

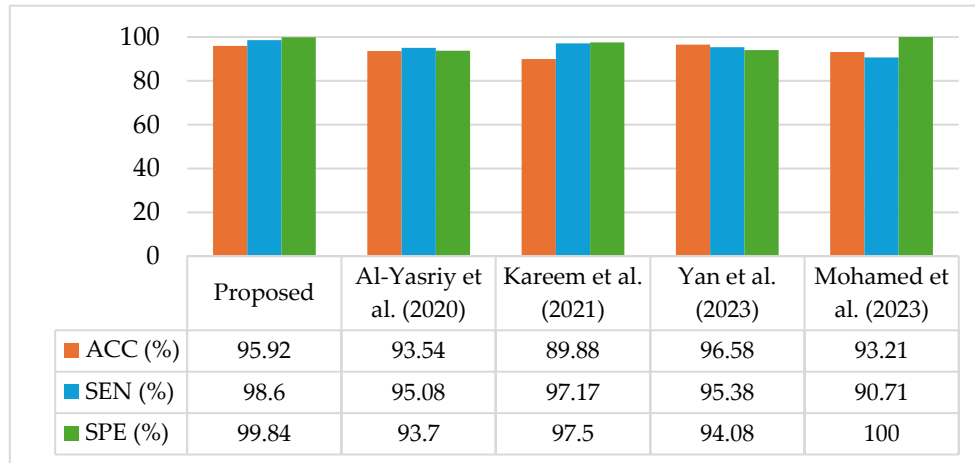
During the analysis, the proposed model was compared with the existing state-of-the-art methods in terms of ACC, SEN, and SPE. A detailed comparison of the proposed model was shown in Table 3. (Al-Yasriy et al., 2020) applied a TL-based method through a pre-trained AlexNet architecture for classification of lung cancer images while using the IQ-OTH/NCCD - Lung Cancer Dataset. The result split dataset into 70% training as well as 30% testing images and the best achieved accuracy after 86 epochs was 93.548%. Following this discussion, (Kareem et al., 2021) evaluated the performance of SVM classifier on the IQ-OTH/NCCD - Lung Cancer Dataset. After applying some pre-processing methods and segmenting cancer nodules using a threshold process, GLCM based features were extracted. Finally, images were grouped using SVM classifier with the accuracy of 89.88%.

In another study, (Yan and Razmjoooy, 2023) normalized dataset by applying some initial pre-processing steps. The result then applied median filtering, contrast enhancement, and data augmentation. Finally, the analysis used CNN with an improved snake optimization method and achieved an accuracy of 96.58% with pre-processing as well as 89.67% without pre-processing steps. The accuracy of the result with extensive pre-processing steps and computationally expensive was slightly higher than the method used in this study. The rationale for this marginal difference could be the number of classes. Classification in the method (Yan and Razmjoooy, 2023) used was performed on two classes, including, lung cancer and normal cases. However, the actual dataset contained three classes, consisting of benign, malignant, and normal cases. (Mohamed et al., 2023) fed the pre-processed images to the Ebola search-optimized CNN. The method used achieved satisfactory results and the accuracy was 93.21%.

The proposed model outperformed the state-of-the-art models in terms of different classification performance measures. The model achieved an accuracy of 95.92%, 98.60% sensitivity, and 99.84% specificity. The comparison of the proposed model with the state-of-the-art models was shown in Table 3. Moreover, the graphical representation of the comparison of the recommended method with the models mentioned earlier was shown in Figure 7. The result could be analyzed that the proposed method achieved promising lung cancer classification results compared to the current state-of-the-art study studies.

Table 3 Comparison of Proposed Model with State-of-the-Art Methods

Ref.	Year	Classes	Performance Measures		
			ACC (%)	SEN (%)	SPE(%)
(Al-Yasriy et al., 2020)	2020	3	93.54	95.08	93.70
(Kareem et al., 2021)	2021	3	89.88	97.14	97.50
(Yan and Razmjoooy, 2023)	2023	2	96.58	95.38	94.08
(Mohamed et al., 2023)	2023	3	93.21	90.71	100.00
Proposed		3	95.92	98.60	99.84

**Figure 7** Analysis of the Proposed Method Compared to Current State-of-the-Art Methods

5.2. Quantitative Analysis of Proposed Method's Average Performance

This section discussed the quantitative analysis of the experiments performed to show the average results. Table 4 compared the results of all kernel flavors of SVM classifiers using 5-fold CV where Shannon Entropy was applied as feature selection method. For the training of proposed algorithms, 50 epochs were used during the analysis. The results showed that the Quadratic SVM had the maximum classification ACC of 95.92%. Accuracy might not be the only proof for the validation of the proposed method. Therefore, some other performance measures were also used for the validation of the proposed method, including, SEN and SPE. The proposed model with the Quadratic SVM achieved 98.60% Sensitivity and 99.84% Specificity. However, the Fine Gaussian SVM achieved the poorest average accuracy of 88.54%.

Table 5 showed the comparison of the results concerning different flavors of SVM classifiers using 5-fold CV where Fuzzy Entropy was used as feature selection method. For these experiments, 50 epochs were used for the training of proposed algorithms. The results showed that the Quadratic SVM had the maximum classification ACC of 94.45%. In addition, the proposed model using the Quadratic SVM achieved 97.78% SEN and 98.53% SPE. The Fine Gaussian SVM achieved the lowest average accuracy compared to the other flavors of SVM.

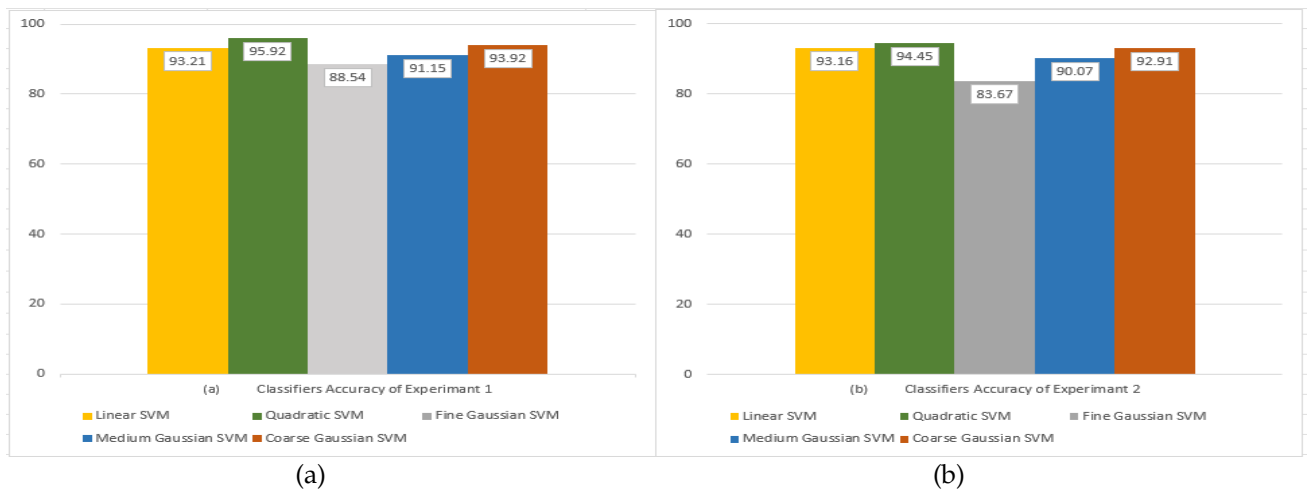
Table 4 Average Quantitative Results of Experiment 1 for $\mathcal{CFinal}_{SE}$.

		$\mathcal{F}_{VGG16}: 8400 \times 1000$ $\mathcal{F}_{LungCFEx24}: 8400 \times 1000$		
		ACC (%)	SEN (%)	SPE (%)
SVM	Linear	93.21	94.38	98.78
	Quadratic	95.92	98.60	99.84
	Fine Gaussian	88.54	86.92	85.75
	Medium Gaussian	91.15	95.16	94.56
	Coarse Gaussian	93.92	94.48	96.17

Table 5 Average Quantitative Results of Experiment 2 for $\mathcal{CFinal}_{FE}$.

		$\mathfrak{F}_{VGG16}: 8400 \times 1000$		
		$\mathfrak{F}_{LungCFEx24}: 8400 \times 1000$		
		ACC (%)	SEN (%)	SPE (%)
SVM	Linear	93.16	94.01	97.94
	Quadratic	94.45	97.78	98.53
	Fine Gaussian	83.67	84.97	84.64
	Medium Gaussian	90.07	94.46	94.54
	Coarse Gaussian	92.91	93.55	95.81

Figure (a), (b) showed a graphical comparison concerning different flavors of SVM classification algorithms for the results of the conducted Experiment 1 and 2 in classification accuracies. During the analysis concerning the performances of classification algorithms according to the achieved ACC, SEN, and SPE, Quadratic SVM was the best in all the experiments. In addition to this result, the Quadratic SVM outperformed other SVM classifiers by achieving an accuracy of 95.92% using $\mathfrak{F}_{VGG16}: 8400 \times 500$, and $\mathfrak{F}_{LungCFEx24}: 8400 \times 500$ with 5-fold cross-validation in Experiment 1.

**Figure 8** Graphical Comparison of SVM Performance: (a) Experiment 1; (b) Experiment 2

6. Conclusion and Future Work

In conclusion, this study proposed a novel CNN architecture and used a TL method for lung cancer classification into benign, malignant, and normal cases. The study aimed to offer medical professionals an additional perspective in diagnosing lung cancer. The result showed the potential of DL for the detection and classification of lung cancer into its respective classes. During the analysis, a hybrid model that used a publicly available IQ-OTH/NCCD - Lung Cancer Dataset containing CT lung images was presented. Dataset was initially pre-processed for performing feature extraction through the proposed LungCFEx24 and VGG16 architectures. In the experiment, LungCFEx24 which was a TL-based pre-trained architecture including CNN and VGG16 was used to form the hybrid model. Entropy-driven selection using Fuzzy as well as Shannon Entropy, and then fusing discriminative DL features was the basis of the proposed model. These selected and fused MFV was input into the ML-based classifiers for the final grouping. The result of the process showed that the hybrid model outperformed the baseline methods, achieving a high accuracy of 95.92%, 98.60% sensitivity, and 99.84% specificity. The current scope of the proposed hybrid model was limited to the detection and classification of lung cancer images obtained from the used dataset. In the future, studies should propose a dense network model leveraging multiple DL methods to simplify the complexities included in feature extraction process. Additionally, classification

accuracy of the model could be further improved by optimizing the extracted and fused features through advanced optimization algorithms, such as the adaptive sine cosine algorithm.

Acknowledgements

The authors are grateful to Prince Sultan University, Riyadh Saudi Arabia for supporting Article Processing Charges (APC) of this publication.

Author Contributions

Conceptualization, Uzair Ishtiaq and Erma Rahayu Mohd Faizal Abdullah; **methodology**, Uzair Ishtiaq and Zubair Ishtiaque; **software**, Uzair Ishtiaq and Amjad Rehman; **validation**, Erma Rahayu Mohd Faizal Abdullah, Noor Ayesha, Fatima Khan Nayer; **formal analysis**, Zubair Ishtiaque, and Amjad Rehman; **investigation**, Erma Rahayu Mohd Faizal Abdullah, Noor Ayesha, and Fatima Khan Nayer; **resources**, Uzair Ishtiaq and Uzair Ishtiaq and Amjad Rehman; **writing—original draft preparation**, Uzair Ishtiaq and Erma Rahayu Mohd Faizal Abdullah; **writing—review and editing**, Zubair Ishtiaque and Noor Ayesha; **supervision**, Erma Rahayu Mohd Faizal Abdullah; **project administration**, Uzair Ishtiaq. All authors have read and agreed to the published version of the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Materials

Information used for conducting this study was acquired from the publicly available IQ-OTH/NCCD Lung Cancer Dataset (Kareem et al., 2021).

References

- Akram, T, Khan, MA, Sharif, M & Yasmin, M 2024, 'Skin lesion segmentation and recognition using multichannel saliency estimation and M-SVM on selected serially fused features'. *Journal of Ambient Intelligence and Humanized Computing*, pp. 1083–1102, <https://doi.org/10.1007/s12652-018-1051-5>
- Al-Absi, HRH, Belhaouari, SB, & Sulaiman, S 2014, 'A computer aided diagnosis system for lung cancer based on statistical and machine learning techniques', *Journal of Computers*, vol. 9, no. 2, pp. 425–431, <http://dx.doi.org/10.4304/jcp.9.2.425-431>
- Aljanabi, MA, Hussain, ZM & Lu, SF 2018, 'An entropy-histogram method for image similarity and face recognition', *Mathematical Problems in Engineering*, no. 1, article 9801308, <https://doi.org/10.1155/2018/9801308>
- Al-Yasriy, HF, Al-Husieny, MS, Mohsen, FY, Khalil, EA & Hassan, ZS 2020, 'Diagnosis of lung cancer based on CT scans using CNN', Paper presented at the IOP conference series: materials science and engineering
- Asuntha, A & Srinivasan, A 2020, 'Deep learning for lung Cancer detection and classification', *Multimedia Tools and Applications*, vol. 79, no. 11, pp. 7731–7762, <https://doi.org/10.1007/s11042-019-08394-3>
- Avci, E & Avci, D 2009, 'An expert system based on fuzzy entropy for automatic threshold selection in image processing', *Expert Systems with Applications*, vol. 36, no. 2, pp. 3077–3085, <https://doi.org/10.1016/j.eswa.2008.01.027>
- Bębas, E, Borowska, M, Derlatka, M, Oczeretko, E, Hładuński, M, Szumowski, P & Mojsak, M 2021, 'Machine-learning-based classification of the histological subtype of non-small-cell lung cancer using MRI texture analysis', *Biomedical Signal Processing and Control*, vol. 66, article 102446, <https://doi.org/10.1016/j.bspc.2021.102446>
- Bhaskar, N, Ganashree, TS & Patra, RK 2023, 'Pulmonary lung nodule detection and classification through image enhancement and deep learning', *International Journal of Biometrics*, vol. 15(3-4), pp. 291–313, <https://doi.org/10.1504/IJBM.2023.130637>
- Fribert, P, Paulová, L, Patáková, P, Rychtera, M & Melzoch, K 2013, 'Alternativní metody separace kapalných biopaliv z média při fermentaci', *Chemické listy*, vol. 107, no. 11, pp. 843–847
- Gayap, HT & Akhloufi, MA 2024, 'Deep machine learning for medical diagnosis, application to lung cancer detection: a review', *BioMedInformatics*, vol. 4, no. 1, pp. 236–284, <https://doi.org/10.3390/biomedinformatics4010015>

Güraksın, GE & Kayadibi, I 2025, 'A Hybrid LECNN architecture: A computer-assisted early diagnosis system for lung cancer using CT images', *International Journal of Computational Intelligence Systems*, vol. 18, no. 1, article 35, <https://doi.org/10.1007/s44196-025-00761-3>

Hashim, FA & Hussien, AG 2022, 'Snake optimizer: A novel meta-heuristic optimization algorithm', *Knowledge-Based Systems*, vol. 242, article 108320, <https://doi.org/10.1016/j.knosys.2022.108320>

Hussain, L, Almarashi, MS, Aziz, W, Habib, N & Saif Abbasi, S-U-R 2024, 'Machine learning-based lungs cancer detection using reconstruction independent component analysis and sparse filter features', *Waves in Random and Complex Media*, vol. 34, no. 1, pp. 226-251, <https://doi.org/10.1080/17455030.2021.1905912>

Ishtiaq, U, Abdul Kareem, S, Abdullah, ERMF, Mujtaba, G, Jahangir, R & Ghafoor, HY 2020, 'Diabetic retinopathy detection through artificial intelligent methods: a review and open issues', *Multimedia Tools and Applications*, vol. 79, pp. 15209-15252, <https://doi.org/10.1007/s11042-018-7044-8>

Ishtiaq, U, Abdullah, ERMF & Ishtiaque, Z 2023, 'A hybrid method for diabetic retinopathy detection based on ensemble-optimized CNN and texture features', *Diagnostics*, vol. 13, no. 10, article 1816, <https://doi.org/10.3390/diagnostics13101816>

Jasmine, MPP, Rajini, KKGK, Hariharan, K, Raj, KU, Ram, KB, Indragandhi, V, Subramaniaswamy, V & Pandya, S 2023, 'Lung diseases detection using various deep learning algorithms', *Journal of healthcare engineering*, vol. 2023, no. 1, article 3563696, <https://doi.org/10.1155/2023/3563696>

Kareem, HF, AL-Husieny, MS, Mohsen, FY, Khalil, EA & Hassan, ZS 2021, 'Evaluation of SVM performance in the detection of lung cancer in marked CT scan dataset', *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 3, pp. 1731-1738, <https://doi.org/10.11591/ijeecs.v21.i3.pp1731-1738>

Kavyashree, C, Vimala, H & Shreyas, J 2024, 'A systematic review of artificial intelligence methods for oral cancer detection', *Healthcare Analytics*, vol. 5, article 100304, <https://doi.org/10.1016/j.health.2024.100304>

Manjupriya, R & Leema, AA 2025, 'Efficient epileptic seizure detection with optimal channel selection and FIXUPACTBI-LSTM deep learning model', *International Journal of Technology*, vol. 16, no. 2, pp. 706-721, <https://doi.org/10.14716/ijtech.v16i2.7333>

Mathivanan, SK, Sonaimuthu, S, Murugesan, S, Rajadurai, H, Shivahare, BD & Shah, MA 2024, 'Employing deep learning and transfer learning for accurate brain tumor detection', *Scientific reports*, vol. 14, no. 1, article 7232, <https://doi.org/10.1038/s41598-024-57970-7>

Mehrotra, R, Namuduri, KR & Ranganathan, N 1992, 'Gabor filter-based edge detection', *Pattern recognition*, vol. 25, no. 12, pp. 1479-1494, [https://doi.org/10.1016/0031-3203\(92\)90121-X](https://doi.org/10.1016/0031-3203(92)90121-X)

Mohamed, TI, Oyelade, ON & Ezugwu, AE 2023, 'Automatic detection and classification of lung cancer CT scans based on deep learning and ebola optimization search algorithm', *PloS one*, vol. 18, no. 8, article e0285796, <https://doi.org/10.1371/journal.pone.0285796>

Mohanaiah, P, Sathyanarayana, P & GuruKumar, L 2013, 'Image texture feature extraction using GLCM method', *International journal of scientific and study publications*, vol.3, no. 5, pp. 1-5

Naghipour, M, Ling, LS & Connie, T 2024, 'A Review of AI techniques in fruit detection and classification: analyzing data, features and AI models used in agricultural industry', *International Journal of Technology*, vol. 15, no. 3, pp. 585-596, <https://doi.org/10.14716/ijtech.v15i3.6404>

Nurfirdausi, AF, Apsari, RA, Wijaya, SK, Prajitno, P & Ibrahim, N 2022, 'Wavelet decomposition and feedforward neural network for classification of acute ischemic stroke based on electroencephalography', *International Journal of Technology*, vol. 13, no. 8, pp. 291-319, <https://doi.org/10.14716/ijtech.v13i8.6132>

Pham, HV, Chu, T, Le, TM, Tran, HM, Tran, HTK, Yen, KN & Dao, SVT 2025, 'Comprehensive evaluation of bankruptcy prediction in taiwanese firms using multiple machine learning models', *International Journal of Technology*, vol. 16, no. 1, pp. 289-309, <https://doi.org/10.14716/ijtech.v16i1.7227>

Pradhan, K & Chawla, P 2024, 'Lung cancer detection using deep learning algorithm', Paper presented at the AIP Conference Proceedings

Rahane, W, Dalvi, H, Magar, Y, Kalane, A & Jondhale, S 2018, 'Lung cancer detection using image processing and machine learning healthcare', Paper presented at the 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)

Reduwan, NH, Aziz, AA, Mohd Razi, R, Abdullah, ERMF, Mazloom Nezhad, SM, Gohain, M & Ibrahim, N 2024, Application of deep learning and feature selection method on external root resorption identification on CBCT images', *BMC Oral Health*, vol. 25, no. 1, article 167, <https://doi.org/10.1186/s12903-024-05030-x>

Siegel, RL, Giaquinto, AN & Jemal, A 2024, 'Cancer statistics, 2024', *CA: a cancer journal for clinicians*, vol. 74, no. 1, pp. 12-49

SR, SC & Rajaguru, H 2019, 'Lung cancer detection using probabilistic neural network with modified crow-search algorithm', *Asian Pacific journal of cancer prevention: APJCP*, vol. 20, no. 7, article 2159, <https://doi.org/10.31557/apjcp.2019.20.7.2159>

Sushanki, S, Bhandari, AK & Singh, AK 2024, 'A review on computational methods for breast cancer detection in ultrasound images using multi-image modalities', *Archives of Computational Methods in Engineering*, vol. 31, no. 3, pp. 1277-1296, <https://doi.org/10.1007/s11831-023-10015-0>

Tey, W, Goh, H, Lim, AH & Phang, C 2023, 'Pre- and post-depressive detection using deep learning and textual-based features', *International Journal of Technology*, vol. 14, no. 6, pp. 291-319, <https://doi.org/10.14716/ijtech.v14i6.6648>

Toğaçar, M, Ergen, B & Cömert, Z 2020, '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, <https://doi.org/10.1016/j.bbe.2019.11.004>

Versaci, M & Morabito, FC 2021, 'Image edge detection: A new method based on fuzzy entropy and fuzzy divergence', *International Journal of Fuzzy Systems*, vol. 23, no. 4, pp. 918-936, <https://doi.org/10.1007/s40815-020-01030-5>

Wahab Sait, AR 2023, 'Lung cancer detection model using deep learning method', *Applied Sciences*, vol. 13, no. 22, article 12510, <https://doi.org/10.3390/app132212510>

Wang, J, Wang, S & Zhang, Y 2025, 'Deep learning on medical image analysis', *CAAI Transactions on Intelligence Technology*, vol. 10, no. 1, pp. 1-35, <https://doi.org/10.1049/cit2.12356>

Wang, L 2022, Deep learning techniques to diagnose lung cancer. *Cancers*, vol. 14, no. 22, article 5569, <https://doi.org/10.3390/cancers14225569>

Wankhade, S & Vigneshwari, S 2023, 'A novel hybrid deep learning method for early detection of lung cancer using neural networks', *Healthcare Analytics*, vol. 3, article 100195, <https://doi.org/10.1016/j.health.2023.100195>

Xu, X, Xu, S, Jin, L & Song, E 2011, 'Characteristic analysis of Otsu threshold and its applications', *Pattern recognition letters*, vol. 32, no. 7, pp. 956-961, <https://doi.org/10.1016/j.patrec.2011.01.021>

Yan, C & Razmjooy, N 2023, 'Optimal lung cancer detection based on CNN optimized and improved Snake optimization algorithm', *Biomedical Signal Processing and Control*, vol. 86, article 105319, <https://doi.org/10.1016/j.bspc.2023.105319>