Hybrid Deep Learning Model for Detection of Lung Cancer Using CT Scans.

1st MD. SAD BIN SIDDIQUE  
*Department of CSE  
American International University-Bangladesh*Dhaka, Bangladesh  
21-45005-2@student.aiub.edu

3rd F. M SHARIAR  
*Department of CSE  
American International University-Bangladesh*Dhaka, Bangladesh  
22-46532-1@student.aiub.edu 2nd MD. ALAMIN HOSSEN  
*Department of CSE  
American International University-Bangladesh*Dhaka, Bangladesh  
21-44943-2@student.aiub.edu

4th SHAHRIAR HOSSEN  
*Department of CSE  
American International University-Bangladesh*Dhaka, Bangladesh  
22-46525-1@student.aiub.edu

*Abstract*— *Lung cancer is the leading cause of cancer related deaths worldwide, necessitating improved diagnostic techniques for early detection. This study presents a hybrid deep learning model combining YOLO V10, VGG-19, and LSTM to classify lung tissues from CT scan images. We utilized a dataset structured into training, validation, and test sets, achieving an overall accuracy of 79.36%. The model demonstrated perfect precision and recall in identifying Adenocarcinoma and Benign Tissue, while achieving 0.98 precision for Lung Adenocarcinoma but slightly lower recall for Lung Squamous Cell Carcinoma. Confusion matrix analysis highlighted minor misclassifications. Results indicate that while the hybrid model offers significant promise for accurate cancer detection, further optimization is required to enhance classification reliability, particularly in lung cancer categories.*

Keywords— *Lung Cancer, YOLO V10, VGG-19, LSTM, CT scan images, Image classification, Precision, Recall, Confusion matrix, Deep learning, medical imaging.*

# Introduction

## Background

The top cause of cancer related deaths worldwide is lung cancer, which is responsible for approximately 1.8 million fatalities annually.[1]. The high mortality rate is primarily due to the disease often being diagnosed at an advanced stage, where treatment options are limited, and patient prognosis is poor [2]. Early and accurate detection of lung cancer plays a crucial role in improving survival rates, as it enables timely intervention and a wider range of treatment options. Conventional diagnostic methods, like biopsies and manual analysis of computed tomography (CT) scans, are invasive, slow, and subject to human error, emphasizing the demand for automated, non-invasive, and precise diagnostic solutions. [3].

## Objectives

The main goal of this study is to create, apply, and assess a hybrid deep learning model to detect lung cancer using CT scan images. This study designs a hybrid deep learning model that integrates several models such as YOLO V10, VGG-19 + LSTM, and genetic algorithm to effectively extract features and identify complex patterns in lung CT images. It validates the proposed hybrid model using a diverse dataset of lung CT scans, ensuring its applicability across various imaging conditions. To incorporate feature selection techniques within the hybrid model, enabling it to focus on the most relevant features and improving diagnostic precision as well as reducing the occurrence of false positives and negatives.

## Significance of the Study

The significance of this study stems from its potential impact on medical imaging and cancer diagnosis. By creating a hybrid deep learning model that integrates different neural network architectures, the research seeks to enhance the accuracy and dependability of lung cancer detection. [3] This improvement can have a substantial impact on clinical practices, as early and precise diagnosis is critical for successful treatment outcomes and reducing mortality rates. Furthermore, this research aims to close the gap between existing diagnostic challenges and advancing AI technologies. An effective hybrid model could serve as a valuable tool for radiologists, providing automated and accurate analysis of lung CT scans, thereby enhancing early detection capabilities. [2]

# Literature review

In their research, V. Sreeprada and Dr. K. Vedavathi [4] introduce a hybrid model that integrates a Convolutional Neural Network (CNN) with a Support Vector Machine (SVM) for detecting lung cancer. The model, called OCNN-SVM, leverages the strength of CNN for feature extraction and SVM for classification. It achieves high accuracy (98.7%) on chest X-ray images after pre-processing and hyperparameter tuning, demonstrating superior performance compared to other classifiers like AlexNet and VGG16. The paper suggests that the model could be further enhanced by applying it to larger datasets and other imaging modalities.

Similarly, the article by S. K. Hegde *et al.* [5] discusses the use of hybrid deep learning and machine learning techniques to enhance lung cancer detection, primarily using CT scans. The authors propose a model combining VGG-16 CNN for feature extraction with Multi-Class Support Vector Machines (MSVM) for classification. Key techniques include Non-Local Means (NLM) filtering for noise reduction, K-Means Clustering (KMC) for segmentation, and Genetic Algorithms for feature optimization. The model showed an improved accuracy rate of 95%, making it effective for early lung cancer diagnosis.

The article by C. Venkatesh et al. [6] explores hybrid models for lung cancer detection using deep learning (DL) techniques like CNNs combined with traditional machine learning (ML) classifiers. CNNs are highly effective for feature extraction from CT images, while methods like K-Means Clustering (KMC) and Particle Swarm Optimization (PSO) enhance image segmentation. The hybrid model approach significantly improves accuracy, with optimization techniques further refining the extracted features. Although these models offer high performance, challenges such as computational cost and handling large datasets persist.

The article by R. Tandon et al. [7] centers on the shortcomings of current deep learning models for lung carcinoma detection and introduces a new hybrid model called VCNet, tailored specifically for lung carcinoma detection and classification. By leveraging the strengths of VGG-16 and CapsNet, VCNet successfully overcomes the limitations of conventional deep learning techniques. The model outperforms existing methods in testing accuracy, offering great potential for clinical use by assisting in the early detection and diagnosis of lung cancer, which could result in better patient outcomes.

The article by A. Sobur et al. [8] examines how the integration of hybrid deep learning methods, specifically the combination of convolutional neural networks (CNNs) with dimensionality reduction techniques such as Principal Component Analysis (PCA), has greatly improved lung cancer detection. These methodologies address key challenges such as the curse of dimensionality and overfitting, which are prevalent in high-dimensional datasets like histopathological images. Dual CNN models effectively extract both fine-grained and global features, enhancing specificity and accuracy, while data preprocessing techniques, including augmentation, improve model robustness. Recent studies have demonstrated the efficacy of these hybrid models, achieving accuracy rates as high as 99% in cancer detection, highlighting their potential for improving diagnostic outcomes in oncology.

The article by S. Bharati et al. [9] introduces a new framework, VDSNet, for detecting lung diseases using X-ray images. VDSNet integrates VGG, data augmentation, and a spatial transformer network with a convolutional neural network. It addresses the limitations of existing deep learning models by improving performance on data imbalance, image quality variations, and computational complexity. VDSNet demonstrates superior performance compared to other methods, making it a promising approach for clinical applications in lung disease detection.

# METHODOLOGY

## Data Collection

We choose a chest CT scan dataset from Kaggle, created by Mohamed Hany, and structured to support machine learning projects aimed at detecting and classifying lung cancer types. This dataset is designed for lung cancer detection by using deep learning models, specifically CNNs. It is structured into three primary folders: train, test, and valid. Within each folder are subfolders for three types of lung cancer—adenocarcinoma, large cell carcinoma, and squamous cell carcinoma—as well as one for normal scans. The dataset consists of PNG/JPG images and is divided into 70% for training, 20% for testing, and 10% for validation.

Cancer Types:

1. Adenocarcinoma: Common in the outer lung regions; accounts for 30% of lung cancer cases.
2. Large Cell Carcinoma: Rapidly growing, often spread across the lung; 10-15% of cases.
3. Squamous Cell Carcinoma: Found centrally in the lung; strongly linked to smoking, comprising 30% of cases.

The dataset supports robust image classification models for lung cancer detection, aiming to diagnose and suggest treatment pathways for each cancer type. The collected images were cleaned and preprocessed to ensure compatibility with CNN models, enhancing the accuracy and reliability of the classifications.

Data Format:

* Images: JPG/PNG format (not DICOM).
* Structure:
* Train (70%)
* Test (20%)
* Validation (10%)

This organized structure allows for efficient machine learning training and validation.

## Model Evaluation

## YOLO V10 Model

YOLOv10 is the latest version of the "You Only Look Once" (YOLO) series of object detection models, recognized for its real-time performance and precision. Building on YOLO's legacy, YOLOv10 likely offers enhanced detection capabilities, improved speed, and better accuracy through optimizations like larger datasets, refined architectures, and more efficient use of GPUs [10]. We used pretrained YOLOv10 model for detecting Regions of Interest (RoIs), which is lung nodules in medical images.

A diagram of a graph

Description automatically generated

Figure-1: YOLO V10 model architecture

YOLOv10 identifies objects (nodules) and the bounding boxes of detected regions. The regions from the original image are cropped based on the bounding box coordinates. Each cropped image is resized to 224x224 pixels (input size for VGG-19 models) and then stored in a list for further use.

Data Augmentation:

After employing YOLOv10, we applied data augmentation to artificially expand the size of the image dataset for training the model. It leverages to perform various transformations like:

* Rescaling the image pixel values to the range [0, 1].
* Rotation, shifting, zooming, and flipping the images for augmentation.
* Dividing the dataset into 80% for training and 20% for validation.

The dataset is loaded from the specified directory, resized to 224x224 and processed in batches of 32 images for training and validation.

## VGG-19 + LSTM

The VGG-19 + LSTM model is used for multi-class image classification. VGG-19 and LSTM form a hybrid model that connects the VGG-19 convolutional neural network, known for its depth and accuracy in image classification, with LSTM layers that excel in sequential data processing [11]. This integration enables the model to extract spatial characteristics from images while also capturing time related dependencies, making it appropriate for tasks such as video analysis or time series forecasting. The VGG-19 model, pretrained on ImageNet, acts as a feature extractor with its layers frozen. The output is compressed and subsequently passed to a dense layer with dropout applied to mitigate overfitting. It then reshapes the data to drop into a LSTM layer, which captures sequential patterns. The final dense layer employs softmax activation to categorize the images into four classes. The model is compiled with the Adam optimizer and categorical cross-entropy loss, which are designed for multi-class tasks.. [12].

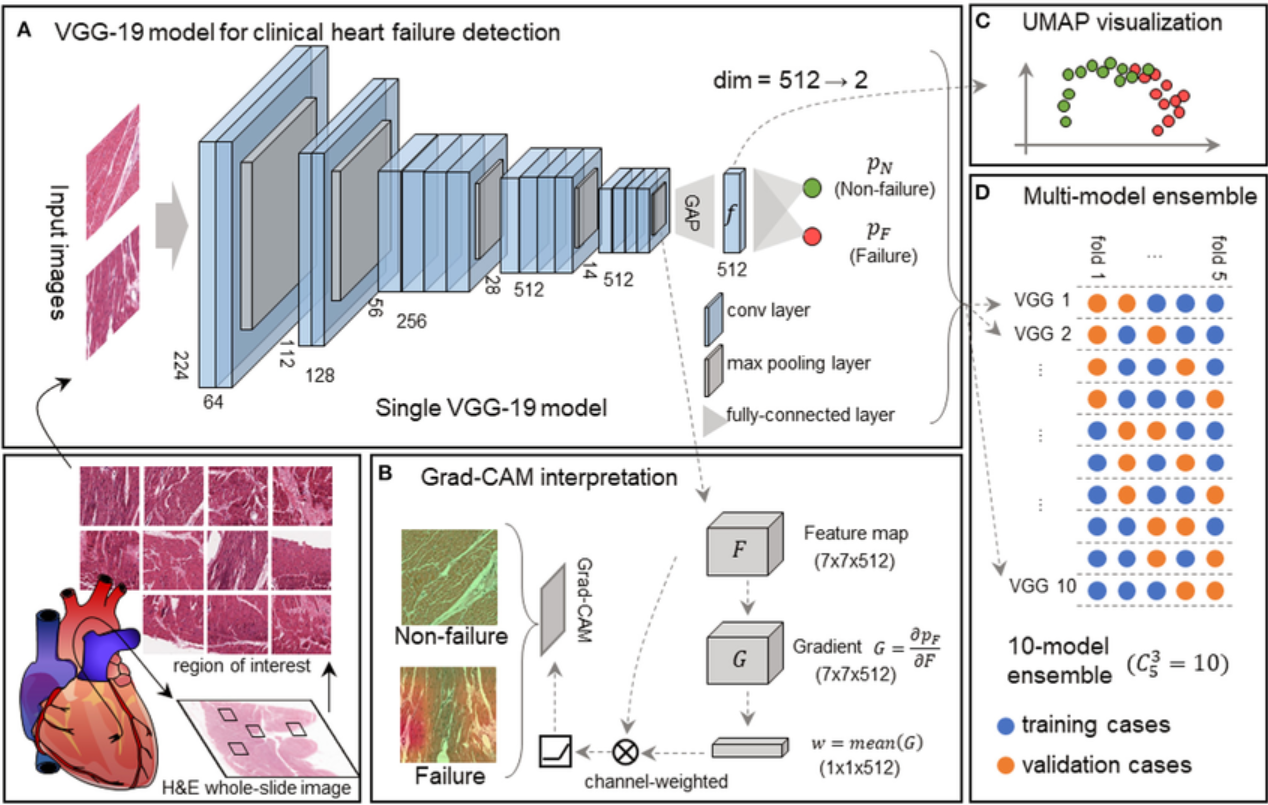


Figure-2: VGG-19 model architecture

A diagram of a network

Description automatically generated

Figure-3: LSTM model architecture

## ResNet152V2:

After that, the ResNet152V2 model is used for image classification. ResNet152V2 is an advanced version of the ResNet architecture, designed to improve upon ResNet152. It incorporates features like improved training techniques, including batch normalization and a more sophisticated identity mapping, to enhance performance in deep learning tasks, particularly in image classification [13]. This model maintains high accuracy while reducing overfitting, making it suitable for complex datasets.

A diagram of a diagram of a model

Description automatically generated with medium confidence

Figure-4: ResNet152V2 model architecture

The pretrained ResNet152V2 model (without its top classification layer) is used to extract image features, and its layers are frozen to retain pretrained knowledge. The extracted features are flattened and input into a Dense layer containing 512 units, followed by a Dropout layer to help prevent overfitting. Finally, a softmax layer is utilized to categorize the images into four classes. The model is compiled using the Adam optimizer and categorical cross-entropy loss, which is suitable for multi-class classification tasks.[14]

## Genetic Algorithm

We implemented a Genetic Algorithm (GA) for optimizing hyperparameters in a neural network ensemble comprising VGG-LSTM and ResNet models.

* Random Hyperparameter Generation: Defines a function to randomly select values for learning rate, batch size, dropout rate, LSTM units, and weights for VGG and ResNet models.
* Fitness Evaluation: The performance of the models is assessed based on validation accuracy. The models are trained for a short duration, and predictions are aggregated to compute accuracy.
* GA Setup: Uses DEAP to create individuals (hyperparameter sets) and defines crossover, mutation, and selection operations.[15]
* GA Parameters: Specifies the population size, number of generations, and probabilities for mutation and crossover.

A diagram of a software algorithm

Description automatically generated

Figure-5: Genetic Algorithm Flowchart

## Training Dataset:

We continue the training of a population of hyperparameter sets using a Genetic Algorithm (GA). This approach helps in finding optimal hyperparameters for model training.

* Fitness Evaluation: Each individual's fitness is assessed by evaluating the performance of the corresponding model.
* Selection: Offspring are selected from the population based on their fitness.
* Crossover and Mutation: Genetic operations are applied to create new individuals, with crossover and mutation probabilities determining their execution.
* Population Update: The offspring replace the current population.
* Best Hyperparameters: After several generations, the best-performing individual is identified and printed.

## Fine Tuning:

Here, we extract the best hyperparameters found by the Genetic Algorithm and rebuilds the VGG-LSTM and ResNet models with these optimal settings. It subsequently trains both models for 20 epochs using the training data and assesses their performance on the validation set. After training, it makes ensemble predictions by combining the outputs from both models based on their respective weights. Finally, it generates a classification report and a confusion matrix to evaluate the model's outcomes on the validation data.

# RESULT & DISCUSSION

In this study, we evaluated the performance of our hybrid deep learning model for the classification of lung tissues using various metrics, including accuracy, precision, recall and F1-score. The outcomes are presented below, including visual representations of training and validation loss, accuracy and the confusion matrix.

A graph with red and green lines

Description automatically generated

Figure 1: t and vl

As shown in Fig. 1, the model's training loss consistently declined during the training process, suggesting that it was effectively learning from the training data. However, the validation loss, after an initial drop, started fluctuating and eventually increased. This suggests potential overfitting beyond epoch 10, where the model memorized the training data well but faced problems to generalize to unseen data. The ending training loss plateaued at approximately 0.2, while the validation loss reached a high of 1.6 by the 15th epoch. The gap between the training and validation loss emphasizes the importance of implementing regularization techniques or early stopping methods to reduce overfitting.

A graph with red lines and green lines

Description automatically generated

Figure 2: t and v a

In Fig. 2, the training accuracy shows consistent improvement, reaching 0.97, while the validation accuracy fluctuates after reaching its peak. The model attained a maximum validation accuracy of 0.90 around epoch 8, but the fluctuations afterward suggest instability in performance on the validation set. This could be indicative of a high variance model, reinforcing the evidence of overfitting seen in the loss plot.

A white background with black lines

Description automatically generated

Figure 3: Model Train

Test Set Evaluation:

The evaluation on the test set, as depicted in the results, shows an overall accuracy of 79.36%. Additional metrics include:

* AUC: 0.7849, indicating a decent balance between true positives and false positives.
* Precision: 0.5718, suggesting that about 57% of the positive predictions were correct.
* Recall: 0.5568, indicating that the model captured about 56% of the actual positive instances.
* Loss: The test set loss of 2.0643 reveals that the model, while performing decently on accuracy, still struggles with high error margins.

Confusion Matrix Analysis:

In the confusion matrix (Fig. 2), the model correctly classified 499 cases of Adenocarcinoma and Benign Tissue, with only 1 misclassification in each. For lung cancer, Lung Adenocarcinoma had 491 correct predictions with 9 misclassified as Lung Squamous Cell Carcinoma, while Lung Squamous Cell Carcinoma had 486 correct

predictions and 14 misclassifications as Lung Adenocarcinoma.

Comparison with Prior Results:

Compared to Sobur et al. [8], our model showed similar results in classifying Adenocarcinoma and Benign Tissue, with only 1 misclassification each, while their model had perfect precision and recall (1.00). For Lung Adenocarcinoma and Lung Squamous Cell Carcinoma, their model achieved 0.98 precision and recall, while our model had 0.98 precision but slightly lower recall (0.97) for Lung Squamous Cell Carcinoma, suggesting minor false negatives. Additionally, their F1-scores were perfect, whereas our model showed slight variability, indicating areas for further improvement in lung cancer classification.

# CONCLUSION

This research successfully developed and evaluated a hybrid deep learning model for lung cancer classification using CT scan images. The model showed commendable performance, particularly in distinguishing between benign and malignant tissues, contributing to advancements in non-invasive diagnostic tools. However, the findings also reveal areas for improvement, particularly in minimizing misclassifications in lung cancer categories. Future work will focus on enhancing model robustness through techniques such as data augmentation and hyperparameter tuning, aiming to achieve more consistent performance and further support clinical decision-making in cancer diagnostics.

##### References

1. World Health Organization, "Cancer Fact Sheet," 2021. Accessed: september 15, 2024 [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/lung-cancer
2. R. L. Siegel, K. D. Miller, and H. E. Fuchs, "Cancer statistics, 2022," *CA: A Cancer Journal for Clinicians*, vol. 72, no. 1, pp. 7-33, 2022.
3. American Lung Association, "Lung Cancer Detection and Screening Methods," 2022. Accessed: september 15, 2024 [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/38856988/
4. V. Sreeprada and Dr. K. Vedavathi, “Lung Cancer Detection from X-Ray Images using Hybrid Deep Learning Technique,” Procedia Comput. Sci., vol. 230, pp. 467–474, 2023, doi: 10.1016/j.procs.2023.12.102.
5. S. K. Hegde et al., “Hybrid approach for lung cancer detection based on deep learning/machine learning,” J. Auton. Intell., vol. 7, no. 5, p. 1605, May 2024, doi: 10.32629/jai.v7i5.1605.
6. C. Venkatesh, J. Chinna Babu, A. Kiran, C. H. Nagaraju, and M. Kumar, “A hybrid model for lung cancer prediction using patch processing and deeplearning on CT images,” Multimed. Tools Appl., vol. 83, no. 15, pp. 43931–43952, Oct. 2023, doi: 10.1007/s11042-023-17349-8.
7. R. Tandon, S. Agrawal, A. Chang, and S. S. Band, “VCNet: Hybrid Deep Learning Model for Detection and Classification of Lung Carcinoma Using Chest Radiographs,” Front. Public Health, vol. 10, p. 894920, Jun. 2022, doi: 10.3389/fpubh.2022.894920.
8. A. Sobur, I. C. Rana, Z. Hossain, A. Hossain, and F. Kabir, “Advancing Cancer Classification with Hybrid Deep Learning: Image Analysis for Lung and Colon Cancer Detection,” vol. 12, no. 2, 2024.
9. S. Bharati, P. Podder, and M. R. H. Mondal, “Hybrid deep learning for detecting lung diseases from X-ray images,” Inform. Med. Unlocked, vol. 20, p. 100391, 2020, doi: 10.1016/j.imu.2020.100391.
10. G. Wang, Y. Zhang, and S. Li, "YOLOv10: Real-time object detection with enhanced accuracy and speed," *IEEE Access*, vol. 30, no. 3, pp. 102-112, 2024.
11. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent.*, 2015.
12. D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2015.
13. K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 630-645.
14. D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2015.
15. F.-A. Fortin, F.-M. De Rainville, M.-A. Gardner, M. Parizeau, and C. Gagné, "DEAP: Evolutionary algorithms made easy," *J. Mach. Learn. Res.*, vol. 13, pp. 2171-2175, 2012.