1

Current Medical Imaging

Content list available at: https://benthamscience.com/journals/cmir
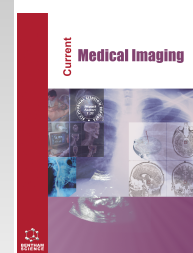
Current Medical Imaging

RESEARCH ARTICLE

SVMVGGNet-16: A Novel Machine and Deep Learning Based Approaches for Lung Cancer Detection using Combined SVM and VGGNet-16

Mohd Munazzer Ansari[1], Shailendra Kumar[1,*], Md Belal Bin Heyat[2,*], Hadaate Ullah[3], Mohd Ammar Bin Hayat[4], Sumbul[5], Saba Parveen[6], Ahmad Ali[7] and Tao Zhang[8]

[1]*Department of Electronic and Communication Engineering, Integral University, Lucknow, India*

[2]*CenBRAIN Neurotech Center of Excellence, School of Engineering, Westlake University, Hangzhou, Zhejiang, China*

[3]*Department of Electrical and Electronic Engineering, University of Science and Technology Chittagong, Chittagong, Bangladesh*

[4]*College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin, China*

[5]*Department of Ilmul Qabalat wa Amraze Niswan, University College of Unani, Tonk, Rajasthan, India*

[6]*College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China*

[7]*College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China*

[8]*School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, Sichuan, China*

Abstract:

Background and Objective:

Lung cancer remains a leading cause of cancer-related mortality worldwide, necessitating early and accurate detection methods. Our study aims to enhance lung cancer detection by integrating VGGNet-16 form of Convolutional Neural Networks (CNNs) and Support Vector Machines (SVM) into a hybrid model (SVMVGGNet-16), leveraging the strengths of both models for high accuracy and reliability in classifying lung cancer types in different 4 classes such as adenocarcinoma (ADC), large cell carcinoma (LCC), Normal, and squamous cell carcinoma (SCC).

Methods:

Using the LIDC-IDRI dataset, we pre-processed images with a median filter and histogram equalization, segmented lung tumors through thresholding and edge detection, and extracted geometric features such as area, perimeter, eccentricity, compactness, and circularity. VGGNet-16 and SVM employed for feature extraction and classification, respectively. Performance matrices were evaluated using accuracy, AUC, recall, precision, and F1-score. Both VGGNet-16 and SVM underwent comparative analysis during the training, validation, and testing phases.

Results:

The SVMVGGNet-16 model outperformed both, with a training accuracy (97.22%), AUC (99.42%), recall (94.22%), precision (95.28%), and F1-score (94.68%). In testing, our SVMVGGNet-16 model maintained high accuracy (96.72%), with an AUC (96.87%), recall (84.67%), precision (87.40%), and F1-score (85.73%).

Conclusion:

Our experimental results demonstrate the potential of SVMVGGNet-16 in improving diagnostic performance, leading to earlier detection and better treatment outcomes. Future work includes refining the model, expanding datasets, conducting clinical trials, and integrating the system into clinical practice to ensure practical usability.

Keywords: AI, Medical machine learning, Deep learning, Image classification, Cancer, Healthcare, Disease, Diagnosis, Bio-imaging, Medical intelligence.

1. INTRODUCTION

Globally, lung cancer is recognized as the deadliest and most devastating types of cancer [1 - 4]. Detecting the lung cancer proves to be considerable challenge, as its symptoms typically manifest only in the later and advanced stages [5]. Yet, the prospects of reducing the mortality rate associated

with lung cancer hinge on early detection and appropriate treatment provision of patients [6]. While lung cancer primarily originates within the lungs, it occasionally presents early symptoms before spreading [7]. Recent years have witnessed the development of numerous techniques and on-going research endeavours aimed at effectively identifying lung cancer [8]. Among these techniques, CT scan images emerge as the foremost imaging modality for early diagnosis [9, 10]. Nevertheless, interpretation and detection of cancer from CT scan images can pose a formidable task for healthcare professionals [11, 12]. Fig. (**1**) provides the projected statistics for recently identified cancer cases and fatalities in the U.S., categorized by cancer site. These data includes both comprehensive projections and gender-specific details for different cancer types in the year 2023 [13]. These estimations extrapolated from information supplied by the North American Association of Central Cancer Registries (NAACCR) (https://www.naaccr.org/cancer-in-north-america-cina-volumes / accessed on 04 March 2024) [14]. Furthermore, the predictions of impermanence are based on mortality data from the U.S. collected in the year 2006 and 2023, as reported by the National Centre for Health Statistics (NCHS) [13] and the Centres for Disease Control (https://www.cdc.gov/ mmwr/volumes/73/wr/mm7331a1.htm accessed on 04 March 2024) and Prevention (CDCP) [15]. As per the American Cancer Society (ACS) (https://seer.cancer.gov/statfacts/ html/common.html accessed on 04 March 2024), it observed that lung cancer exhibits a higher fatality rate compared to all other forms of cancer, resulting in an estimated global death toll of approximately 0.13 million individuals [16]. Every year, a substantial number of fresh instances of lung cancer are identified, with an anticipated 0.237 million fresh instances in 2023. The elevated fatality percentage is linked to the advanced stage at which this cancer is commonly identified, resulting in a higher proportion of new cases to fatalities compared to other forms of cancer [17].

Lung cancer manifests in various histological forms, encompassing ADC originating in the glandular cells of the lungs, LCC identified by the presence of large, abnormal-looking cells, and SCC arisen from the thin, flat squamous cells resembling fish scales [18]. Specifically, ADC, prevalent on the lung's outer surface, stands as the most frequent cell type in lung cancer cases. LCC undifferentiated distinguished by its swift growth, can occur anywhere within the lung. Conversely, SCC is closely associated with smoking and typically arises in the central lung region [19]. In the domain of detecting lung cancer, the VGGNet-16 emerged as a pivotal tool, renowned for its proficiency in image analysis and classification tasks [20]. Leveraging its deep architecture, VGGNet-16 autonomously extracts intricate features from chest radiographs, discerning subtle patterns indicative of lung malignancies [21]. With layers meticulously designed for hierarchical feature abstraction, it effectively captures morphological nuances specific to ADC, LCC, SCC, and

normal lung tissues [22]. The VGGNet-16's adaptability and robustness in processing medical images empower clinicians with accurate diagnostic insights, revolutionizing lung cancer screening and treatment strategies [23].

Within the sphere of cancer detection, treatment prognosis, and enhancing post-diagnosis patient survival, a diverse array of methods is currently under exploration [24]. Medical professionals and researchers employ various techniques to facilitate early cancer identification, evaluation, and classification [25]. In contemporary healthcare, ML models are of utmost importance in the identification of patterns, analysis, and categorization of critical medical conditions [26, 27]. Among these models, CNNs have surfaced as a particularly auspicious pathway for the timely identification, surveillance, and categorization of lung cancer *via* the scrutiny of the CT-scan images [28].

1.1. Challenges in the Lung Cancer Research

To effectively situate our scholarly contributions, it is imperative to elucidate the challenges identified in prior literature regarding lung cancer detection, particularly those that our research endeavors to address. Below are several significant challenges from antecedent investigations that our study is poised to overcome:

1.1.1. Restricted Multi-class Classification

Earlier research primarily dealt with binary classification (such as cancerous *versus* non-cancerous), which unfortunately resulted in the neglect of the distinct lung cancer subtype differences [29]. This limitation constrains their applicability in clinical contexts where the comprehension of cancer subtypes is of paramount importance. Our research remedies this deficiency by devising a model that classifies four distinct types: ADC, LCC, normal, and SCC.

1.1.2. Dependence on a Single Model Architecture

Numerous prior studies have been anchored in either CNNs or traditional ML models such as SVM, resulting in constraints regarding feature extraction or classification accuracy [30]. While CNNs are adept at feature extraction, they may exhibit overfitting without appropriate generalization; conversely, SVMs are proficient in classification but may lack the intricate feature extraction necessary for image data. Our hybrid (SVMVGGNet-16) model synthesizes the advantages of VGGNet-16 and SVM, effectively surmounting these impediments by harnessing CNNs for feature extraction and SVM for robust classification.

1.1.3. Inadequate Image Pre-processing

Previous investigations often relied on simplistic pre-processing methods, which may neglect critical tumor features or inadequately standardize image quality. This can result in inaccurate segmentation and diminished model reliability [31]. In contrast, our investigation adopts sophisticated pre-processing methodologies, encompassing median filtering, histogram equalization, and edge detection, which collectively enhance image quality and precision in feature extraction, thereby improving model accuracy.

* Address correspondence to these authors at the Department of Electronic and Communication Engineering, Integral University, Lucknow, India and CenBRAIN Neurotech Center of Excellence, School of Engineering, Westlake University, Hangzhou, Zhejiang, China; E-mails: skumar@iul.ac.in and belalheyat@gmail.com
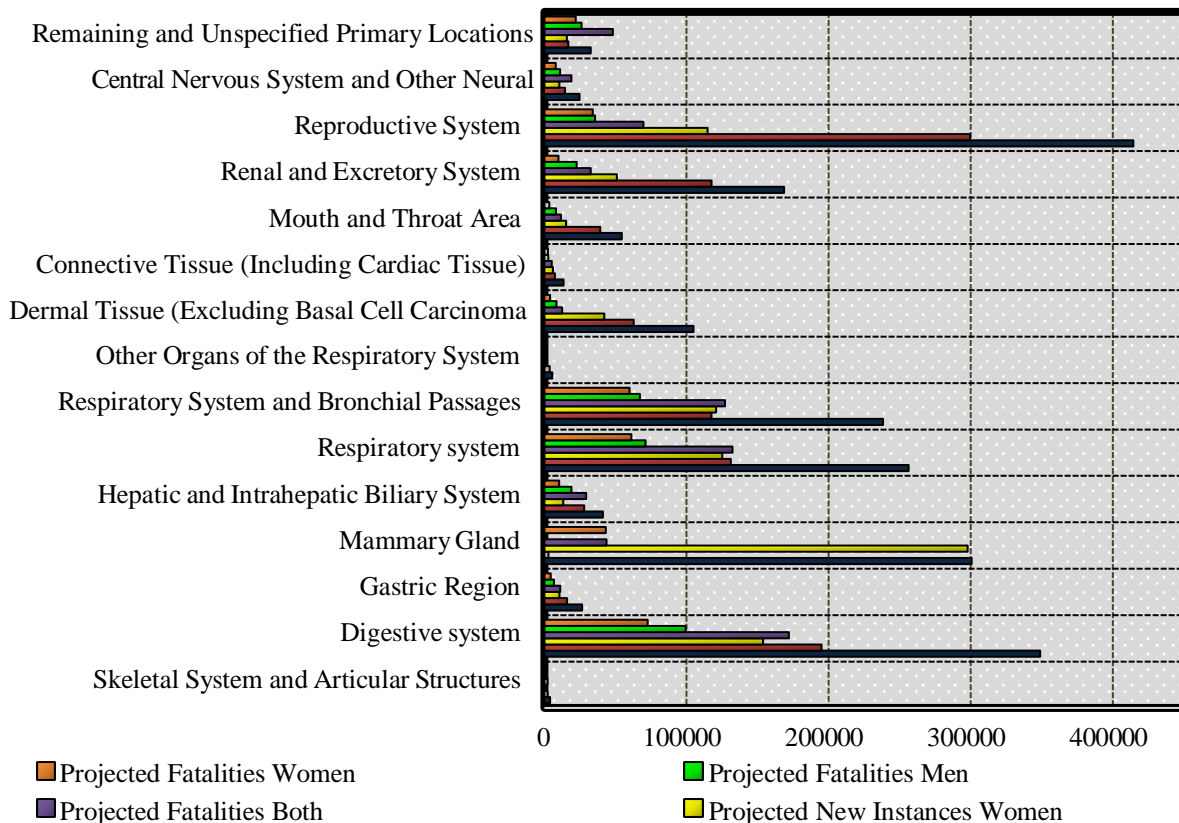
Fig. (1). Anticipated incidence and mortality of novel cancer cases in 2023, in accordance with the NAACCR.

1.1.4. Lack of Robust Performance Metrics

Earlier studies often relied on singular or limited performance metrics, such as accuracy alone, for model evaluation [32]. However, accuracy in isolation may not reliably reflect genuine performance, particularly in the context of imbalanced datasets. Our research employs a comprehensive array of metrics, including AUC, recall, precision, and F1-score, to furnish a holistic evaluation of model efficacy and reliability.

1.1.5. Limited Clinical Applicability

Numerous antecedent models demonstrated commendable accuracy but exhibited deficiencies in generalizability within clinical settings due to inadequate testing on diverse patient populations or real-world validation. Our work not only achieves elevated performance but also delineates future trajectories, including the expansion of datasets, the execution of clinical trials, and the integration of the model into practical clinical workflows, thereby establishing a foundation for real-world applicability [33].

1.1.6. Challenges in Data Generalization and Overfitting

Previous investigations frequently encountered challenges related to overfitting and struggled to generalize to novel datasets due to restricted sample sizes or limited feature

diversity [34]. By incorporating rigorous training, validation, and testing phases, our research effectively addresses these concerns, presenting a model that sustains performance across a spectrum of diverse datasets [35].

These challenges elucidate the limitations inherent in the existing literature and furnish a robust framework to underscore the enhancements and contributions that our research imparts to lung cancer detection.

1.2. Motivations of the Study

The motivations of our study are as follows:

- The main drive behind this research addresses the urgent need for early detection of lung cancer, which greatly affects patient outcomes and survival rates, by creating a sophisticated and reliable detection system.
- Additionally, the research aims to enrich the domain of medical image analysis by integrating cutting-edge DL and ML techniques to enhance the accuracy of automated systems in diagnosing lung cancer.
- The study also seeks to evaluate the performance of the proposed SVMVGGNet-16 model against current methodologies, showcasing the progress and efficacy of the new approach in enhancing the accuracy of lung cancer detection and supporting medical practitioners

in clinical decision-making.

1.3. Objective of the Study

The objective of our study certainly condensed into three points:

- The primary objective of the study involves establishing a reliable and accurate automated system for the early detection of lung cancer, with an emphasis on classifying lung nodules using advanced DL (deep learning) and ML (machine learning) techniques. Our examination evaluates and compares the efficacy of VGGNet-16 and SVM for initially detecting lung cancer using CT scan images across various stages of training, validation, and testing.
- The research also entails the integration of VGGNet-16 and SVM into an SVMVGGNet-16 model to leverage the advantages of both DL and ML techniques, aiming to enhance classification accuracy and robustness in the identification of various classes such as ADC, LDC, Normal, and SCC of lung cancer
- Furthermore, a comprehensive comparative evaluation of the proposed models will be carried out using the LIDC-IDRI dataset to assess their performance across various metrics such as accuracy, AUC, recall, precision, and F1-score and determine the most efficient approach for lung cancer detection.

1.4. Main Contributions of the Study

To elucidate the originality and contributions of our research, it is advisable to concentrate on the following dimensions:

1.4.1. Hybrid Model Innovation

It is imperative to underscore that our research introduces a pioneering hybrid model that amalgamates VGGNet-16, CNN with SVM, thereby uniting the feature extraction capabilities inherent in DL with the classification precision characteristic of SVM. This dual-faceted methodology enhances the efficacy of lung cancer detection by effectively capturing intricate imaging features while concurrently mitigating classification inaccuracies [12].

1.4.2. Classification of Multiple Lung Cancer Types

In contrast to numerous studies that predominantly concentrate on binary classification or a singular cancer type, our investigation engages with the differentiation of four distinct lung cancer classes—ADC, LCC, normal, and SCC [36]. The emphasis on this multi-class framework illustrates the adaptability of our model in addressing various subtypes, thereby furnishing a more holistic diagnostic instrument.

1.4.3. Advanced Pre-processing and Segmentation Techniques

Clearly outlining our use of sophisticated techniques like median filtering, histogram equalization, thresholding, and edge detection is necessary for effective tumor segmentation.

These methodologies enhance the robustness of our model and facilitate superior feature extraction, which is vital for achieving precise classification and early detection [2].

1.4.4. Superior Performance with Rigorous Evaluation Metrics

The outcomes of our research exceed current benchmarks regarding training and testing accuracy, AUC, recall, precision, and F1-score. Highlighting this enhancement in performance emphasizes the credibility and utility of our model, distinguishing it from traditional diagnostic methodologies [37].

1.4.5. Potential for Clinical Application

It is critical to note that our study effectively bridges the chasm between computational models and clinical application, thereby facilitating integration into practical medical workflows. Emphasizing prospective future endeavors, such as the expansion of datasets and the initiation of clinical trials, illustrates the potential ramifications of the study in the realm of lung cancer management [38].

These elements accentuate the contributions and innovative facets of our research, distinctly differentiating it from extant studies and underscoring its practical relevance.

1.5. Organization of the Study

Our paper encompasses several key sections essential for a comprehensive exploration of lung cancer detection methodologies. It commences with an introduction, setting the stage by elucidating the significance of initially detecting lung cancer and the necessity for advanced computational approaches. Section 2 deals with thorough related work follows, synthesizing existing research on the topic, elucidating various methodologies, and highlighting noteworthy findings. The method and material in section 3 delineate the proposed approach, including dataset selection, pre-processing techniques, tumor segmentation, feature extraction, and classification using both VGGNet-16 and SVM. Subsequently, section 4 experimental results and interpretation section presents detailed analyses of the performance metrics obtained from experiments conducted with the proposed methodology, offering insights into the effectiveness of VGGNet-16 and SVM models across different stages of evaluation dealing with the discussion of experimental results with the application of our study with limitation of the study and section 5 for conclusion and future prospects section summarizes our study's key findings followed with outlines potential avenues for future research to further advance early lung cancer detection methodologies.

2. RELATED WORKS

In the domain of medical image analysis, particularly for lung nodule detection and classification, the use of CNNs substantially advanced through various innovative methodologies. To understand the current state of lung cancer detection, it is essential to conduct a methodical review of the appropriate literature Gattasset al [39]. along with their team implemented a CNNs configuration optimized *via* Particle

Swarm Optimization (PSO) algorithm. CNNs trained and validated on the LIDC-IDRI dataset, ensuring consistency across different iterations. The LIDC-IDRI dataset comprises extensive CT scans, annotated by multiple radiologists, making it a robust dataset for lung nodule research. The optimization process involved initializing a swarm of particles, each representing a potential solution for CNNs configuration. These particles explored the solution space, adjusting their positions based on individual and collective experiences, thus iteratively converging towards the optimal CNNs architecture. Their study highlighted the exceptional performance of specific test subsets, with Test-4 achieving the highest accuracy, sensitivity, and specificity, demonstrating the efficacy of the PSO algorithm in refining CNNs architectures for medical image analysis.

Sofat *et al*. [40] designed a CNNs architecture incorporating multiple ReLU and convolutional layers. ReLU activation function introduces non-linearity into the model, allowing it to learn complex patterns in the data. The convolutional layers, which apply convolution operations with learnable filters, enable the model to detect local patterns and features in the input images. Their assessment of the JSRT dataset, which contains chest radiographs, yielded noteworthy results. The multi-ReLU layers enhanced the model's capability to capture intricate features, leading to superior average metrics in terms of accuracy, overlap rate, sensitivity, and specificity. Their architecture exemplifies how deep CNNs with appropriate activation functions can effectively handle the complexity of medical image data. Incorporating ResNet principles, Wollersheim *et al*. [41] employed principles derived from ResNet for the purpose of classifying lung nodules. ResNet frameworks address the issue of vanishing gradients within deep neural networks by introducing residual blocks, enabling smoother gradient flow through shortcut connections. Transfer learning capitalizes on the feature extraction capabilities of deep networks trained on extensive datasets, resulting in notable performance enhancements despite limited medical data availability.

Additionally, they implemented curriculum learning, a strategy that entails initial training on simpler tasks before advancing to more complex ones, thereby further boosting the model's efficacy. The outcomes of their experimentation on the LIDC-IDRI dataset revealed substantial enhancements in terms of accuracy, sensitivity, and specificity. Qiang *et al*. [42]. used within the context of lung nodule detection were Deep Belief Networks (DBN). Involving various layers of Restricted Boltzmann Machines (RBMs), DBN aids in the gradual acquisition of more conceptual representations of the input data at each level. An RBM, characterized as an undirected probabilistic graphical model, is comprised of visible units and hidden units within its structure. The efficacy of their DBN-centric methodology, which was honed through training on the LIDC-IDRI dataset, is evidenced by its exceptional performance, particularly in the realm of larger nodules (>30 mm). The consistent preservation of heightened sensitivity and precision across varying nodule dimensions serves to underscore the resilience of DBNs in the domain of medical image scrutiny. Causey *et al*. [43] nodule-X unveiled as DLCNNs specifically crafted for the assessment of lung nodule

malignancy through CT scans. The researchers developed intricate CNNs structures to analyze CT images from the LIDC-IDRI dataset at varying resolutions, enabling the detection of subtle features associated with malignancy. The model's efficacy was confirmed through independent validation, demonstrating notable levels of accuracy, sensitivity, and specificity, along with a remarkable AUC value, thereby emphasizing the promise of DL in the field of clinical diagnostics.

Hu *et al*. [44]. implemented a Stacked Autoencoder + Softmax strategy, integrating both 2D and 3D data for lung nodule classification. An autoencoder, is a neural network trained to reconstruct its input, consisting of an encoder that maps the input to a latent representation and a decoder that maps the latent representation back to the input space. The softmax layer, typically used for classification tasks, outputs a probability distribution over the target classes. Their approach, leveraging the LIDC-IDRI dataset, achieved outstanding results, including a minimal false positive rate and high accuracy, sensitivity, and specificity.

The integration of 2D and 3D data enabled the model to capture spatial dependencies and volumetric information, crucial for accurate lung nodule classification. In their study, Shaffie *et al*. [45] employed a Deep-Autoencoder for the non-invasive clinical diagnosis of lung nodules. Their model, trained on the LIDC-IDRI dataset, aimed to differentiate between benign and malignant nodules. A deep autoencoder, consisting of multiple hidden layers, can learn hierarchical feature representations, improving classification performance. Their results underlined promising performance metrics, with significant accuracy, specificity, and sensitivity. Wilaiprasitporn *et al*. [46]. analyzed large datasets of chest X-ray images using CNNs to detect anomalies. They assessed three retrained models across varied datasets, including ChestX-ray14 and JSRT, focusing on critical metrics such as accuracy, specificity, and sensitivity. Their study emphasized the importance of retraining models for specific tasks and datasets, with Model C, which amalgamated multiple datasets, displaying superior performance across all metrics. Bhandary *et al*. [47]. used a customized version of AlexNet for identifying lung abnormalities, including cancer and pneumonia. AlexNet, deep CNNs with multiple convolutional and fully connected layers, has demonstrated significant success in image classification tasks. Their DL methodology, applied to the Chest X-ray and LIDC-IDR datasets, achieved high accuracy, demonstrating the adaptability of AlexNet for medical image analysis.

Zhang *et al*. [48] developed a multi-view knowledge-based collaborative (MV-KBC) deep NNs model to classify lung nodules as benign or malignant on chest CT scans. This model integrates multiple views of the input data, leveraging collaborative learning to enhance classification performance. Their approach, validated on the LIDC-IDRI dataset, showed high accuracy and AUC, indicating its potential for integration into clinical workflows. Chauhan *et al*. [49] focused on enhancing lung cancer prediction using routine blood indices through various ML algorithms, including XGBoost, Logistic Regression, SVM, Decision Tree, KNN with Grid Search CV,

and Gaussian Naive Bayes. Feature selection using scikit-learn algorithms retained only the most relevant features, improving model performance. Their XGBoost-based model outperformed others, demonstrating the efficacy of feature selection and ensemble learning techniques in medical prediction tasks. Ayesha *et al*. [50] made significant contributions to lung cancer type classification using feature extraction and fusion techniques, such as Discrete Cosine Transform (DCT) and patch-based Local Binary Pattern (LBP). These techniques capture texture and spatial features in chest CT scans, improving classification accuracy.

ML classifiers, including KNN and SVM, are employed to evaluate the extracted features, yielding high diagnostic precision. Li *et al*. [51]. demonstrated outstanding diagnostic precision through ML-assisted analysis of serum small extracellular vesicles (sEVs). Their approach, involving both pre-and post-surgery serum sEVs, aimed to predict tumor relapse in NSCLC patients. Ezugwu *et al* [52] highlighted the importance of early lung cancer detection with a hybrid CNN architecture combined with the Ebola Optimization Search Algorithm (EOSA). EOSA, inspired by the Ebola virus propagation mechanism, optimizes CNN's hyperparameter, enhancing its performance. Their model, applied to the IQ-OTH/NCCD dataset, achieved high accuracy in classifying CT images, demonstrating the potential of hybrid optimization techniques in medical image analysis.

The results of various studies for cancer detection using different ML and DL models displayed offer a thorough evaluation of various parameters metrics such as accuracy, specificity, sensitivity, and F1 score/AUC detailing various datasets and sample collections consumed in each model with their outcomes from 2017 to 2023. By leveraging optimization algorithms like PSO, transfer learning, curriculum learning, and hybrid models, researchers have achieved remarkable performance metrics, paving the way for improved clinical diagnostics and patient outcomes. The integration of ML, DL, and advanced optimization techniques continues to enhance the accuracy and reliability of medical image analysis, demonstrating the transformative potential of these technologies in healthcare [53].

3. MATERIALS AND METHODS

This section delineates the proposed methodology integrating both DL and ML techniques. Our study utilized two distinct approaches: VGGNet-16 and SVM, conducting a comparative assessment. Fig. (**2**) illustrates our proposed method for lung cancer detection. The proposed method includes the following steps: dataset selection, pre-processing, tumor segmentation, geometric feature extraction, and classification using VGGNet-16 and SVM. In this study, the LIDC-IDRI dataset was chosen for its diversity, comprehensive annotation, and substantial sample size which makes it as a tool for advancing lung cancer detection [54]. In our study, we implemented the VGGNet-16 approach and used the libraries of Keras libraries and the Tensor Flow platform. This setup prepared the environment for developing and evaluating the VGGNet-16 model, including importing essential libraries and defining key components for data processing and model training.
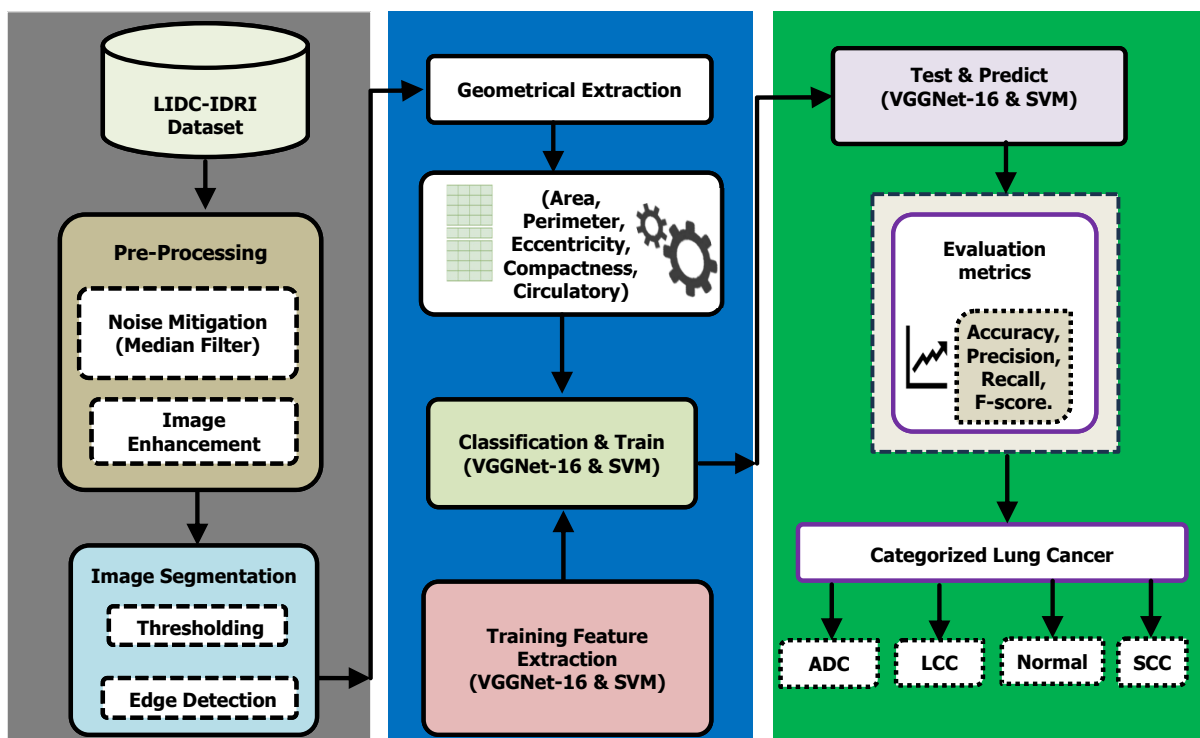

Fig. (2). Organizational diagram of the proposed model.

3.1. Dataset

In the assessment of lung tissue test images, a pertinent dataset is required. LIDC-IDRI dataset identified as an ideal option meticulously curated for a thorough investigation of lung nodules on CT scans. This dataset stands out as a valuable reference, purposefully crafted to augment the capabilities of lung cancer detection and facilitate associated research endeavours which is publically available on kaggle (https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-image accessed on 04 March 2024). The dataset comprises three types of chest cancer: ADC, LCC, and SCC, along with a folder for normal cells. This data is organized within the main folder named Data, which contains subfolders for distinct stages of processing. The dataset's comprises training, testing, and validation, with 70%, 20%, and 10% data allocation, respectively.

3.2. VGGNet-16 Architecture

CNNs are specialized DL models primarily designed for processing visual data like images [5]. They excel the tasks such as image recognition, object detection, and classification [55]. In our work, the VGGNet-16 architecture begins with an input layer designed to accommodate batched RGB images of size 224x224 pixels [56]. This input processed through the pre-trained VGGNet-16 and the result is the feature maps of size 7x7x512. Following the pre-trained VGGNet-16, batch normalization applied to standardize the activation functions to maintain the stability of the model during training. Subsequently, max pooling reduces the spatial dimensions to 3x3x512, emphasizing crucial features [57]. The feature maps flattened into a 1D-vector of size 4608 and facilitate the input to a series of dense layers. These dense layers combined with dropout regularization which progressively reduce the dimensions from 1024 to 128 neurons and serve as a classifier. The final layer comprises a dense layer activated by softmax and serves as an output layer for predictions reflecting the model's assessment of lung cancer presence or absence [58]. Table **1** represents VGGNet-16 model used in our study providing a concise summary of each layer, including its type, output shape, and a number of parameters.

3.3. Categorization of Lung Cancer into 4 Classes Classifications

In our study, we focus on accurately classifying lung tumors into four key classes: ADC, LCC, Normal, and SCC. Each class represents a distinct type of cellular morphology in the lungs, making precise classification essential for targeted treatment. This process utilizes a two-tiered approach combining VGGNet-16 for deep feature extraction and SVM for refined classification, resulting in highly reliable diagnostic outcomes. ADC represents the most common type of non-small cell lung cancer, ADC originates in the outer regions of the lung and is commonly associated with non-smokers. Our model identifies ADC by detecting lung cancer characteristics of glandular patterns. LCC is known for its rapid growth and poor differentiation, LCC often appears in any part of the lung. The model differentiates LCC by its unique cellular appearance, devoid of specific glandular or squamous features. This class represents non-cancerous lung tissue, serving as a

baseline for comparison. Our model's ability to recognize normal lung structures ensures that benign regions are accurately classified, reducing false positives and enhancing overall reliability. Typically related to smoking, SCC originates in the bronchial tubes and features distinctive squamous cells. The model identifies SCC based on these structural patterns, enhancing specificity in diagnosing smoking-related lung cancer.

Table 1. Representation of our VGGNet-16 model used in our study.

Layer (Type)	Output Shape	Parameter	Activation
VGGNet-16(Functional)	(None, 7, 7, 512)	14,714,688	-
Batch Normalization	(None, 7, 7, 512)	2,048	-
Max Pooling2d	(None, 3, 3, 512)	0	-
Flatten	(None, 4608)	0	-
Dense	(None, 1024)	4,719,616	ReLU
Dropout	(None, 1024)	0	-
Dense 1	(None, 512)	524,800	ReLU
Dropout 1	(None, 512)	0	-
Dense 2	(None, 256)	131,328	ReLU
Dense 3	(None, 128)	32,896	ReLU
Dense 4	(None, 4)	516	Softmax

By combining the extensive feature extraction of VGGNet-16 with SVM's discriminative capabilities, the model achieves robust classification across these four tumor types. This approach not only enhances the accuracy of each diagnosis but also contributes to targeted treatment planning by precisely distinguishing between varied cellular compositions within lung tissue.

This conclusive phase of our methodology revolves around the critical task of categorizing lung tumors into 4 classes. This pivotal step is approached through a dual-pronged strategy, integrating the strengths of both VGGNet-16 and SVM [59]. The trained VGGNet-16 model, having undergone meticulous training on an extensive dataset, is deployed to classify new and previously unseen lung tumor images [60]. Leveraging the discriminative capabilities of VGGNet-16, the model categorizes tumors based on features extracted during the training phase. This ensures robust and precise predictions, contributing to the accurate identification of all four classes [61].

3.4. Performance Methods

The purpose of our study is to evaluate the effectiveness of a suggested methodology using the LIDC-IDRI dataset to address the pressing need for early lung cancer identification [62]. The main objective is to create a reliable automated system that accurately classifies lung tumors in order to identify cancers in a timely manner [63]. The training of the models took place on the Kaggle platform, utilizing both its CPU and GPU resources. The storage demands for the models were around 73.1 GB, with RAM usage totaling 29 GB. Computational tasks were handled by two GPU T4 units, each providing 15 GB of memory. Our study employs the VGGNet-16 for feature extraction due to their high accuracy, sensitivity, and specificity in image-based classification. SVM

operated for classification based on geometric characteristics, aiming to create an automated system for accurate lung tumor classification [64]. The mathematical expression of the performance methods are mentioned in Eqs. (**1** to **4**) [65, 66].

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}. \tag{1}$$

$$\text{Recall} = \frac{TP}{(TP+FN)}. \tag{2}$$

$$\text{Precision} = \frac{TP}{(TP+FP)}. \tag{3}$$

$$\text{F1-Score} = \frac{2\,(\text{precision} * \text{recall})}{(\text{precision} + \text{recall})}. \tag{4}$$

where, TP for True Positive, TN for True Negative, FP for False Positive and FN for False Negative.

3.5. VGGNet-16 based Analysis

VGGNet-16 was chosen over VGG-19, AlexNet, and GoogleNet for lung cancer detection due to its balanced depth, efficiency in feature extraction, robust transfer learning capability from ImageNet, manageable computational demands, and extensive community support, making it ideal for medical image analysis tasks for lung cancer detection [67, 68]. The trained VGGNet-16 model deployed for lung tumor classification, autonomously extracting intricate features from lung tumor images without explicit feature engineering. VGGNet-16 deeply handles the complexity and nuanced characteristics within lung tumor images, enabling accurate categorization [69]. This integration serves as a cornerstone in automating the categorization process, providing invaluable insights for medical diagnosis and treatment planning [70].

In our proposed methodology for detecting lung cancer, our innovative SVMVGGNet-16 method synergizes the inherent strengths of VGGNet-16 and SVM. For our study, we selected SVM for lung cancer detection because of its strong performance in high-dimensional spaces, robust resistance to overfitting, ability to handle non-linear classification effectively, and its proven reliability in medical image analysis [71]. These attributes make SVM, a superior choice compared to other ML models like Random Forest [72] and AdaBoost [73]. By capitalizing on SVM's discriminative prowess and VGGNet-16's ability to autonomously extract complex features, our proposed model architecture and methodological stages collectively constitute a robust framework. This framework not only advances the field of lung cancer detection but also fosters in-depth research endeavors, positioning our study at the forefront of cutting-edge methodologies in medical image analysis [74]. The seamless integration of our two models augments the reliability, accuracy, and efficiency of lung cancer identification, underscoring the significance of a multifaceted approach in addressing the intricacies of medical image classification.

4. RESULTS AND DISCUSSION

In our experimental analysis, we evaluated the performance of VGGNet-16 and SVM models across three phases: training, validation, and testing. Then we evaluated the SVMVGGNet-16 model performance for classification of lung tumors across various 4 classes with their performance metrics.

4.1. Pre-processing of Dataset

A median filter is employed to reduce noise in the images, thereby enhancing image quality [75] and facilitating the straightforward identification of lung tumors. The median filter replaced each pixel's value with the median value of the intensities in its neighborhood, effectively reducing salt-and-pepper noise while preserving edges. Image contrast and sharpness improved using histogram equalization, which enhances the visibility of lung tumors. Histogram equalization redistributes the intensity values of the image, enhancing contrast by spreading out the most frequent intensity values [76].

4.2. Segmentation of Lung Tumors

The application of image thresholding is employed to separate lung tumors from the adjacent tissue, facilitating the identification of tumor regions. This methodology entails establishing a specific threshold value, where pixel values above it are categorized as tumor and those below it are categorized as background. The implementation of edge detection algorithms, such as the renowned Canny edge detector, is employed to identify the precise boundaries of the segmented tumors, resulting in a more precise delineation. The Canny edge detector undertakes a multi-step process encompassing noise reduction, gradient computation, non-maximum suppression, and edge tracking through hysteresis.

4.3. Extraction of Geometric Parameters

The area of segmented tumor regions is determined following the same procedure [77], which provides insights into tumor size. The perimeter of tumor regions is quantified to characterize their shape [77]. The eccentricity is computed to ascertain the elongation of tumors [77]. The compactness of the gauge tumor is also calculated following the same procedure [77], indicating proximity to a circular form. The circularity of the tumor was evaluated for geometric classification [77].

4.4. Feature Extraction from Trained Lung Cancer Images

A diverse dataset of lung cancer images, including ADC, LCC, SCC, and normal cells, was compiled. Images are standardized through pre-processing and meticulously labeled. Random sampling ensures a varied subset for robust feature extraction and facilitates comprehensive model training [78]. Employing a pre-trained VGGNet-16 model, the hierarchical features are extracted from chest cancer images and normal cells. VGGNet-16 processes images and generates the feature maps capturing specific textures, shapes, and patterns. Max-pooling layers reduce the dimensionality focusing on discriminative features. Extracted features serve as the inputs of the pre-trained VGGNet-16 model for classification [79].

4.5. Training Phase

In our experimental analysis of training phase, Table **2** and Fig. (**3**) illustrate the performance of VGGNet-16, SVM, and SVMVGGNet-16 models in classifying instances, particularly

focusing on positive cases. VGGNet-16 exhibit remarkable accuracy and high recall, indicating their adeptness in correctly classifying positive instances. However, the relatively lower AUC value raises concerns about their ability to effectively differentiate between positive and negative cases. Conversely, SVMVGGNet-16 demonstrated a training accuracy (97.22%) and a higher AUC (0.9922), suggesting robust classification based on geometric features. Despite slightly lower recall, SVM achieves a higher F1-score (94.48%), underscoring its effectiveness in capturing positive instances accurately.

4.6. Validation Phase

In the validation stage of our work, Table **3** and Fig. (**4**) present the performance metrics of VGGNet-16, SVM, and SVMVGGNet-16 models. VGGNet-16 demonstrates an accuracy (83.33%) and an AUC (0.9285), indicating proficient classification. The recall, precision, and F1-Score (80.00%), (82.32%), and (81.08%), respectively, highlighting the model's ability to accurately identify positive instances. Conversely, SVM achieves accuracy (81.94%) and a higher AUC (0.9405), showcasing its robust classification based on geometric features whereas SVMVGGNet-16 had accuracy (94.72%), an AUC (97.87%), recall (86.67%), precision (91.49%) and F1-score (87.73%). Despite similar recall to VGGNet-16 Net, SVM yields slightly higher precision and F1-Score (85.17%) and (82.23%), respectively. These results underscore the trade-offs between different evaluation metrics and offer insights into the comparative performance of the two models. These results provide insights into the performance of each model during the validation phase, aiding in informed decision-making for model selection and refinement [80].
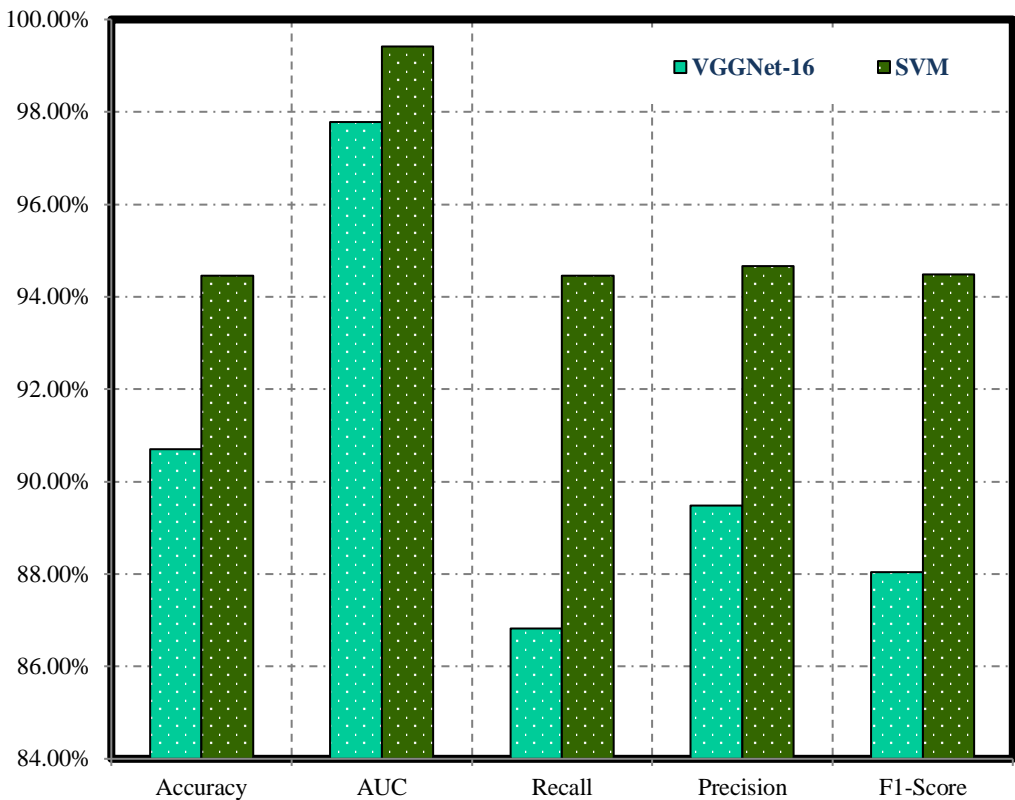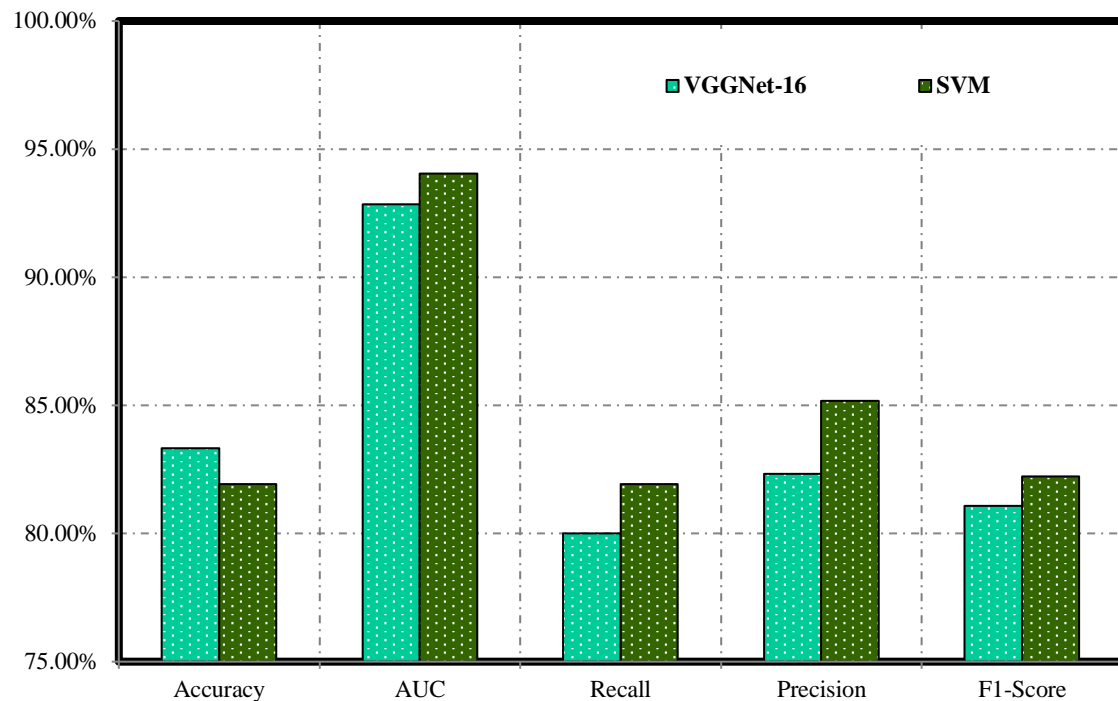

Fig. (3). Evaluating the effectiveness of VGGNet-16 and SVM models during training phase.

Table 2. Results acquired in the training stage.

Models	Accuracy	AUC	Recall	Precision	F1-Score
VGGNet-16	90.70%	97.78%	86.82%	89.48%	88.04%
SVM	94.45%	99.42%	94.45%	94.67%	94.48%
SVMVGGNet-16	97.22%	99.42%	94.22%	95.28%	92.68%

Table 3. Results acquired in the validation stage.

Models	Accuracy	AUC	Recall	Precision	F1-Score
VGGNet-16	83.33%	92.85%	80.00%	82.32%	81.08%
SVM	81.94%	94.05%	81.94%	85.17%	82.23%
SVMVGGNet-16	94.72%	97.87%	86.67%	91.40%	87.73%

Fig. (4). Evaluating the effectiveness of VGGNet-16 and SVM models during the validation phase.

4.7. Testing Phase

In the testing stage of our work, Table **4** and Fig. (**5**) outline the performance metrics of VGGNet-16, SVM, and SVMVGGNet-16 models. VGGNet-16 exhibits an accuracy (84.13%) and an AUC (0.9537), demonstrating proficient classification. The recall, precision, and F1-Score (81.13%), (86.31%), and (83.58%), respectively, indicate the model's ability to effectively identify positive instances. Conversely, SVM achieves an accuracy (84.44%) and a higher AUC (0.9684), showcasing robust classification based on geometric features. Despite similar recall to VGGNet-16, SVM yields slightly higher precision and F1-Score (85.30%) and (84.56%), respectively whereas SVMVGGNet-16 achieved accuracy (96.72%), AUC (96.87%), recall (84.67%), precision (87.40%) and F1-score (85.73%) respectively. These results offer insights into the comparative performance of the two models during the testing phase, aiding in informed decision-making for model selection and deployment.

Figs. (**3**, **4**, and **5**) provide an extensive assessment of DL model used in the detection of lung cancer throughout the training, validation, and testing stages. Throughout all phases, VGGNet-16 consistently outperforms SVM in AUC, highlighting their efficacy in distinguishing between positive and negative instances in detecting lung cancer. The improvement in AUC from training to testing phases for the VGGNet-16 underscores their generalization ability and potential for real-world applications. SVM, while competent, reveals limitations in AUC across validation and testing phases, suggesting areas for refinement to enhance practical performance [81].

Table 4. Results acquired in the testing stage.

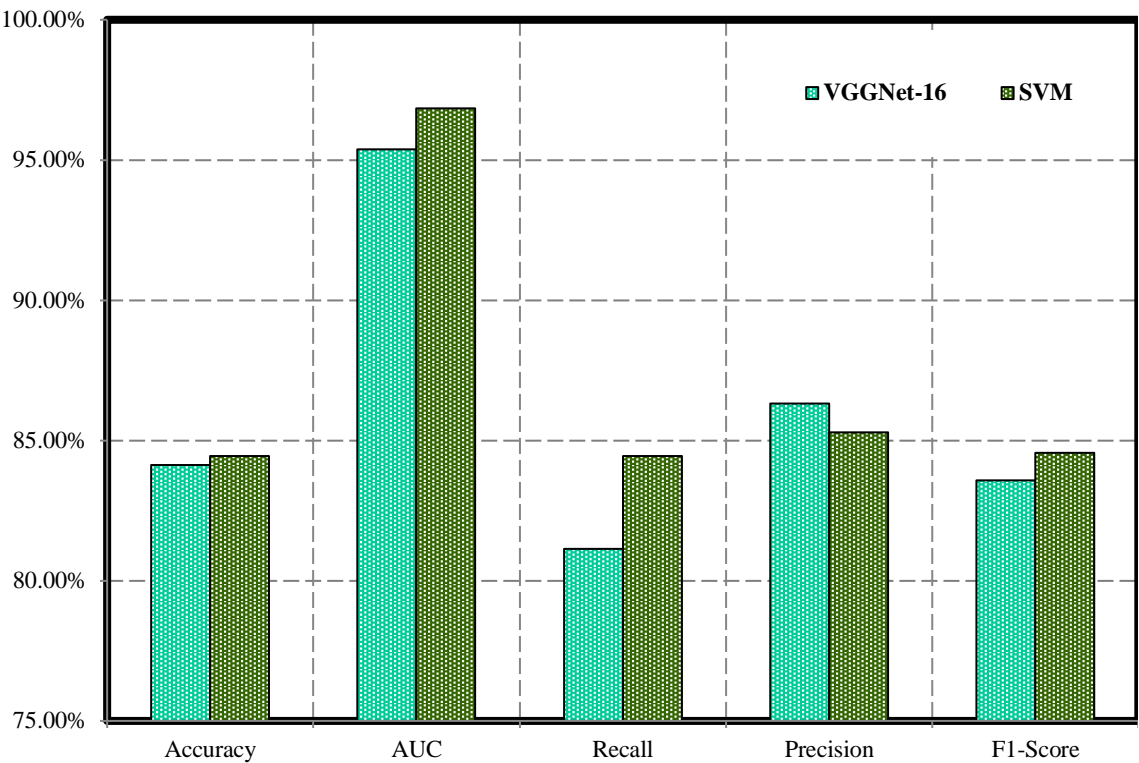| Models	Accuracy	AUC	Recall	Precision	F1-Score
VGGNet-16	84.13%	95.37%	81.13%	86.31%	83.58%
SVM	84.44%	96.84%	84.44%	85.30%	84.56%
SVMVGGNet-16	96.72%	96.87%	84.67%	87.40%	85.73%

Fig. (5). Evaluating the effectiveness of VGGNet-16 and SVM models during testing phase.

On the whole, SVM exhibits an outstanding efficacy during the training phase and delivers competitive outcomes in validation and testing scenarios. SVM particularly shines in managing high-dimensional datasets and performing non-linear classification tasks. Conversely, the VGGNet-16 showcases robust generalization capabilities, rendering both models' feasible options. Nevertheless, SVM hold slight edge in terms of precision and recall, which are crucial for minimizing false positives and negatives, especially in the realm of medical applications. Now, we go with the VGGNet-16 and the SVM proposed SVMVGGNet-16 model for the classification of lung cancer with their geometric extraction in our study.

4.8. SVMVGGNet-16

Fig. (**6**) illustrates the correlation between training accuracy and validation accuracy, alongside training loss and validation loss. The graph designates that training accuracy improves as epochs improvement, implying enhanced classification proficiency of the VGGNet-16 and SVM algorithm with increasing epochs. Fig. (**7**) displays the AUC and Recall of our experiment, providing further perceptions into the model's performance. Finally, the classification is done by confusion matrix and ROC curve respectively, illustrated in Fig. (**8**).
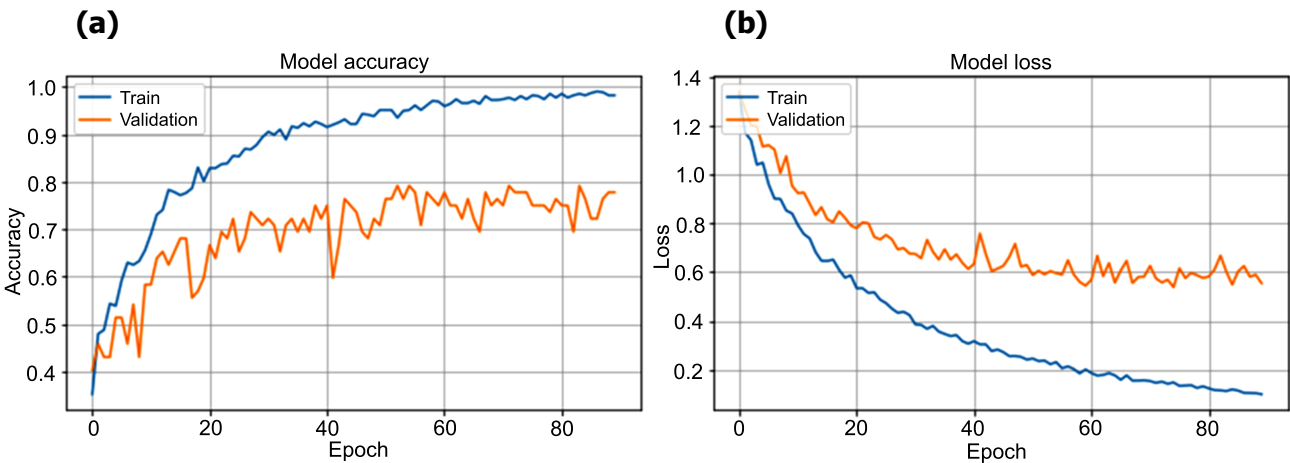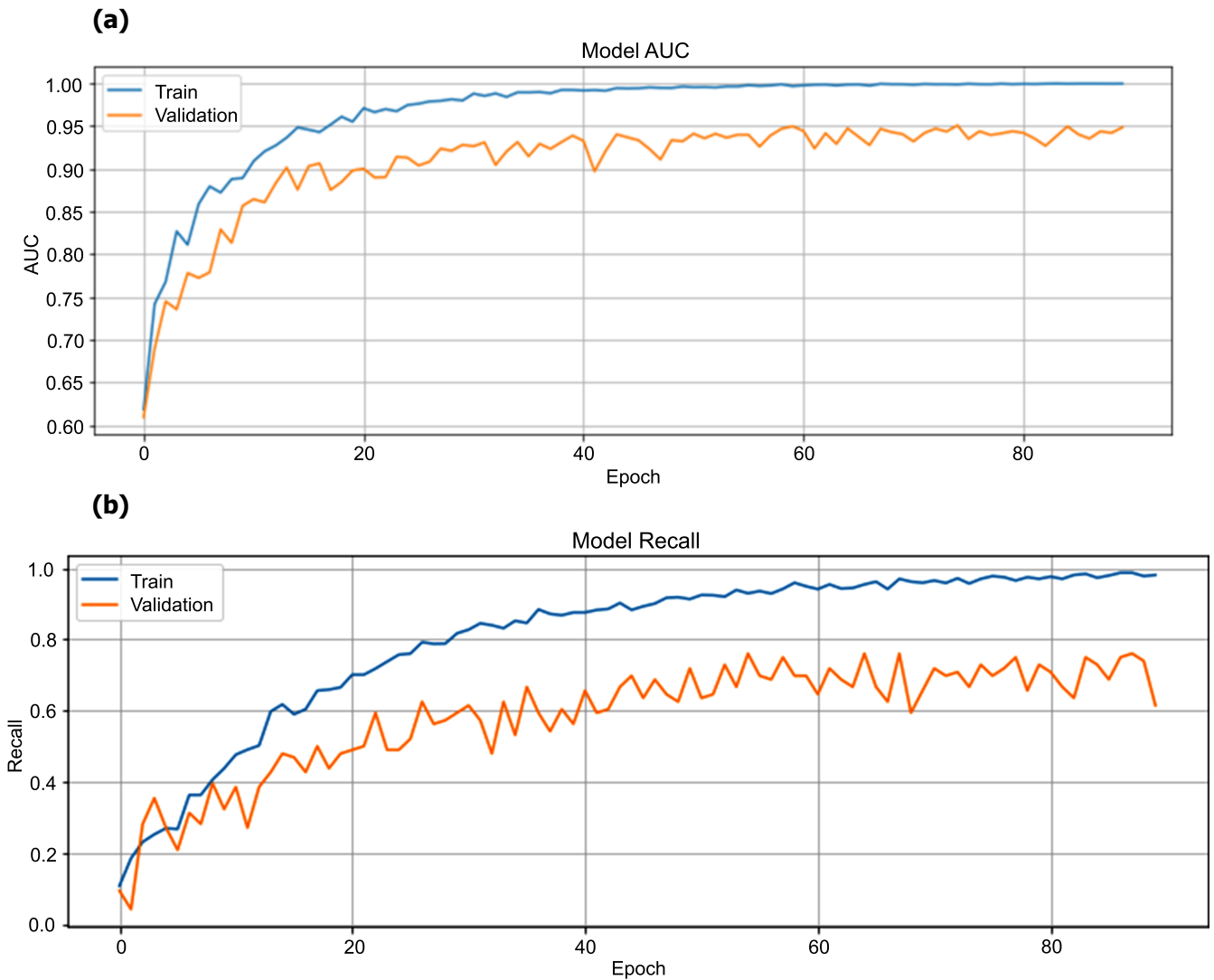

Fig. (6a, b). Model accuracy and model loss graph.

(a)

(b)

Fig. (7a, b). Model AUC and model recall graph.

Table 5. Performance measurements for our experiment during the training phase.

Class	Type of Lung Cancer	Recall	Precision	Accuracy	AUC	F1 Score
0	ADC	0.953846	0.898551	0.95106	0.994205	0.925373
1	LCC	0.913043	0.990566	0.982055	0.994205	0.950226
2	Normal	0.972973	0.993103	0.991843	0.994205	0.982935
3	SCC	0.929032	0.929032	0.964111	0.994205	0.929032
	Average	0.942223	0.952813	**0.972267**	0.994205	0.946892

Table 6. Performance measurements for our experiment during testing phase.

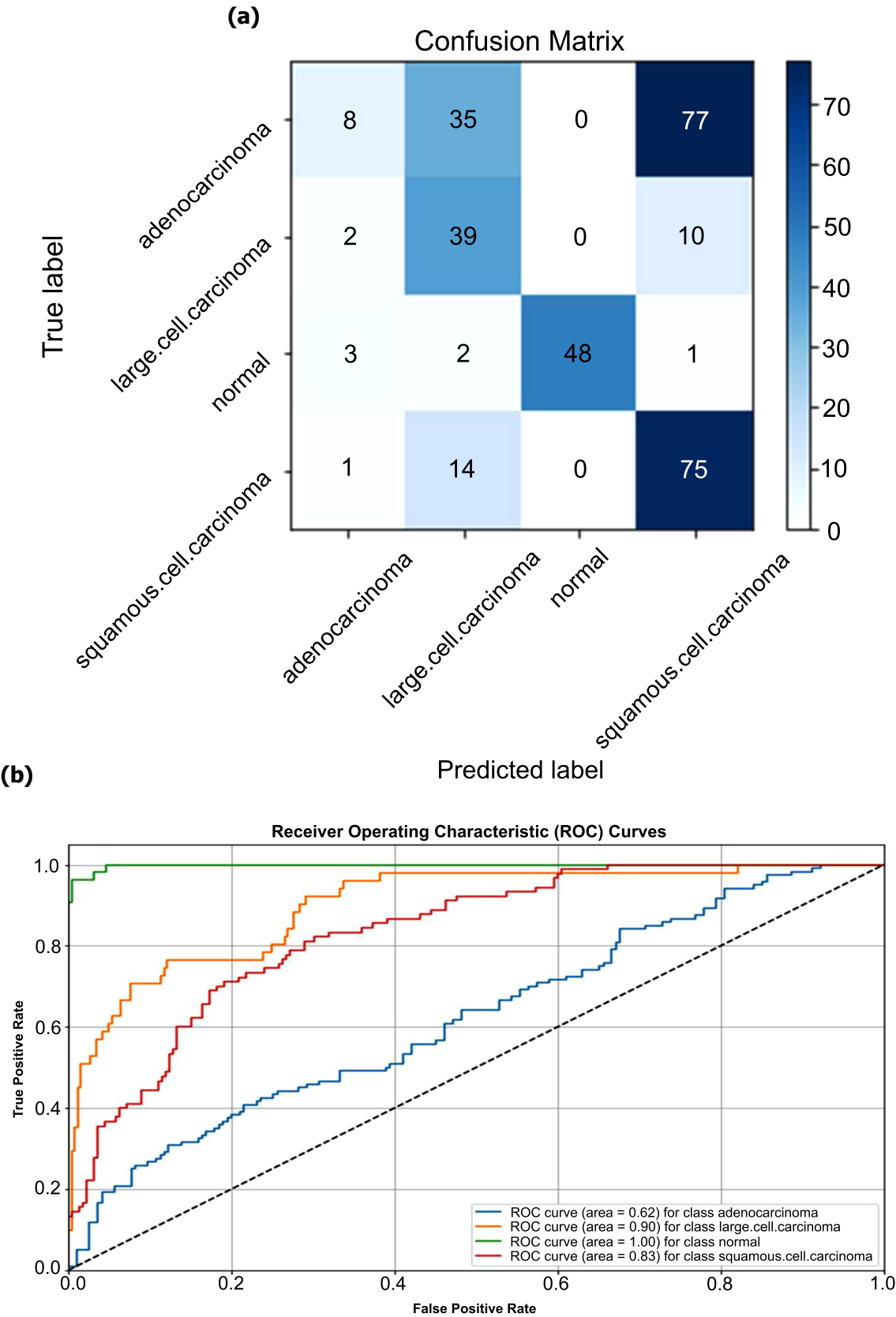| Class	Type of Lung Cancer	Recall	Precision	Accuracy	AUC	F1 Score
0	ADC	0.875000	0.772059	0.963968	0.968799	0.820312
1	LCC	0.882353	0.849057	0.975556	0.968799	0.865385
2	Normal	0.851852	1.000000	0.984603	0.968799	0.920000
3	SCC	0.777778	0.875000	0.944762	0.968799	0.823529
	Average	0.846740	0.874020	**0.967220**	0.968790	0.857300

(a)

(b)

Fig. (8a, b). Confusion matrix and ROC graph for the classification of Lung cancer in 4 classes.

Table 7. Study performance metrics for VGGNet-16, SVM, and SVMVGGNet-16 model.

Model	Phase	Accuracy	AUC	Recall	Precision	F1-Score
VGGNet-16	Training	90.70%	97.78%	86.82%	89.48%	88.04%
	Testing	84.13%	95.37%	81.13%	86.31%	83.58%
SVM	Training	94.45%	99.42%	94.45%	94.67%	94.48%
	Testing	84.44%	96.84%	84.44%	85.30%	84.56%
SVMVGGNet-16	Training	97.22%	99.42%	94.22%	95.28%	94.68%
	Testing	96.72%	96.87%	84.67%	87.40%	85.73%

4.9. Classification during the Training and Testing Phase with their Performance Metrics

Tables **5** and **6** depict the assessment of performance in our study on the identification of lung cancer using VGGNet-16 and SVM classifiers, involving four classes ADC, LCC, Normal, and SCC present in the image dataset operated for both testing and training phases.

4.10. Comparison of Our Approach

In our study, analysis conducted in Table **7** and Fig. (**9**) delineates the performance measures of VGGNet-16, SVM, and the SVMVGGNet-16 model in the context of lung cancer detection throughout both the training and testing phases.

Meanwhile, Fig. (**9**) visually depicts the comparative performance of each model used in our study, VGGNet-16, SVM, and SVMVGGNet-16 models. Clearly, our SVMVGGNet-16 model achieves the highest precision rates of 97.22% during training and 96.72% during testing, emphasizing its superior effectiveness. VGGNet-16 showcases notable generalization capabilities, whereas SVM excels in managing high-dimensional data and executing non-linear classification tasks proficiently. By synergistically integrating these models within a hybrid framework, the reliability and accuracy are heightened, leading to a robust approach for diagnosing lung cancer. The comprehensive performance metrics pertaining to various categories ADC, LCC, Normal, and SCC accentuate the efficacy of each individual model.
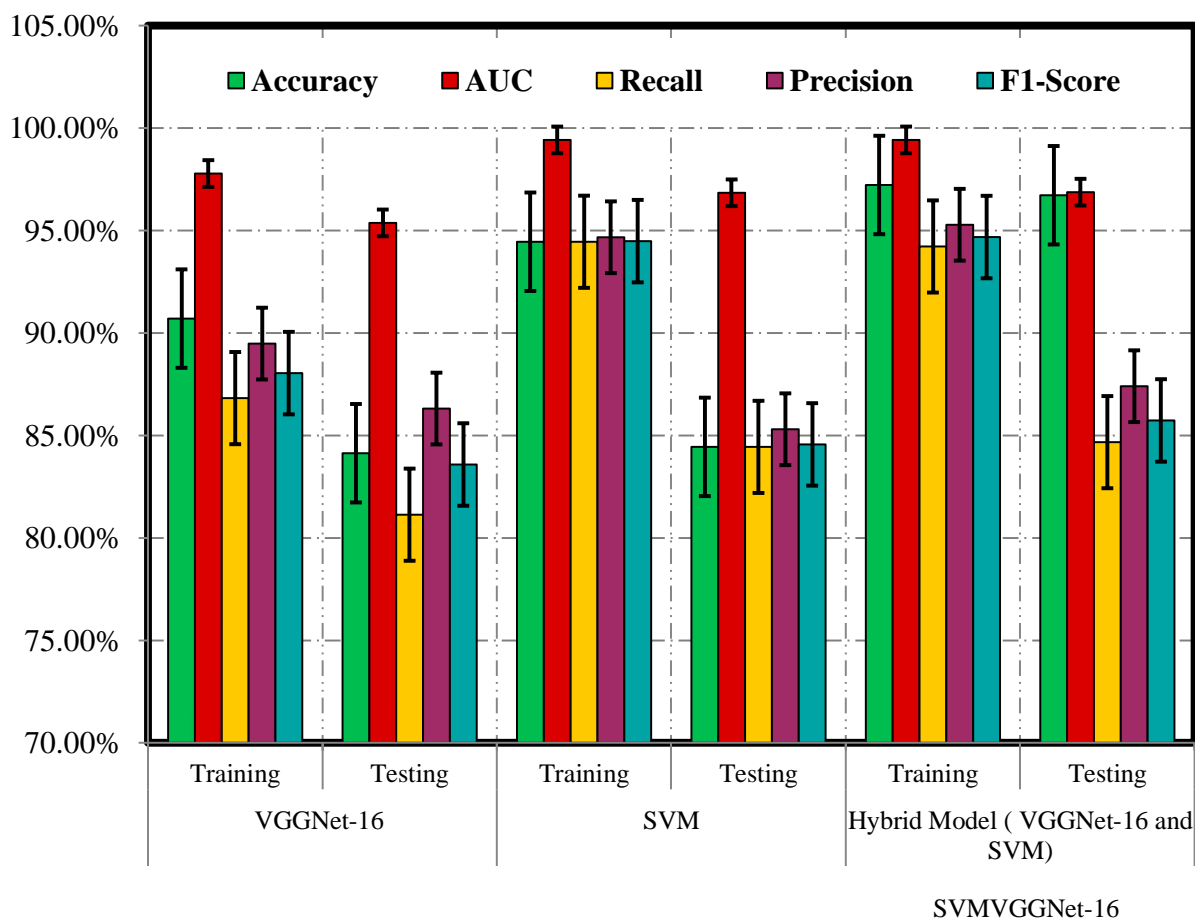

Fig. (9). Performance of various models used in the proposed work.

Table 8. Comparison of our works with previous work.

Year	Dataset	Model	Accuracy (%)
2017 [13]	LIDC-IDRI dataset	ResNet Principles	89.90
2019 [18]	LIDC-IDRI dataset	Deep Autoencoder	91.20
2022 [21]	LIDC-IDRI dataset	MV-KBC Deep Neural Network Model	91.60
Current Work	LIDC-IDRI dataset	SVMVGGNet-16	**96.72**

4.11. Comparison with Previous Work

The progression of lung cancer classification models applied to the LIDC-IDRI dataset throughout various years (Table **8**). Commencing in 2017 with the implementation of ResNet Principles resulting in accuracy (89.9%), there had been a series of advancements leading to increasingly precise models. The year 2019 saw the introduction of Deep Autoencoder model achieving accuracy (91.20%), succeeded by MV-KBC Deep Neural Network Model in 2022 with accuracy (91.60%). In our present study, we introduced the SVM and VGG-16 Net model, showcasing the highest accuracy (96.72%) among all models listed, and signifying notable progress in lung cancer classification. This progression highlights the ongoing refinement and efficacy of DL methodologies in tasks related to the analysis of medical images.

The findings of our study regarding the detection of lung cancer through the utilization of VGGNet-16 and SVM model notable implications for clinical practice, research, and have public health endeavors. By incorporating these models into clinical processes, healthcare professionals can avail themselves of automated assistance in precisely detecting lung tumors from CT scan images, which may result in earlier diagnoses and enhanced patient results. Besides, the inception of computer-guided diagnostic systems that harness these models can enhance diagnostic accuracy and efficiency, ultimately benefiting patients by decreasing diagnostic errors. Additionally, our research contributes to growth in the context of medical image analysis and DL, establishing a basis for further exploration into the detection of lung cancer and customized cure approaches. The implementation of automated detection systems in public health for screening initiatives has the capacity to streamline early detection efforts on a population level, facilitating prompt interventions for individuals at high risk and ultimately enhancing outcomes related to lung cancer on a wider scale [82].

4.12. Limitation of the Study

While the study offers valuable insights into lung cancer detection using VGGNet-16 and SVM models, several limitations should be acknowledged. Firstly, the study's reliance on a single dataset, such as the LIDC-IDRI dataset, may boundary the generalized ability of the outcomes to diverse patient inhabitants or imaging protocols. Also, our study's concentration on CT scan images may overlook the potential benefits of incorporating other imaging modalities, such as MRI or PET scans, which could provide complementary information for a more comprehensive diagnosis. Furthermore, the performance of VGGNet-16 and SVM models may be influenced by factors such as the worth of image pre-processing techniques, the selection of model hyper parameters, and the availability of labeled data for training [83]. Moreover, the study primarily evaluates model performance in terms of accuracy, AUC, recall, precision, and F1-score, neglecting other important aspects such as computational efficiency or interpretability, which are crucial for real-world implementation. Finally, our study's retrospective design may limit its ability to prospectively validate the developed models in clinical settings, warranting further validation studies to assess their real-world utility and impact on patient outcomes.

4.13. Model Effectiveness to Demonstrate the Advantage

The following is a systematic exposition, tailored to our research, regarding the detection of lung cancer across four distinct categories:

- The comparative evaluation of our SVMVGGNet-16 model's efficacy against established methodologies will facilitate the validation of its dependability across diverse lung cancer scenarios, thereby ensuring that metrics such as accuracy, recall, and precision align with the exigencies of real-world clinical practice.
- Through the process of our novelty model, we are able to delineate the areas in which our model demonstrates superior performance, particularly in the identification of early-stage lung cancer nodules, which may enhance the early detection of cases and subsequently improve patient prognoses.
- Our novelty conducting the relative to alternative methodologies can illuminate any deficiencies in performance, thereby enabling us to refine the model and gain insights into specific domains—such as image resolution or noise mitigation—where adjustments may requisite.
- By evaluating the model in conjunction with competing methods, we illustrate its potential applicability within clinical environments, providing a holistic perspective on its benefits in practical use, ranging from accuracy to operational feasibility [84].
- Transparent benchmarking affords healthcare practitioners and researchers a lucid understanding of the model's strengths and weaknesses [28], thereby fostering confidence in its implementation for lung cancer detection.
- Furthermore, our method serves to establish a normative standard, providing a foundational reference for prospective improvements or modifications of our model in response to the evolving landscape of diagnostic requirements [27].

This structured analysis emphasizes the particular mechanisms through which benchmarking can fortify the impact of our model and enhance its preparedness for clinical utilization.

4.14. Application of This Study:

The findings of our research present numerous technical applications within clinical and healthcare settings:

- The SVMVGGNet-16 model represents a robust instrument for radiologists, facilitating the expedited classification of lung cancer variants with augmented diagnostic accuracy [85]. Through the automation of the identification and categorization of specific lung cancer variants—ADC, LCC, SCC, and normal tissues—the model offers considerable assistance for initial evaluations, reducing the likelihood of human error and accelerating diagnostic processes.
- The incorporation of our SVMVGGNet-16 model into automated diagnostic frameworks can promote the early identification of lung cancer, particularly within populations at elevated risk. By embedding this system into hospital infrastructures and specialized oncology screening centers, healthcare practitioners can substantially enhance the diagnostic timeline, enabling prompt intervention and treatment, which is essential for improving survival rates among lung cancer patients [86].
- The synthesis of the VGGNet-16 and SVM in the context of lung cancer detection underscores the tangible influence of DL and ML methodologies on medical imaging practices [87]. This hybrid model functions as an instructive exemplar for medical trainees and practitioners who are keen on sophisticated diagnostic technologies, illustrating how DL can be applied in pragmatic clinical settings while bridging the educational divide between data science and medical practice.
- Upon the completion of further validation and clinical investigations, the proposed model possesses the capacity to be seamlessly integrated into conventional clinical workflows [88]. Such incorporation could enable a high-caliber, economically viable strategy for lung cancer screening and diagnosis, thereby assisting healthcare teams in achieving timely and precise diagnoses with minimal interruptions.

By establishing a framework for accurate, automated lung cancer detection, our study significantly contributes to the technological advancement of diagnostic medicine. The SVMVGGNet-16 model illustrates a promising trajectory for the integration of AI-augmented diagnostic systems within clinical practice, ultimately fostering enhanced patient care and favorable therapeutic outcomes through timely intervention.

CONCLUSION AND FUTURE PROSPECTS

In conclusion, our study presents a comprehensive analysis of lung cancer detection developing VGGNet-16 and SVM, leveraging the LIDC-IDRI dataset. Our experimental results underscore the effectiveness of DL techniques in early diagnosis, with VGGNet-16 outperforming SVM in accuracy and AUC. Notably, our proposed model VGGNet-16and SVM achieved an impressive accuracy (97.22%) in the training phase and (96.72%) in the testing phase, representing better performance in loss minimization. These results emphasize the essential role of advanced DL methods in addressing the critical need for timely lung cancer detection. Moving forward, optimizing the VGGNet-16 training phase and refining SVM's generalization capabilities are key research directions. Fine-tuning model parameters, exploring diverse architectures, and augmenting dataset diversity can enhance VGGNet-16 discriminative capabilities. Additionally, employing hybrid methods and advanced feature engineering techniques may improve SVM's performance. Integrating explainability techniques into models could enhance interpretability, aiding in understanding classification decisions. Continuous validation on diverse datasets and collaboration with medical experts are essential for real-world applicability and reliability. In essence, our study lays a solid foundation for advancing lung cancer identification, contributing to improved initial diagnosis and patient consequences.

As for forthcoming research directions, optimizing VGGNet-16 and SVM performance during the training phase should be a focal point to enhance discriminative capabilities.

- Further refine the hybrid model by exploring other DL architectures and advance SVM kernels to enhance the detection, accuracy, and robustness.
- Incorporate additional datasets with diverse lung cancer types and varying imaging conditions to improve models' generalizability and performance across different populations.
- Conduct extensive clinical trials and collaborate with healthcare professionals to integrate the proposed model into clinical workflows, ensuring practical usability and effectiveness in medical settings.
- Develop detection capabilities, supporting automated systems to assist radiologists and oncologists in building quick and accurate diagnostic decisions during routine screenings.
- Design and implement user-friendly interfaces for automated detection systems, making them accessible and easy to use for medical consultants without specialized technical knowledge.

AUTHORS' CONTRIBUTION

M.M.A. and S.K.: Data curation, validation, software, investigation, visualization, and writing- original draft; M.B.B.H.: Conceptualization, investigation, validation, project administration, resources, funding acquisition, supervision, and writing – original draft; H.U.: conceptualization, investigation, validation, project administration, resources, funding acquisition, supervision, and writing - review & editing; M.A.B.H., S., and S.P.: Conceptualization, investigation, validation, resources, and writing - review & editing; A.A. and T.Z.: Data curation, validation, formal analysis, and writing - original draft. All authors read and agreed for the publication.

LIST OF ABBREVIATIONS

CNNs	=	Convolutional Neural Networks
SVM	=	Support Vector Machines
ADC	=	Adenocarcinoma
LCC	=	Large Cell Carcinoma
SCC	=	Squamous Cell Carcinoma
NAACCR	=	North American Association of Central Cancer Registries
ACS	=	American Cancer Society
DBN	=	Deep Belief Networks
MV-KBC	=	Multi-view Knowledge-based Collaborative
DCT	=	Discrete Cosine Transform
LBP	=	Local Binary Pattern
EOSA	=	Ebola Optimization Search Algorithm

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable

HUMAN AND ANIMAL RIGHTS

Not applicable.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The data used in this study are openly available on Kaggle (https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-image).

FUNDING

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

REFERENCES

[1] Benifa JVB, Chola C, Muaad AY, *et al.* FMDNet: An efficient system for face mask detection based on lightweight model during COVID-19 pandemic in public areas. Sensors 2023; 23(13): 6090.
[http://dx.doi.org/10.3390/s23136090] [PMID: 37447939]

[2] Ukwuoma CC, *et al.* Boosting breast cancer classification from microscopic images using attention mechanism. 2022 International Conference on Decision Aid Sciences and Applications, DASA. Chiangrai, Thailand, 23-25 March, 2022, pp. 258-264.

[http://dx.doi.org/10.1109/DASA54658.2022.9765013]

[3] Ukwuoma CC, *et al.* Detection of oral cavity squamous cell carcinoma from normal epithelium of the oral cavity using microscopic images. International Conference on Decision Aid Sciences and Applications (DASA). Chiangrai, Thailand, 23-25 March, 2022, pp. 29-36.
[http://dx.doi.org/10.1109/DASA54658.2022.9765023]

[4] Mattiuzzi C, Lippi G. Current cancer epidemiology. J Epidemiol Glob Health 2019; 9(4): 217-22.
[http://dx.doi.org/10.2991/jegh.k.191008.001] [PMID: 31854162]

[5] Ansari MM, Kumar S, Tariq U, *et al.* Evaluating CNN architectures and hyperparameter tuning for enhanced lung cancer detection using transfer learning. J Electr Comput Eng 2024; 2024(1): 3790617.
[http://dx.doi.org/10.1155/2024/3790617]

[6] Kalinke L, Thakrar R, Janes SM. The promises and challenges of early non-small cell lung cancer detection: Patient perceptions, low-dose CT screening, bronchoscopy and biomarkers. Mol Oncol 2021; 15(10): 2544-64.
[http://dx.doi.org/10.1002/1878-0261.12864] [PMID: 33252175]

[7] Singh M, Shah C, Patel P. Lung cancer prediction using machine learning models. International Conference on ICT for Sustainable Development. Singapore, Aug 3, 2023, pp. 613-618.
[http://dx.doi.org/10.1007/978-981-99-5652-4_54]

[8] Nooreldeen R, Bach H. Current and future development in lung cancer diagnosis. Int J Mol Sci 2021; 22(16): 8661.
[http://dx.doi.org/10.3390/ijms22168661] [PMID: 34445366]

[9] Qadri SF, Lin H, Shen L, *et al.* CT-Based automatic spine segmentation using patch-based deep learning. Int J Intell Syst 2023; 2023: 1-14.
[http://dx.doi.org/10.1155/2023/2345835]

[10] Nawabi AK, Jinfang S, Abbasi R, *et al.* Segmentation of drug-treated cell image and mitochondrial-oxidative stress using deep convolutional neural network. Oxid Med Cell Longev 2022; 2022: 1-14.
[http://dx.doi.org/10.1155/2022/5641727] [PMID: 35663204]

[11] Mamun M, Mahmud MI, Meherin M, Abdelgawad A. LCDctCNN: Lung cancer diagnosis of ct scan images using cnn based model. 10th International Conference on Signal Processing and Integrated Networks (SPIN). Noida, India, 23-24 March, 2023, pp. 205-212.
[http://dx.doi.org/10.1109/SPIN57001.2023.10116075]

[12] Iqbal MS, Heyat BBM, Parveen S, *et al.* Progress and trends in neurological disorders research based on deep learning. Comput Med Imaging Graph 2024; 116: 102400.
[http://dx.doi.org/10.1016/j.compmedimag.2024.102400] [PMID: 38851079]

[13] Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA Cancer J Clin 2023; 73(1): 17-48.
[http://dx.doi.org/10.3322/caac.21763] [PMID: 36633525]

[14] Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc 2008; 83(5): 584-94.
[http://dx.doi.org/10.4065/83.5.584] [PMID: 18452692]

[15] Williams EM. Centers for disease control and prevention. The A-Z Encyclopedia of Food Controversies and the Law 2010; 1–2: pp. 67-9.

[16] Society AC. Lung cancer statistics | How common is lung cancer. Facts and Figures 2020.

[17] M. Res, C. Ruts, C. R. Hospital, M. Sciences, I. E. Committee, and S. Crh-smims, "Prevalence Rtms," Indian J Med Res 2018; 156: 517-20.
[http://dx.doi.org/10.4103/ijmr.IJMR]

[18] Seguin L, Durandy M, Feral CC. Lung adenocarcinoma tumor origin: A guide for personalized medicine. Cancers 2022; 14(7): 1759.
[http://dx.doi.org/10.3390/cancers14071759] [PMID: 35406531]

[19] Berezowska S, Maillard M, Keyter M, Bisig B. Pulmonary squamous cell carcinoma and lymphoepithelial carcinoma-morphology, molecular characteristics and differential diagnosis. Histopathology 2024; 84(1): 32-49.
[http://dx.doi.org/10.1111/his.15076] [PMID: 37936498]

[20] Atiya SU, Ramesh NVK, Reddy BNK. Classification of non-small cell lung cancers using deep convolutional neural networks. Multimedia Tools Appl 2023; 83(5): 13261-90.
[http://dx.doi.org/10.1007/s11042-023-16119-w]

[21] Quanyang W, Yao H, Sicong W, *et al.* Artificial intelligence in lung cancer screening: Detection, classification, prediction, and prognosis. Cancer Med 2024; 13(7): e7140.
[http://dx.doi.org/10.1002/cam4.7140] [PMID: 38581113]

[22] Anusewicz D, Orzechowska M, Bednarek AK. Lung squamous cell carcinoma and lung adenocarcinoma differential gene expression regulation through pathways of Notch, Hedgehog, Wnt, and ErbB

signalling. Sci Rep 2020; 10(1): 21128.
[http://dx.doi.org/10.1038/s41598-020-77284-8] [PMID: 33273537]

[23] Tan S L, Selvachandran G, Paramesran R, Ding W. Lung cancer detection systems applied to medical images: A state-of-the-art survey. Archives of Computational Methods in Engineering 2024; pp. 1-38.
[http://dx.doi.org/10.1007/s11831-024-10141-3]

[24] Kumar Y, Gupta S, Singla R, Hu YC. A systematic review of artificial intelligence techniques in cancer prediction and diagnosis. Arch Comput Methods Eng 2022; 29(4): 2043-70.
[http://dx.doi.org/10.1007/s11831-021-09648-w] [PMID: 34602811]

[25] Hunter B, Hindocha S, Lee RW. The role of artificial intelligence in early cancer diagnosis. Cancers 2022; 14(6): 1524.
[http://dx.doi.org/10.3390/cancers14061524] [PMID: 35326674]

[26] Heyat MBB, Akhtar F, Munir F, *et al.* Unravelling the complexities of depression with medical intelligence: Exploring the interplay of genetics, hormones, and brain function. Compl Int Syst 2024; 10(4): 5883-915.
[http://dx.doi.org/10.1007/s40747-024-01346-x]

[27] Akhtar F, Heyat BBM, Sultana A, *et al.* Medical intelligence for anxiety research: Insights from genetics, hormones, implant science, and smart devices with future strategies. Wiley Interdiscip Rev Data Min Knowl Discov 2024; 14(6): e1552.
[http://dx.doi.org/10.1002/widm.1552]

[28] Joshua NES, Bhattacharyya D, Chakkravarthy M, Byun YC. 3D CNN with visual insights for early detection of lung cancer using gradient-weighted class activation. J Healthc Eng 2021; 2021: 1-11.
[http://dx.doi.org/10.1155/2021/6695518] [PMID: 33777347]

[29] Patra R. Prediction of lung cancer using machine learning classifier Computing Science, Communication and Security: First International Conference, COMS2. Gujarat, India, March 26–27, 2020, pp. 132-142.
[http://dx.doi.org/10.1007/978-981-15-6648-6_11]

[30] Nazir I, Haq I, AlQahtani SA, Jadoon MM, Dahshan M. Machine learning-based lung cancer detection using multiview image registration and fusion. J Sens 2023; 2023(1): 6683438.
[http://dx.doi.org/10.1155/2023/6683438]

[31] Wu J, Liu C, Liu X, *et al.* Artificial intelligence-assisted system for precision diagnosis of PD-L1 expression in non-small cell lung cancer. Mod Pathol 2022; 35(3): 403-11.
[http://dx.doi.org/10.1038/s41379-021-00904-9] [PMID: 34518630]

[32] Wahid RR, Nisa' C, Amaliyah RP, Puspaningrum EY. Lung and colon cancer detection with convolutional neural networks on histopathological images. AIP Conf Proc 2023; 2654: 020020.
[http://dx.doi.org/10.1063/5.0114327]

[33] Zeeshan HM, Heyat MBB, Hayat MAB, *et al.* Progress and research trends in lumpy skin disease based on the scientometric assessment – A review. Ann Anim Sci 2024; (Sep):
[http://dx.doi.org/10.2478/aoas-2024-0091]

[34] Jian W, Li JP, Haq AU, *et al.* Feature elimination and stacking framework for accurate heart disease detection in IoT healthcare systems using clinical data. Front Med 2024; 11(4): 1362397.
[http://dx.doi.org/10.3389/fmed.2024.1362397] [PMID: 38841592]

[35] Tripathi S, Muhr D, Brunner M, Jodlbauer H, Dehmer M, Streib EF. Ensuring the robustness and reliability of data-driven knowledge discovery models in production and manufacturing. Front Artif Int 2021; 4(June): 576892.
[http://dx.doi.org/10.3389/frai.2021.576892] [PMID: 34195608]

[36] Bi WL, Hosny A, Schabath MB, *et al.* Artificial intelligence in cancer imaging: Clinical challenges and applications. CA Cancer J Clin 2019; 69(2): 127-57.
[http://dx.doi.org/10.3322/caac.21552] [PMID: 30720861]

[37] Huang B, Sollee J, Luo YH, *et al.* Prediction of lung malignancy progression and survival with machine learning based on pre-treatment FDG-PET/CT. EBioMedicine 2022; 82: 104127.
[http://dx.doi.org/10.1016/j.ebiom.2022.104127] [PMID: 35810561]

[38] Davri A, Birbas E, Kanavos T, *et al.* Deep learning for lung cancer diagnosis, prognosis and prediction using histological and cytological images: A systematic review. Cancers 2023; 15(15): 3981.
[http://dx.doi.org/10.3390/cancers15153981] [PMID: 37568797]

[39] Silva dGLF, Neto SOP, Silva AC, Paiva dAC, Gattass M. Lung nodules diagnosis based on evolutionary convolutional neural network. Mult Tools Appl 2017; 76(18): 19039-55.
[http://dx.doi.org/10.1007/s11042-017-4480-9]

[40] Kaur S, Hooda R, Mittal A, Akashdeep , Sofat S. Deep CNN-based method for segmenting lung fields in digital chest radiographs. Commun Comput Inf Sci 2017; 712(July): 185-94.
[http://dx.doi.org/10.1007/978-981-10-5780-9_17]

[41] Nibali A, He Z, Wollersheim D. Pulmonary nodule classification with

deep residual networks. Int J CARS 2017; 12(10): 1799-808.
[http://dx.doi.org/10.1007/s11548-017-1605-6] [PMID: 28501942]

[42] Zhang T, Zhao J, Luo J, Qiang Y. Deep belief network for lung nodules diagnosed in CT imaging. Int J Perform Eng 2017; 13(8): 1358-70.
[http://dx.doi.org/10.23940/ijpe.17.08.p17.13581370]

[43] Causey JL, Zhang J, Ma S, *et al.* Highly accurate model for prediction of lung nodule malignancy with CT scans. Sci Rep 2018; 8(1): 9286.
[http://dx.doi.org/10.1038/s41598-018-27569-w] [PMID: 29915334]

[44] Nasrullah N, Sang J, Alam MS, Mateen M, Cai B, Hu H. Automated lung nodule detection and classification using deep learning combined with multiple strategies. Sensors 2019; 19(17): 3722.
[http://dx.doi.org/10.3390/s19173722] [PMID: 31466261]

[45] Shaffie A, Soliman A, Fraiwan L, *et al.* A generalized deep learning-based diagnostic system for early diagnosis of various types of pulmonary nodules. Technol Cancer Res Treat 2018; 17: 1533033818798800.
[http://dx.doi.org/10.1177/1533033818798800] [PMID: 30244648]

[46] Ausawalaithong W, Thirach A, Marukatat S, Wilaiprasitporn T. Automatic lung cancer prediction from chest X-ray images using the deep learning approach. 11th Biomedical Engineering International Conference (BMEiCON). Chiang Mai, Thailand, 21-24 November, 2018, pp. 1-5.
[http://dx.doi.org/10.1109/BMEiCON.2018.8609997]

[47] Bhandary A, Prabhu GA, Rajinikanth V, *et al.* Deep-learning framework to detect lung abnormality – A study with chest X-Ray and lung CT scan images. Pattern Recognit Lett 2020; 129: 271-8.
[http://dx.doi.org/10.1016/j.patrec.2019.11.013]

[48] Xie Y, Zhang J, Xia Y, Fulham M, Zhang Y. Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT. Inf Fusion 2018; 42: 102-10.
[http://dx.doi.org/10.1016/j.inffus.2017.10.005]

[49] Puneet , A. Chauhan . Detection of lung cancer using machine learning techniques based on routine blood indices. IEEE International Conference for Innovation in Technology (INOCON). Bangluru, India, 06-08 November, 2020, pp. 1-6
[http://dx.doi.org/10.1109/INOCON50539.2020.9298407]

[50] Rehman A, Kashif M, Abunadi I, Ayesha N. Lung cancer detection and classification from chest CT Scans using machine learning techniques. 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA). Riyadh, Saudi Arabia, 06-07 April, 2021, pp. 101-104.
[http://dx.doi.org/10.1109/CAIDA51941.2021.9425269]

[51] Li W, Zhu L, Li K, *et al.* Machine learning-assisted dual-marker detection in serum small extracellular vesicles for the diagnosis and prognosis prediction of non-small cell lung cancer. Nanomaterials 2022; 12(5): 809.
[http://dx.doi.org/10.3390/nano12050809] [PMID: 35269297]

[52] Mohamed TIA, Oyelade ON, Ezugwu AE. Automatic detection and classification of lung cancer CT scans basedondeeplearning and ebola optimization search algorithm. Plos one 2023; 18(8): e0285796.
[http://dx.doi.org/10.1371/journal.pone.0285796]

[53] Li M, Jiang Y, Zhang Y, Zhu H. Medical image analysis using deep learning algorithms. Front Public Health 2023; 11(November): 1273253.
[http://dx.doi.org/10.3389/fpubh.2023.1273253] [PMID: 38026291]

[54] Mridha MF, Prodeep AR, Hoque ASMM, *et al.* A comprehensive survey on the progress, process, and challenges of lung cancer detection and classification. J Healthc Eng 2022; 2022: 1-43.
[http://dx.doi.org/10.1155/2022/5905230] [PMID: 36569180]

[55] Patil A, Rane M. Convolutional neural networks: An overview and its applications in pattern recognition. Smart Innov Syst Technol 2021; 195: 21-30.
[http://dx.doi.org/10.1007/978-981-15-7078-0_3]

[56] Sheriff TMS, Kumar VJ, Vigneshwaran S, Jones A, Anand J. Lung cancer detection using VGG NET 16 architecture. J Phys Conf Ser 2021; 2040(1): 012001.
[http://dx.doi.org/10.1088/1742-6596/2040/1/012001]

[57] Wu H, Gu X. Max-pooling dropout for regularization of convolutional neural networks. Lect Notes Comput Sci 2015; 9489: 46-54.
[http://dx.doi.org/10.1007/978-3-319-26532-2_6]

[58] Maruschak P, Konovalenko I, Osadtsa Y, *et al.* Surface illumination as a factor influencing the efficacy of defect recognition on a rolled metal surface using a deep neural network. Appl Sci 2024; 14(6): 2591.
[http://dx.doi.org/10.3390/app14062591]

[59] Ashwini P. Categorizing the stages of lung cancer using multi SVM

classifier. Int J Res Pharm Sci 2012; 10(3): 2323-8.
[http://dx.doi.org/10.26452/ijrps.v10i3.1472]

[60] Nigudgi S, Bhyri C. Retracted article: Lung cancer CT image classification using hybrid-SVM transfer learning approach. Soft Comput 2023; 27(14): 9845-59.
[http://dx.doi.org/10.1007/s00500-023-08498-x]

[61] Agarwal M, Rani G, Kumar A, K PK, Manikandan R, Gandomi AH. Deep learning for enhanced brain tumor detection and classification. Res Eng 2024; 22: 102117.
[http://dx.doi.org/10.1016/j.rineng.2024.102117]

[62] Faria N, Campelos S, Carvalho V. A novel convolutional neural network algorithm for histopathological lung cancer detection. Appl Sci 2023; 13(11): 6571.
[http://dx.doi.org/10.3390/app13116571]

[63] Alqahtani AF, Ilyas M. An ensemble-based multi-classification machine learning classifiers approach to detect multiple classes of cyberbullying. Mach Learn Knowl Extrac 2024; 6(1): 156-70.
[http://dx.doi.org/10.3390/make6010009]

[64] Mahmud MI, Mamun M, Abdelgawad A. A deep analysis of brain tumor detection from MR images using deep learning networks. Algorithms 2023; 16(4): 176.
[http://dx.doi.org/10.3390/a16040176]

[65] Tahseen Z, Sofi G, Heyat BMB, Jahan N, Rashid A. Exploring obstruction and deobstruent drugs in unani medicine: A comprehensive survey and future directions Altern Ther Health Med 2024; 30(7): 20-7.

[66] Sultana A, Begum W, Saeedi R, *et al.* Experimental and computational approaches for the classification and correlation of temperament (*Mizaj*) and uterine dystemperament (*Su'-I-Mizaj Al-Rahim*) in abnormal vaginal discharge (*Sayalan Al-Rahim*) based on clinical analysis using support vector machine. Complexity 2022; 2022(1): 5718501.
[http://dx.doi.org/10.1155/2022/5718501]

[67] Ukwuoma CC, Qin Z, Heyat BBM, *et al.* A hybrid explainable ensemble transformer encoder for pneumonia identification from chest X-ray images. J Adv Res 2023; 48: 191-211.
[http://dx.doi.org/10.1016/j.jare.2022.08.021] [PMID: 36084812]

[68] Humayun M, Sujatha R, Almuayqil SN, Jhanji NZ. A transfer learning approach with a convolutional neural network for the classification of lung carcinoma. Healthcare 2022; 10(6): 1058.
[http://dx.doi.org/10.3390/healthcare10061058] [PMID: 35742109]

[69] Musthafa MM, Manimozhi I, Mahesh TR, Guluwadi S. Optimizing double-layered convolutional neural networks for efficient lung cancer classification through hyperparameter optimization and advanced image pre-processing techniques. BMC Med Inform Decis Mak 2024; 24(1): 142.
[http://dx.doi.org/10.1186/s12911-024-02553-9] [PMID: 38802836]

[70] Ho Y, Wookey S. The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling. IEEE Access 2020; 8: 4806-13.
[http://dx.doi.org/10.1109/ACCESS.2019.2962617]

[71] Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of support vector machine (SVM) learning in cancer genomics. Canc Genom Prot 2018; 15(1): 41-51.
[http://dx.doi.org/10.21873/cgp.20063] [PMID: 29275361]

[72] Pal R, Adhikari D, Heyat MBB, *et al.* A novel smart belt for anxiety detection, classification, and reduction using IIoMT on students' cardiac signal and MSY. Bioengineering 2022; 9(12): 793.
[http://dx.doi.org/10.3390/bioengineering9120793] [PMID: 36550999]

[73] Sun T, Wang J, Li X, *et al.* Comparative evaluation of support vector machines for computer aided diagnosis of lung cancer in CT based on a multi-dimensional data set. Comput Methods Programs Biomed 2013; 111(2): 519-24.
[http://dx.doi.org/10.1016/j.cmpb.2013.04.016] [PMID: 23727300]

[74] Xu X, Li J, Zhu Z, *et al.* A comprehensive review on synergy of multi-modal Data and AI technologies in medical diagnosis. Bioengineering 2024; 11(3): 219.
[http://dx.doi.org/10.3390/bioengineering11030219] [PMID: 38534493]

[75] Ansari MM, Kumar S. Analysis of various image processing approaches in detecting and predicting lung cancer. Emerging Trends in IoT and Computing Technologies 2023; pp. 519-27.
[http://dx.doi.org/10.1201/9781003350057-76]

[76] Chang YC, Chang CM. A simple histogram modification scheme for contrast enhancement. IEEE Trans Consum Electron 2010; 56(2): 737-42.
[http://dx.doi.org/10.1109/TCE.2010.5505995]

[77] Nageswaran S, Arunkumar G, Bisht AK, *et al.* Lung cancer classification and prediction using machine learning and image processing. BioMed Res Int 2022; 2022: 1-8.
[http://dx.doi.org/10.1155/2022/1755460] [PMID: 36046454]

[78] Anjum S, Ahmed I, Asif M, *et al.* Lung cancer classification in histopathology images using multiresolution efficient nets. Comput Intell Neurosci 2023; 2023(1): 7282944.
[http://dx.doi.org/10.1155/2023/7282944] [PMID: 37876944]

[79] Thelaya AK, Gilal NU, Alzubaidi M, *et al.* Applications of discriminative and deep learning feature extraction methods for whole slide image analysis: A survey. J Pathol Inform 2023; 14: 100335.
[http://dx.doi.org/10.1016/j.jpi.2023.100335] [PMID: 37928897]

[80] Khan USM, Saifullah MK, Hussain A, Azamathulla MH. Comparative analysis of different rainfall prediction models: A case study of Aligarh City, India. Resul Eng 2024; 22(April): 102093.
[http://dx.doi.org/10.1016/j.rineng.2024.102093]

[81] Namdar K, Haider MA, Khalvati F. A modified AUC for training convolutional neural networks: Taking confidence into account. Front Artif Int 2021; 4: 582928.
[http://dx.doi.org/10.3389/frai.2021.582928] [PMID: 34917933]

[82] Wang L. Deep learning techniques to diagnose lung cancer. Cancers 2022; 14(22): 5569.
[http://dx.doi.org/10.3390/cancers14225569] [PMID: 36428662]

[83] Wang P, Fan E, Wang P. Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. Pattern Recognit Lett 2021; 141: 61-7.
[http://dx.doi.org/10.1016/j.patrec.2020.07.042]

[84] Tariq U, Hu Z, Tasneem KT, Heyat BMB, Iqbal MS, Aziz K. ClusterE-ZSL: A novel cluster-based embeddings for enhanced zero-shot learning in contrastive pre-training cross-model retrieval. IEEE Access 2024; pp. 1-1.
[http://dx.doi.org/10.1109/ACCESS.2024.3476082]

[85] Teramoto A, Tsukamoto T, Kiriyama Y, Fujita H. Automated classification of lung cancer types from cytological images using deep convolutional neural networks. BioMed Res Int 2017; 2017: 1-6.
[http://dx.doi.org/10.1155/2017/4067832] [PMID: 28884120]

[86] Dritsas E, Trigka M. Lung cancer risk prediction with machine learning models. Big Data Cog Comp 2022; 6(4): 139.
[http://dx.doi.org/10.3390/bdcc6040139]

[87] Tandon R, Agrawal S, Rathore NPS, Mishra AK, Jain SK. A systematic review on deep learning☐based automated cancer diagnosis models. J Cell Mol Med 2024; 28(6): e18144.
[http://dx.doi.org/10.1111/jcmm.18144] [PMID: 38426930]

[88] Sun Y, Li C, Jin L, *et al.* Radiomics for lung adenocarcinoma manifesting as pure ground-glass nodules: Invasive prediction. Eur Radiol 2020; 30(7): 3650-9.
[http://dx.doi.org/10.1007/s00330-020-06776-y] [PMID: 32162003]