Cureus Journal of Computer Science

Open Access Original Article

Deep Hybrid Neural Network: Unveiling Lung Cancer With Deep Hybrid Intelligence from CT Scans

Received 10/25/2024
Review began 11/05/2024
Review ended 02/10/2025
Published 02/10/2025

DOI:
https://doi.org/10.7759/s44389-024-02008-2

Chinmayee Nayak [1] , Alakananda Tripathy [1] , Manoranjan Parhi [2]

1. Centre for Artificial Intelligence and Machine Learning, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, IND 2. Centre for Data Science, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, IND

Corresponding authors: Chinmayee Nayak, chinmayeenayak@soa.ac.in, Alakananda Tripathy, alakanandatripathy@soa.ac.in

Abstract

Lung cancer remains one of the main causes of cancer-related mortality globally, often diagnosed at late stages due to non-specific early symptoms. Traditional diagnostic methods, including biomarker analysis, biopsies, and image analysis, have long evaluation periods, subjectivity, and interobserver variability. To improve diagnostic accuracy and efficiency, this study introduces a deep hybrid neural network (DHNN) for lung cancer diagnosis from histopathology images. The proposed model incorporates EfficientNet as the basis architecture and utilizes data augmentation techniques to optimize performance. The suggested DHNN outperformed the traditional convolutional neural network (94.03%) by achieving an accuracy of 98.86%. It demonstrated exceptional precision, identifying malignant patients with higher reliability while reducing misclassification rates. This strategy provides a strong and effective substitute for current techniques, speeding up diagnosis and maybe enhancing patient outcomes.

Categories: Health Informatics, Machine Learning (ML)
Keywords: hybrid deep learning technique, fully connected layer, steam convolution layer, flattened layer, global average pooling, lungs cancer, data augmentation, cnn, maxpooling

Introduction

Lung cancer is the deadliest cancer in the world, with an annual fatality rate of 1,51,900 [1]. The unchecked proliferation and spread of abnormal cells throughout the body formally determine cancer. Although risk factors such as smoking, alcohol use, genetic predisposition, and immune system deficiencies have been linked to lung cancer, the precise etiology of the disease is still mostly unclear. As technology advances, the prevalence of cancer is increasing worldwide. The goal of this study is to accurately classify the subtypes of lung cancer by utilizing the efficacy of convolutional neural networks (CNNs) in digital pathology. Conventional diagnosis depends on closely examining CT scans, which is a laborious and often inaccurate process [2]. Modern optimization strategies and sophisticated image processing techniques are crucial for producing precise and useful imaging to overcome this difficulty. The proposed method aims to improve doctors' ability to identify lung nodules early, enabling prompt treatment of primary forms of lung cancer. Cancerous cells can develop into nodules or tumors that require immediate intervention because they multiply more rapidly than healthy cells [3]. Raoof et al. [4] explain the spread of lung cancer cells to other areas of the body. The study discusses various machine learning (ML) models for detecting the spread of lung cancer. A support vector machine (SVM) model was considered to be the best among all the other models for lung cancer prediction. The study also explores the conditions under which men and women are more likely to be affected by lung cancer [5]. It is interesting to note that women, even those who do not smoke, have a markedly higher risk of developing lung cancer.

Approximately 75% of lung cancer cases are non-small-cell lung cancer. It should be noted that not all lung nodules are cancerous; nodules that are less than 3 mm in diameter are usually benign, while those more than 4 mm have the ability to develop into cancer. A deep neural network, ConvNets, was used to feed image data to the model and classify cancer types in a more optimized way. Song et al. [6] suggested that computer-aided diagnosis (CAD) technologies can be used for the early detection of the disease in a more accurate way than medical professionals. Various deep neural networks were used to perform the classification task. The CT scan image database was utilized to evaluate the model. The CNN architecture outperformed all other deep learning models. Metrics like accuracy, speed, and automation level are used to assess how well CAD systems work [7].

Manual identification of CT scans is often not enough to get accurate information about disease progression [8]. As a result, early and accurate detection of lung cancer depends on ML methods. In order to isolate and find tiny nodules inside the pulmonary vasculature, the initial technique for identifying pulmonary nodules entails eliminating the lung endothelium from the CT imaging [9]. A major challenge in detecting lung cancer is the lack of an efficient screening technique, which leads to most diagnoses being made at advanced stages, such as Stage III or IV. For subtyping and staging, radiologists typically manually examine and interpret medical images of patients who may have cancer. This manual process is susceptible to subjectivity, inter-observer variability, misclassification of cancer subtypes, and substantial time consumption [10]. A deep CNN model was developed to improve the precision and effectiveness of lung cancer diagnosis and prediction to overcome these limitations.

For lung tumors to have the best prognosis and treatment, which can greatly increase survival rates, accurate diagnosis and prediction are essential. The following are some advantages of using deep learning to diagnose lung cancer:

• Reducing the likelihood of misdiagnosis by managing large datasets and producing extremely accurate forecasts.

• Facilitating early detection of lung cancer, which increases the chances of a successful outcome.

• Enabling more personalized care, as deep learning can predict how each patient will respond to treatments, allowing for more efficient and customized treatment regimens for each individual.

Deep learning has the potential to enhance patient care and decrease lung cancer mortality through early diagnosis and personalized treatment. The proposed method can diagnose lung cancer from images in under 5 s, achieving an accuracy of 98.86%. This study is organized as follows: first, a review of related work, followed by a detailed explanation of the methodologies used, then the presentation and discussion of results, and concluding with a final summary and the implications of the findings.

Materials And Methods

Related work

The importance of medical imaging technologies in the diagnosis and treatment monitoring of early-stage lung cancer is discussed by Wang [11]. Many imaging modalities, including CT, MRI, PET, chest X-rays, and molecular imaging, are commonly used to identify lung cancer. However, these techniques have limitations, especially when it comes to automatically classifying images of cancer, which reduces their effectiveness for patients with other illnesses. Recent advancements in deep learning-based imaging methods for early lung cancer detection are the main focus of this research.

A study by Javed et al. [12] explores how deep learning techniques are essential tools in lung cancer diagnosis, utilizing a range of imaging modalities, including whole slide imaging and X-rays. It provides a systematic review of research conducted from 2015 to 2024 on improving lung cancer prognosis and therapy through advanced detection methods and quality evaluations. This study also highlights the CNN, due to its multilayer architecture and ability to convey local weights. CNNs perform better in the detection and classification of lung cancer.

A study by Asuntha and Srinivasan [13] discusses that one of the leading causes of death worldwide is lung cancer. The goal of the study is to highlight how CT scan images can be used to identify and evaluate the severity of lung cancer progression. A real-time hospital dataset is used to perform the task. To extract texture and intensity information from the image data, the study uses effective feature extraction methods, such as wavelet transformation analysis and the invariant feature transform method. Fuzzy particle swarm optimization-based CNN, a modified CNN model, is used to perform the experiment, and then conventional CNN is used for final classification. The study ultimately demonstrates that the approaches used to categorize the images are more better than other alternative approaches.

According to a study [14], deep learning-based systems combined with conventional CAD schemes investigate pattern recognition based on neural network techniques to detect lung cancer. The study also assesses several deep neural networks, where CNNs are identified as the most widely used technique for detecting lung cancer more accurately. It sums up by addressing the challenges of implementing deep neural networks for lung cancer prognosis and identifying the advantages and drawbacks of recently used methods in this field.

Subramanian et al. [15] trained a CNN using ImageNet models, namely LeNet, and VGG-16, using a dataset of CT scans in JPG format. With an accuracy of 99.52%, the suggested model uses AlexNet and applies the features from its final fully connected layer to a softmax classifier. The softmax layer, in conjunction with AlexNet, provides a reliable and long-lasting model for lung cancer detection.

The use of low-dose computed tomography (LDCT) scans for early cancer diagnosis has been discussed by Lakshmanaprabu et al. [16]. This technique can reduce cancer mortality to some extent compared to conventional radiography. The research presents ML-based techniques for formulating disease progression evaluation in order to detect lung cancer accurately. A deep neural network is used to create a unique split technique to distinguish between different classes. To enhance network, multiple networks and data dynamics with constraint-handling techniques are incorporated. A hierarchical semantic CNN was used to improve the quality of visual acuity.

The current study introduces a machine-based diagnostic technique for lung cancer detection from CT scan images [17]. This study used LDA and an optimal neural network to analyze lung CT scan images. The modified gravitational search algorithm optimization technique is applied to the neural networks for better prediction accuracy. The results suggested that the classifier yields 96.2% recall, 94.2% precision, and 94.56% accuracy, according to the comparison results. The early identification of lung cancer is the focus of the article [18]. To improve early detection, CT scans are being used as part of global automation programs. Using sophisticated image processing techniques, such as image segmentation algorithms and dimensionality reduction, the study focuses on diagnosing lung cancer occurrences.

A study [19] uses a variety of architectures, including AlexNet, VGG16, and VGG19, to extract deep learning features after pre-processing the raw data to improve image contrast. The feature set is optimized using a genetic algorithm to enhance detection accuracy. Several classifiers are examined to detect lung nodules; the VGG19 architecture and an SVM classifier yields the best results. The I-ELCAP database is used to evaluate the system, employing 320 LDCT images from 50 individuals. . A study [20] presents an automatic identification method for detecting lung tumors in CT scans from the obtained dataset, with an emphasis on areas of interest. The system processes DICOM (Digital Imaging and Communications in Medicine) images of size 512 × 512 and uses the watershed algorithm in conjunction with filters (median, Gaussian, and Gabor) to segment lung areas. For feature extraction, deep learning models are employed; AlexNet uses 227 × 227 × 3 "fc7" layers. With 100% accuracy and specificity, a multi-class SVM classifier demonstrated the best results, showing its efficacy in deep learning-based nodule detection.

Using a deep hybrid deep learning model that incorporates several data types and assessment indicators, the suggested study proposes a novel method for lung cancer diagnosis. The study shows that the deep hybrid neural network (DHNN) achieves higher accuracy in lung cancer diagnostics by comparing its performance with other previously analyzed models. This study fills a crucial gap in lung cancer identification, as implemented models often fail to accurately and early detect the disease and have difficulty integrating various data sources. By addressing these issues, this study aims to provide a more thorough and reliable method for diagnosing lung cancer.

This section outlines the research process, including the dataset preparation process and the model architecture. The dataset used in this study contains 1018 lung images, consisting of 600 images of lung cancer samples and 518 images of non-malignant lung tissue. The Lung Image Database Consortium is one of the public databases from which the CT images were initially obtained [21]. Compared to MRI and PET images, lung CT images have less noise. To enhance the deep learning model's resilience, augmentation techniques were applied to increase the dataset's size to 11,040 images in both classes.

In this study, a CNN was used to classify lung cancer from medical imaging data. The CNN architecture, as shown in Figure *1*, was designed to differentiate between cancerous and non-cancerous lung cells by processing input images through many convolutional layers, followed by pooling and fully connected layers. While the CNN model consistently extracts significant features for classification, its performance is limited in capturing complex patterns in lung tissue images. To overcome this constraint, we proposed DHNN, which combines additional augmentation layers for improved accuracy, with EfficientNetB0 used as a feature extractor. The CNN and DHNN were compared, and the results showed that the proposed DHNN performs better than the traditional CNN in terms of lung categorization accuracy.

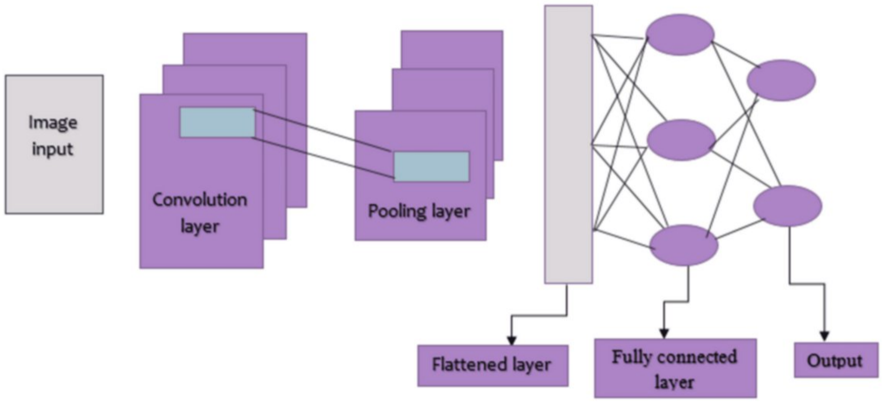

FIGURE 1: CNN architecture for image classification

CNN: Convolutional Neural Network

The CNN architecture, shown in Figure 1, was designed to distinguish between cancerous and non-cancerous lung cells as follows:

$$\text{dimension of output image} = \left\lfloor \frac{N + 2P - F}{S} + 1 \right\rfloor \times \left\lfloor \frac{N + 2P - F}{S} + 1 \right\rfloor \qquad (1)$$

where N = input pixel number, F = filters number in pixel, P = padding, and S = stride.

The non-linear activation function, ReLU, is applied to each layer of the convolution operation. To refine the behavior of neurons in neural networks, a selective activation function was used, as defined in Equation 2:

$$f(x) = \max(0, x) \qquad (2)$$

Each convolution operation's output was passed through the activation function, ReLU, which is theoretically defined as $f(x) = \max(0,x)$ in order to precisely simulate how neurons would behave in the actual world. The choice of ReLU as the activation function ensures that not all neurons are activated simultaneously, enabling selective activation during backpropagation. This increases learning efficiency by preventing all neurons from activating simultaneously. The data were subsequently subjected to the pooling layer, which was used to downsample the feature maps. This procedure introduced translation invariance, which improved the model's ability to tolerate slight shifts and distortions in lung cancer tissue images, in addition to reducing the dimensionality of the feature maps, which decreased the number of parameters needed to be learned in later layers. The model's capacity to detect and categorize lung cancer is improved by this series of steps.

To downsample the feature maps and reduce dimensionality while preserving important information, the study used MaxPooling. A flattening layer was then applied to the output, converting the pooled data into a one-dimensional vector that could be fed into the fully connected layers of the feedforward neural network. Learnable weights were used in these fully connected layers to link every node in the subsequent layer to every node in the previous layer. Thus, the output layer, where classification took place, was mapped to the characteristics extracted from the pooling layers. The output layer employed the SoftMax activation function for lung cancer classification, which allowed the model to assign probabilities to each class. The binary cross-entropy loss function, commonly employed in classification problems, was used to train the model. The following is the definition of the loss function for binary classification:

$$\text{loss} = -\frac{1}{N} \sum_{i=1}^{N} \left[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \right] \qquad (3)$$

where N = sample number, y_i = actual probability of the outcome, and $\hat{y}_i$ = predicted probability of the outcome.

Comparison with the proposed DHNN model

Although CNNs have demonstrated significant performance in lung cancer classification, their ability to capture complex patterns and achieve high classification accuracy remains limited. To address these limitations, we propose DHNN, which integrates EfficientNetB0 as the feature extractor along with additional enhancement layers for improved performance. The DHNN model employs a sophisticated data augmentation technique to enhance generalization and robustness. Given a dataset $\mathrm{D} = \{(x_i, y_i)\}_{i=1}^{N}$, augmentation techniques such as flipping, zooming, scaling, translation, and rotation are applied, defined as:

$$x_i' = \mathcal{F}_r(\mathcal{F}_t(\mathcal{F}_s(\mathcal{F}_z(\mathcal{F}_f(x_i))))) \qquad (4)$$

where the dataset displays variability due to the different transformation functions $\mathcal{F}$. The DHNN design optimizes depth, width, and resolution by substituting EfficientNetB0 for traditional CNN layers and utilizing compound scaling. In order to enhance generalization and avoid overfitting, dropout layers and batch normalization are also included. The Adam optimizer is used to update the model's weights:

$$\mathbf{W}^{(t+1)} \leftarrow \mathbf{W}^{(t)} - \eta \nabla \mathcal{L}(\mathbf{W}^{(t)}) \qquad (5)$$

Experimental results demonstrate that the DHNN outperforms the traditional CNN in terms of classification accuracy and loss. The introduction of EfficientNetB0 as a feature extractor significantly enhances the model's ability to detect subtle cancerous patterns in lung tissue images.

Evaluation metrics

Standard assessment measures like accuracy and loss were used to compare CNN with DHNN's performance over different epochs. The trained model's ultimate validation loss is calculated as Equation 3. DHNN is an advanced deep learning architecture designed for detecting lung cancer using medical imaging data. The model's foundation architecture is EfficientNetB0, and it incorporates a robust data augmentation process to enhance model generalization. Given a dataset:

$$\mathbf{D} = \{(x_i, y_i)\}_{i=1}^{N} \qquad (6)$$

The procedure starts by applying a series of transformations, including flipping, zooming, scaling, translation, and rotation, which are mathematically represented as follows: where x_i represents medical images and y_i indicates related labels:

$$x_i' = \mathcal{F}_r(\mathcal{F}_t(\mathcal{F}_s(\mathcal{F}_z(\mathcal{F}_f(x_i))))) \qquad (7)$$

The original dataset is then mixed with the enhanced dataset '**D**', guaranteeing a rich and varied training set. The dataset is then divided into training and validation sets:

$$(\mathbf{D}_{\text{train}}, \mathbf{D}_{\text{val}}) \leftarrow \text{Split}(\mathbf{D}, \text{test ratio} = 0.2) \qquad (8)$$

The EfficientNetB0 model is initialized with pre-trained weights, and additional layers, including Dropout, Global Average Pooling, and Fully Connected layers, are added for classification. The Adam optimizer is used to optimize the model parameters during training, updating weights at each epoch in the manner described below:

$$\mathbf{W}^{(t+1)} \leftarrow \mathbf{W}^{(t)} - \eta \nabla \mathcal{L}(\mathbf{W}^{(t)}) \qquad (9)$$

The loss function computed is the binary cross-entropy loss, computed as:

$$\mathbf{L} = -\frac{1}{m} \sum_{i=1}^{m} \left[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \right] \qquad (10)$$

where $\hat{y}_i$ presents the predicted probability of lung cancer presence. The validation accuracy is calculated using

$$\text{Accuracy} = \frac{1}{m} \sum_{i=1}^{m} \nVdash(\hat{y}_i = y_i) \qquad (11)$$

where the indicator function is **L**. The ultimate validation loss after training is calculated as follows:

$$\mathbf{L}_{\text{val}} = -\frac{1}{m} \sum_{i=1}^{m} \left[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \right] \qquad (12)$$

Finally, the trained model is returned along with the classification accuracy, ensuring a robust system for lung cancer diagnosis in medical imaging. Algorithm 1 depicts the step-by-step process of the proposed DHNN model, and Figure 2 depicts the step-by-step architecture of the proposed model.

Algorithm 1: Deep Hybrid Neural Network Model (DHNN)

{

Require Dataset: $\mathbf{D} = \{(x_i, y_i)\}_{i=1}^{N}$ with images and labels

Ensure Trained EfficientNetB0 model for Prediction

State: **Step 1: Data Augmentation**

For each $x_i \in \mathbf{D}$

State: Apply transformations: $x_i' = \mathcal{T}(x_i)$, where

$$x_i' = \mathcal{F}_r(\mathcal{F}_t(\mathcal{F}_s(\mathcal{F}_z(\mathcal{F}_f(x_i)))))$$

State: Augmented images $\mathbf{D}' = \{(x_i', y_i)\}_{i=1}^{kN}$

End far

State: $\mathbf{D} \leftarrow \mathbf{D} \cup \mathbf{D}$

State: **Step 2: Train-Test Split**

State $(\mathbf{D}_{\text{train}}, \mathbf{D}_{\text{val}}) \leftarrow \text{Split}(\mathbf{D}, \text{test ratio} = 0.2)$

State: **Step 3: Model Initialization**

State: Load EfficientNetB0 base model with ImageNet weights

State: Include Dropout, Global Average Pooling, and Fully Connected layers.

State: **Step 4: Training**

For each epoch $t = 1$ to $\mathcal{T}$

State: Update weights with Adam optimizer:

$$\mathbf{W}^{(t+1)} \leftarrow \mathbf{W}^{(t)} - \eta \nabla \mathcal{L}(\mathbf{W}^{(t)})$$

State: Compute loss:

$$\mathbf{L} = -\frac{1}{m} \sum_{i=1}^{m} \left[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \right]$$

State: Compute validation accuracy:

$$\text{Accuracy} = \frac{1}{m} \sum_{i=1}^{m} \nVdash (\hat{y}_i = y_i)$$

End for

State: **Step 5: Evaluation**

State: Compute final validation loss:

$$\mathbf{L}_{\text{val}} = -\frac{1}{m} \sum_{i=1}^{m} \left[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \right]$$

State: Plot training curves for loss and accuracy

State: **Step 6: Output Results**

State: Return trained model and classification accuracy

}

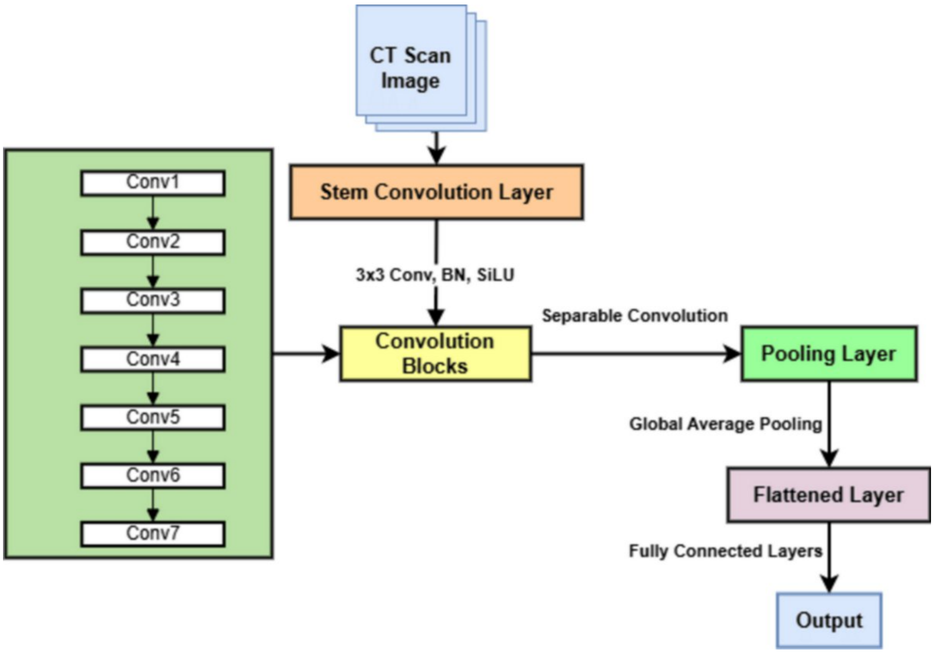

FIGURE 2: Proposed DHNN model architecture for image classification

DHNN: Deep Hybrid Neural Network

Results

To evaluate the effectiveness of the proposed DHNN model, we conducted in-depth experiments and compared the results with a conventional CNN model. A total of 11,040 images were used, equally distributed between healthy cells and those affected by the serous carcinoma subtype. Figure *3* shows a sample image of a lung cancer CT scan, and Figure *4* shows the image after partial data augmentation. Hyperparameter tuning was done during the training process by gradually increasing the number of epochs in increments of 20. Model convergence and performance stability were assessed by continuously monitoring both training and validation accuracies. Table *1* presents the configuration settings of the DHNN model.

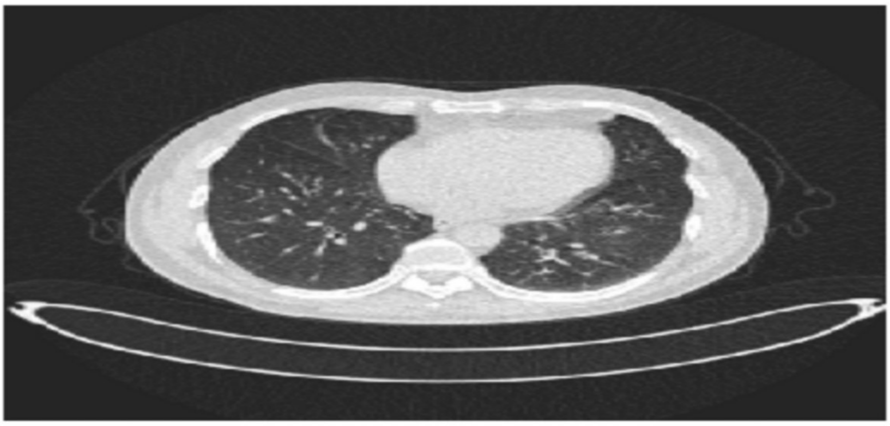

FIGURE 3: Sample image of affected lung cancer

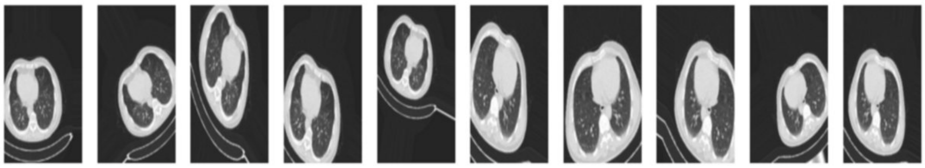

FIGURE 4: Preprocessed data after augmentation

Configuration Settings	Values
Input Image Size	(224, 224, 3)
Data Augmentation	Rotation, translation, scaling, zoom, flipping
Pre-Trained Model	EfficientNetB0 (ImageNet weights)
CNN Branch	3 convolutional layers
First Convolution Layer	Filters: 32, Activation: ReLU, Kernel: (3,3)
Second Convolution Layer	Filters: 64, Activation: ReLU, Kernel: (3,3)
Third Convolution Layer	Filters: 128, Activation: ReLU, Kernel: (3,3)
Pooling	MaxPooling (2,2) after each conv layer
Flattening	Converts feature maps to 1D vector
Fully Connected Layer	256 neurons, Activation: ReLU
EfficientNetB0 Branch	Global average pooling
Feature Fusion	Concatenation of CNN and EfficientNetB0 features
Dropout	0.5 (to prevent overfitting)
Output Layer	1 neuron, Activation: Sigmoid
Loss Function	Binary cross-entropy
Optimizer	Adam
Learning Rate	Adaptive (from Adam optimizer)
Batch Size	32 images per batch
Epochs	20
Train-Test Split	80% training, 20% validation
Evaluation Metrics	Accuracy, validation loss

TABLE 1: DHNN image classification configuration settings

CNN: Convolution Neural Network; DHNN: Deep Hybrid Neural Network Model

The improved performance of the DHNN model can be attributed to the combination of traditional CNN layers and the EfficientNetB0 branch, which enhanced feature extraction capabilities. The feature combination strategy effectively integrated deep spatial features from CNN with EfficientNetB0's global feature representations, leading to better generalization and robustness. Additionally, the application of data augmentation, dropout regularization, and adaptive optimization contributed to minimizing overfitting and improving classification accuracy. Table 2 illustrates the accuracies and loss of the proposed DHNN model across epochs.

Epochs	Accuracy (Training)	Accuracy (Validation)
2.5	98.03	93.43
5.0	98.46	84.16
7.5	98.55	90.16
10.0	98.72	98.78
20.0	98.56	98.86

TABLE 2: The recorded training and validation accuracies along with the corresponding epochs

A comparative analysis of the training and validation loss curves highlights the effectiveness of the DHNN model. The CNN model exhibited steady fluctuations and overfitting in later epochs, while the DHNN model maintained a stable and steadily decreasing loss, indicating superior learning dynamics. This suggests that the hybrid model efficiently captures discriminative features crucial for lung cancer identification. As shown in Figure *5*, the CNN model initially achieved moderate performance, reaching a training accuracy of 94.03% and a validation accuracy of 93.56% at 20 epochs. However, the proposed DHNN model demonstrated a stable improvement in performance with increased training epochs. By the 20th epoch, the DHNN model achieved a remarkable training accuracy of 98.56% and a validation accuracy of 98.86%, significantly outperforming the baseline CNN model, as demonstrated in Figure *6*.

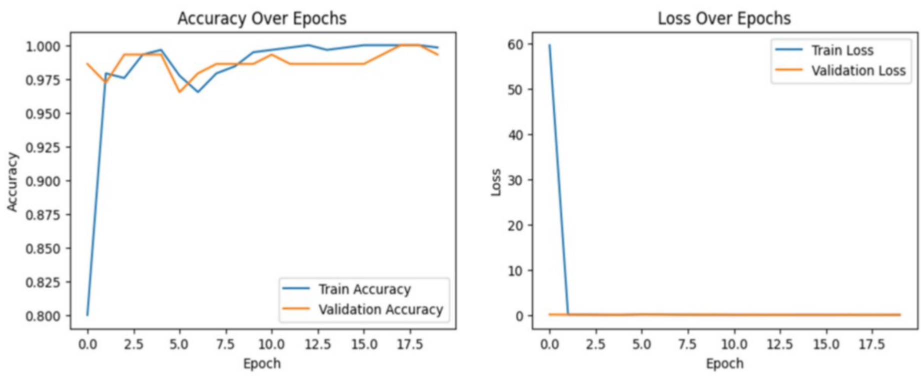

FIGURE 5: CNN model accuracy and loss for training and validation after 20 epochs

CNN: Convolution Neural Network

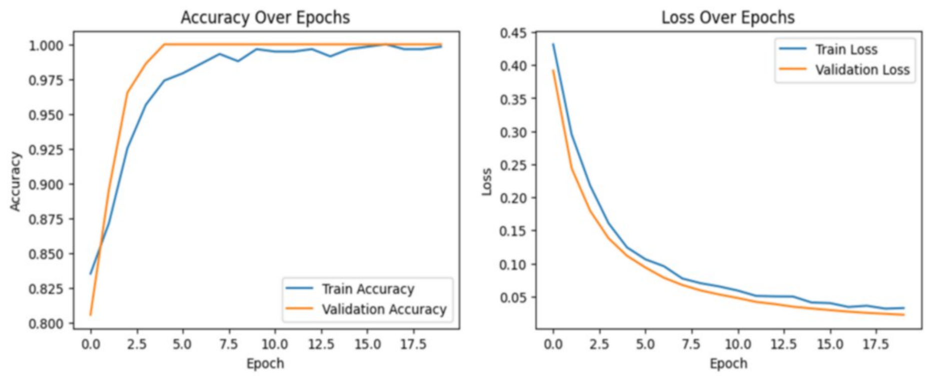

FIGURE 6: Proposed deep hybrid neural network model accuracy and loss for training and validation after 20 epochs

Following the training phase, the model was tested using images from the test dataset. For each image, the algorithm provided a classification percentage, indicating the likelihood of the sample being either lung cancer or healthy cells. The model's accuracy shows a steady increase in both training and validation performances, as illustrated in Figure *5*. The graph shows an increase in accuracy from 80.00% for training and 97.73% for validation at 2.5 epochs to 98.56% and 98.86%, respectively, after 20 epochs, underscoring the model's improved performance over time in detecting lung cancer.

Table *3* compares the training and validation accuracy, along with the loss, of the proposed DHNN model over a number of epochs and compares its performance with a traditional CNN. While the CNN model demonstrated moderate accuracy, the DHNN model achieved superior performance, attaining a validation accuracy of 98.86%. This improvement can be attributed to the integration of EfficientNetB0 with CNN, which enhanced feature extraction and generalization. Furthermore, the use of batch normalization after every convolutional layer effectively improved the model's resilience even further. The comparative investigation highlights the advantages of the proposed DHNN model in lung cancer detection over traditional CNN architectures.

Although the exact network architecture used in the study is not specified, it is likely that the model incorporates multiple deep neural network layers and employs techniques like batch normalization to reduce overfitting. This combination of deep learning's feature extraction capabilities with conventional algorithms allows for a more robust and accurate classification. The aim of this work was to propose a deep learning model to accurately distinguish between healthy lung tissue and lung cancer cells.

Model	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
Deep Hybrid Neural Network (DHNN)	98.86	98.56	0.025	0.03
Convolutional Neural Network (CNN)	94.03	93.56	3.125	4.56

TABLE 3: Comparison of performance for proposed DHNN model to CNN model

Discussion

There are some limitations in the study, including the inability to capture the dynamic nature of disease progression, which may affect the model's applicability to different patient populations. Furthermore, the absence of crucial medical information, such as genetic profiles and other relevant factors, may affect the model's effectiveness in accurately detecting lung cancer. To address these challenges, future studies should focus on partnering with medical professionals and research centers to validate the model using real-world medical data and incorporating comprehensive datasets for more accurate evaluations. Additionally, the study highlights areas for further research, including addressing the limitations related to advancing the interpretability of the model and validating it across diverse demographic groups to ensure broader applicability.

Conclusions

This study focuses on the implementation of deep learning models for the early detection of lung cancer. Our proposed work demonstrates significant accuracy of 98.86% for early detection. This work can further be extended into the fields of radiology and oncology by highlighting the potential for real-world applications in the identification and improvement of lung cancer treatment.

Considering certain drawbacks, the study presents the potential for further investigation to strengthen the findings. Improving the quality of data, testing the models on more independent datasets, and enhancing generalization through studies incorporating a range of demographics are some of these possibilities. Our work concludes by highlighting the significance of advanced computational methods for timely detection of lung cancer. By addressing current problems and improving these models to create more accurate and widely applicable clinical practice tools, we can enhance cancer diagnosis, outcome prediction, and individualized treatment strategies.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Acquisition, analysis, or interpretation of data: Alakananda Tripathy, Chinmayee Nayak, Manoranjan Parhi

Critical review of the manuscript for important intellectual content: Alakananda Tripathy, Manoranjan Parhi

Supervision: Alakananda Tripathy, Manoranjan Parhi

Concept and design: Chinmayee Nayak

Drafting of the manuscript: Chinmayee Nayak

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2018. CA: A Cancer Journal for Clinicians. 2018, 68:7-30. 10.3322/caac.21442
2. Martin MD, Kanne JP, Broderick LS, Kazerooni EA, Meyer CA: Lung-RADS: Pushing the limits. RadioGraphics. 2017, 37:1975-1993. 10.1148/rg.2017170051
3. Diciotti S, Picozzi G, Falchini M, Mascalchi M, Villari N, Valli G: 3-D segmentation algorithm of small lung nodules in spiral CT images. IEEE Transactions on Information Technology in Biomedicine. 2008, 12:7-19. 10.1109/titb.2007.899504
4. Raoof SS, Jabbar MA, Fathima SA: Lung cancer prediction using machine learning: A comprehensive approach. 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bangalore, India. 2020, 108-115. 10.1109/ICIMIA48430.2020.9074947
5. Ciompi F, Chung K, van Riel S, et al.: Towards automatic pulmonary nodule management in lung cancer screening with deep learning. Scientific Reports. 2017, 7:46479. 10.1038/srep46479
6. Song Q , L Zhao , X Luo , X Dou : Using deep learning for classification of lung nodules on computed tomography images. Journal of Healthcare Engineering. 2017, 2017:8314740. 10.1155/2017/8314740
7. Khan S, Islam N, Jan Z, Ud Din I, Rodrigues JJCC: A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. Pattern Recognition Letters. 2019, 125:1-6. 10.1016/j.patrec.2019.03.022
8. Hamad YA, Simonov K, Naeem MB: Breast cancer detection and classification using artificial neural networks. In 2018 1st Annual International Conference on Information and Sciences (AiCIS). 2018, 51-57. 10.1109/AiCIS.2018.00022
9. Bhatia S, Sinha Y, Goel L: Lung cancer detection: A deep learning approach. Soft Computing for Problem Solving. Advances in Intelligent Systems and Computing. Bansal J, Das K, Nagar A, Deep K, Ojha A (ed): Springer, Singapore; 2019. 817:699-705. 10.1007/978-981-13-1595-4_55
10. Palani D, Venkatalakshmi K: An IoT based predictive modelling for predicting lung cancer using fuzzy cluster based segmentation and classification. Journal of Medical Systems. 2019, 43:21. 10.1007/s10916-018-1139-7
11. Wang L: Deep learning techniques to diagnose lung cancer. Cancers. 2022, 14:5569. 10.3390/cancers14225569
12. Javed R, Abbas T, Khan AH, Daud A, Bukhari A, Alharbey R: Deep learning for lungs cancer detection: A review.

Artificial Intelligence Review. 2024, 57:197. 10.1007/s10462-024-10807-1
13. Shatnawi, Mohammad Q and Abuein, Qusai and Al-Quraan, Romesaa: Deep learning-based approach to diagnose lung cancer using CT-scan images. Intelligence-Based Medicine. 2025, 11:11-188. 10.1016/j.ibmed.2024.100188
14. Das S, Majumder S: Lung cancer detection using deep learning network: A comparative analysis. 2020 Fifth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN). 2020, 30-35. 10.1109/ICRCICN50933.2020.9296197
15. Subramanian RR, Mourya RN, Reddy VPT, Reddy BN, Amara S: Lung cancer prediction using deep learning framework. International Journal of Control and Automation. 2020, 13:154-160.
16. Lakshmanaprabu SK, Mohanty SN, Shankar K, Arunkumar N, Ramirez G: Optimal deep learning model for classification of lung cancer on CT images. Future Generation Computer Systems. 2019, 92:374-382. 10.1016/j.future.2018.10.009
17. Elnakib A, Amer HM, Abou-Chadi FEZ: Early lung cancer detection using deep learning optimization. International Journal of Online and Biomedical Engineering. 2020, 16:82-94. 10.3991/ijoe.v16i06.13657
18. Nayak SK, Garanayak M, Swain SK, Panda SK, Godavarthi D: An intelligent disease prediction and drug recommendation prototype by using multiple approaches of machine learning algorithms. IEEE Access. 2023, 11:99304-99318. 10.1109/ACCESS.2023.3314332
19. Venkatesh C, and Chinna Babu, J and Kiran, Ajmeera and Nagaraju, CH and Kumar, Manoj : A hybrid model for lung cancer prediction using patch processing and deeplearning on CT images. Multimedia Tools and Applications. 2024, 83:43931--43952. 10.1007/s11042-023-17349-8
20. Lakshmi L, Devi KDS, Gupta S, Reddy KAN, Grandhi SK, Panda SK: WOMT: Wasserstein distribution based minimization of false positives in breast tumor classification using deep learning. IEEE Access. 2023, 11:57831-57842. 10.1109/access.2023.3279496
21. Asuntha A, Srinivasan A: Deep learning for lung cancer detection and classification. Multimedia Tools and Applications. 2020, 79:7731-7762 . 10.1007/s11042-019-08394-3