JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 07 Issue 02
2024

Research Article
https://doi.org/10.56979/702/2024

Enhanced Lung Cancer Detection and Classification with mRMR-Based Hybrid Deep Learning Model

Shahroz Zafar[1], Jawad Ahmad[1], Zeeshan Mubeen[2*], and Gohar Mumtaz[1]

[1]Faculty of Computer Science and Information Technology, Superior University, Lahore, 54000, Pakistan.
[2]Riphah International University, Lahore, 54000, Pakistan.
[*]Corresponding Author: Zeeshan Mubeen. Email: zeeshan.mubeen@riphah.edu.pk

Received: May 12, 2024 **Accepted:** August 20, 2024 **Published:** September 01, 2024

Abstract: Among the most common and fatal malignant tumors worldwide is lung cancer (LC). Generally, it has a poorer five-year survival rate than many other well-known tumors. Pulmonary nodules in the lung are an indicator of the most deadly and lethal kind of lung cancer. Improving a patient's chances of survival requires early detection and evaluation of lung cancer. In the field of lung cancer, broad use of deep machine learning techniques has led to notable progress in recent years in reaching high performance in early diagnosis and prognostic prediction. This study presented a novel hybrid deep learning model by applying the method, of two pre-trained deep model architectures ResNet101 and SqueezeNet obtain feature mappings from the CT images in the dataset. Out of the seven deep-learning CNN architectures that were tested, these two were chosen for evaluation based on their superior performance. Minimizing Redundancy and Maximize Relevance (mRMR) is also used to extract the best features from both of models in order to improve the computational efficiency and performance of the proposed technique. As a result, characteristics that have little bearing on accuracy are removed. All features are ranked to create a new set of feature maps. Next, the technique of feature concatenation is implemented. The best feature map is obtained and then classified using SVM and KNN, two machine learning (ML) classifiers. The accuracy of the newly presented hybrid model was 99.09% with SVM. The results of the experiment show that the suggested hybrid model performed exceptionally well in terms of accuracy on the IQ-OTH/NCCD dataset.

Keywords: Hybrid Deep Learning; mRMR; Transfer Learning; Lung Cancer Detection and Classification; Machine Learning.

1. Introduction

Cancer is now known to be a million-person silent killer and a serious global public health concern. As per the World Health Organization (WHO) survey report [1], it is ranked as the second leading cause of death with an estimated 10 million fatalities in 2020. There are currently over a hundred different forms of cancer, and each has its own traits and behaviors [2]. Any age can be affected by cancer, and it can affect any part of the body. Lung, stomach, liver, colorectal, and breast cancers are the most prevalent cancer types [3]. When pulmonary nodules are present, lung cancer (LC) is one of the deadliest and most deadly diseases. Its main reason is frequently related to the lung tissue's cells growing out of control. It is possible that lung cancer symptoms won't be immediately noticeable in the early stages [4]. Therefore, identifying pulmonary nodules accurately and promptly is crucial to increasing the survival rate of patients with lung cancer. As of right now, the most common method for locating pulmonary nodules is to perform chest imaging tests, which include MRIs, CT scans, X-rays, and other imaging modalities. The most often used diagnostic procedure among these is the CT scan, which is appreciated for its low cost and high resolution [5]. Moreover, the high sensitivity of CT scans has led to their widespread adoption for lung cancer screening. Several lung cancer datasets, both from public and private sources, are accessible, aiding in the advancement of techniques for early lung cancer detection. For instance, Zheng et al. [9] investigated the

utilization of maximum intensity projection images to identify lung nodules within CT scan images using CNN, achieving a sensitivity of 94.2% using the dataset known as Lung Nodule Analysis 2016 (LUNA16). We employed the publicly available benchmark dataset, specifically the To assess the suggested design, the Iraq-Oncology Teaching Hospital/National Center for Cancer Diseases (IQ-OTH/NCCD) lung cancer dataset will be used.

Primary lung cancer can be histologically categorized into two main classes: Non-Small Cell Lung Carcinoma (NSCLC) and Small Cell Lung Carcinoma (SCLC). NSCLC accounts for roughly 80% of diagnoses, while SCLC constitutes about 20% of all lung cancer cases. Within NSCLC, tumors are further divided into three subclasses: Large cell cancer (LCC), squamous cell carcinoma (SCC), and adenocarcinoma [6]. This piece concentrates on non-small cell lung cancer, or NSCLC. Analyzing medical images for the classification of lung cancer comprises two primary methods: algorithms for machine learning capable of learning to classify cancer by utilizing extracted image features, and the development of deep learning models specifically personalized for image analysis [7]. Convolutional neural networks (CNNs) are the most often utilized neural networks for image and video categorization applications. CNNs are particularly favored because of their ability to create hierarchical learning representations from data. Transfer learning (TL) stands as another extensively applied method aimed at addressing the CNN model's shortcomings. Pre-trained CNN models in TL, frequently serve as a starting point, with enhancements made by further training the model using a new dataset. This process involves unfreezing certain layers within using the fresh data to train the already-trained model. Alternatively, it can entail incorporating additional layers into the model that has already been trained and training the model from beginning to end. Furthermore, the catalog of valuable studies could be significantly expanded, indicating from these studies that TL methods exhibit commendable performance [8].

This study introduces a hybrid method for classifying lung cancer from CT scans. Initially, feature extraction was conducted utilizing two distinct pre-trained CNN architectures. These extracted features were reduced by utilizing mRMR, a novel approach for optimal feature selection and then optimal features were categorized by two ML classifiers, and their outcomes were analyzed.

A feature map of dimensions 1097 × 3048 was generated by merging features from two architectures, as ResNet101 contributing a 1097 × 2048 features and SqueezeNet 1097 × 1000. Through the utilization of mRMR feature optimizer, the number of features within this map was condensed to 1097 × 1000. Subsequently, these optimized features were subjected to classification by two distinct machine learning (ML) algorithms: SVM and KNN. Figure 1 shows the suggested hybrid model's workflow. This hybrid approach serves to mitigate potential manual errors, thereby reducing doctors' workload in diagnosing lung cancer. Furthermore, it eliminates the necessity for specialized expertise in initial diagnosis.

The main contributions of the proposed study are as follows.

- We conducted preprocessing on the publicly available benchmark dataset IQ-OTH/NCCD to minimize computational time.
- We proposed an innovative hybrid transfer learning framework, employing ResNet101 and SqueezeNet for feature extraction.
- The feature map was reduced by utilizing mRMR, feature reduction to enhance the model's performance.
- The introduced hybrid model exhibits exceptional performance in classifying lung cancers using a limited set of selected features.
- To describe the effectiveness of the proposed hybrid model over alternative classification models, we conducted a comparison, evaluating the novel-designed model against other sophisticated methods based on their execution time and computational complexity.
- Two different classifiers were used for feature extraction, feature fusion, dimension reduction, and classification in the suggested approach. Based on the experimental outcomes employing the SVM classifier, an impressive accuracy of 99.09% is attained.

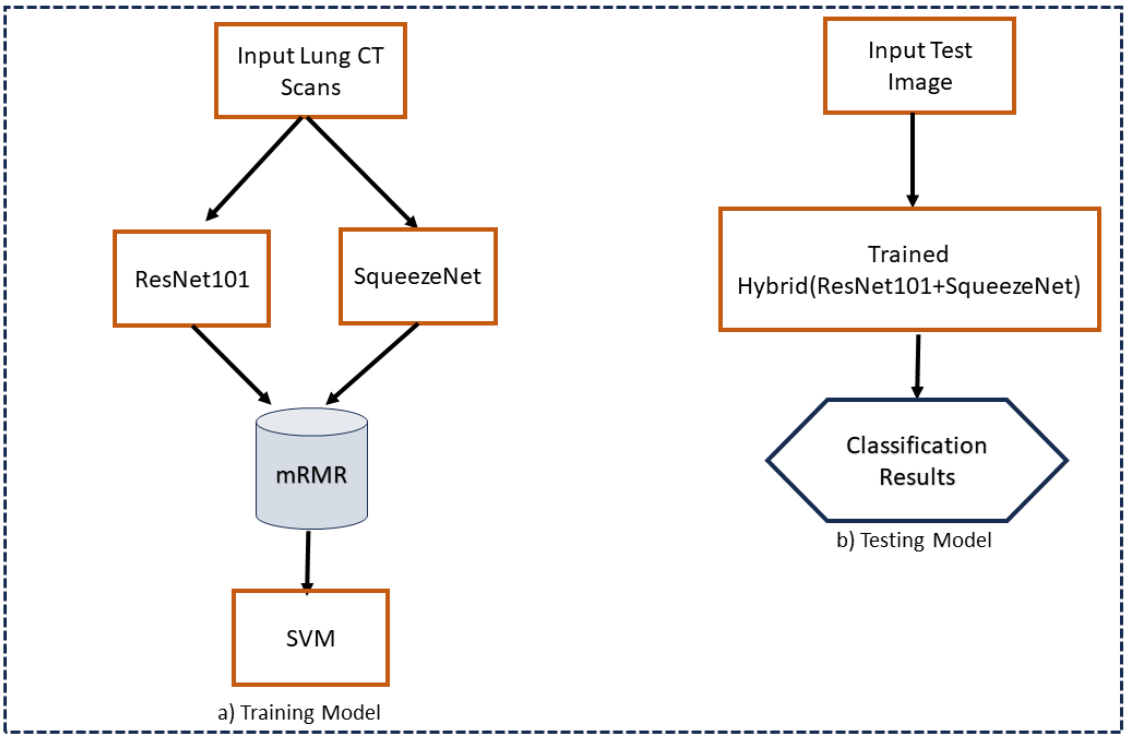


Figure 1. The proposed hybrid model's workflow

The next sections of the article are structured as follows: Section 2 encompasses relevant literature in the realm of lung cancer detection and classification. Section 3 presents a comprehensive breakdown of the organizational framework of the proposed hybrid methodology. The experimental findings compared with state-of-the-art methods are detailed in Section 4. Section 5 delves deeply into the comprehensive exploration and comparison of the proposed method using previous research that is available in the literature. In conclusion, Section 6 presents the results, emphasizing the developed approach.

2. Related Work

This section reviews recent research on deep learning for lung cancer diagnosis and nodule detection, utilizing sources like Elsevier, Springer, Scopus, and Web of Science. We selected 20 relevant papers from 2020 to 2024. Surekha Nigudgi et al. [28] proposed a hybrid model using AlexNet, VGG, and GoogleNet for feature extraction and SVM for classification, achieving 97% accuracy on 1190 CT images. Chen et al. [12] developed a chatbot employing CNN and NLP to distinguish lung nodules, achieving 88% validation accuracy with a Hierarchical Model (HBAM). Thakur et al. [10] summarized CAD systems using CNNs for nodule detection, discussing models like MVCNN and 3D CNN. Lai et al. [11] constructed a DNN for predicting NSCLC patient survival with 75.44% accuracy and an AUC of 0.8163. Jiang et al. [13] introduced a 3D dual-path network for pulmonary nodule classification, utilizing contextual and spatial attention. Riquelme and Akhloufi [14] [44] discussed deep CAD methods aiding radiologists in nodule identification. Tekade and Rajeswari [15] developed a 12-layer CNN for classifying lung cancer, comparing it with VGG16 and Inception-V3. Chaunzwa and associates. [17] used radiomics and CNNs to forecast lung cancer histology. Blanc-Durand et al. [18] explored deep learning for anthropometric measurements in NSCLC prognosis. Raza et al. [19] introduced a DECNN for classifying lung and liver cancer, achieving 99.8% accuracy on static data.

Dodia et al. [20] presented the RFR V-Net for lung nodule classification, combining SqueezeNet and ResNet. Sun and Pang [21] introduced a transformer-based segmentation approach for lung cancer analysis. Lyu [22] proposed an ensemble of four DCNN architectures, achieving 99% test accuracy. Nanglia et al. [23] explored hybrid approaches combining SVMs and NNs for lung cancer classification. Data scarcity and computational resource limitations pose challenges in deep learning for medical imaging, often mitigated by transfer learning. This approach, using pre-trained models, is effective in addressing these constraints.

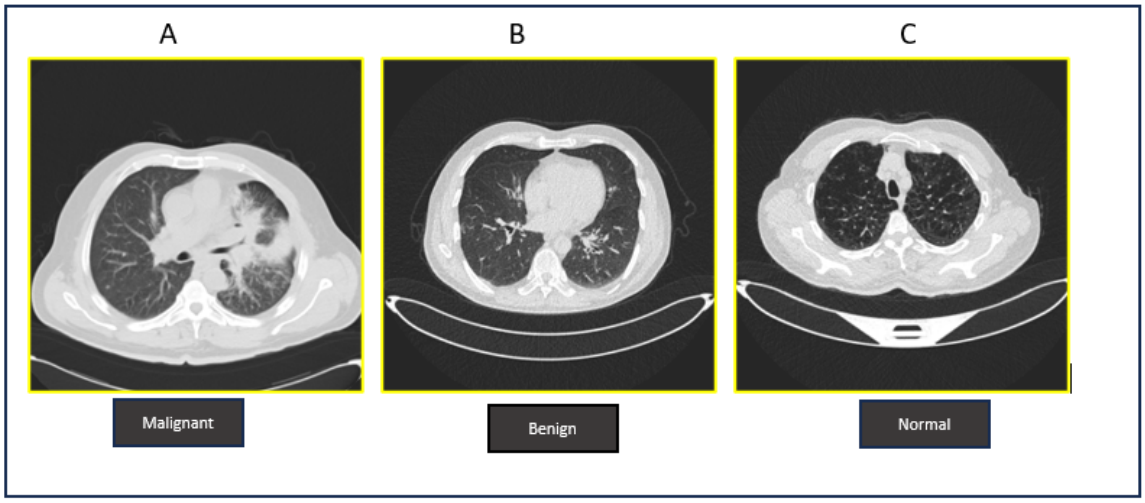


Figure 2. Sample image of each class from the dataset.

3. Materials and Methods

This section provides a comprehensive exploration of the dataset employed and all methodologies applied in constructing the proposed architecture. The architecture encompasses several components, including pre-processing of CT scans of lung images, extracting deep features from pre-trained CNNs, utilizing the mRMR approach, and employing machine learning classifiers.

3.1. Dataset

The used dataset, collected over more than three months in 2019 from the "Iraq-Oncology Teaching Hospital and the National Center for Cancer Diseases," comprises CT scans of individuals. The dataset, which included 1097 human chest images representing CT scan slices from 110 cases, was annotated by a number of radiologists and oncologists. The scans showed variations in age, gender, educational attainment, place of residence, and living arrangement. These patients were categorized as benign, malignant, and normal, as shown in Fig. 2. 110 instances in all were included in the investigation; Fifty were declared cancerous, fifteen benign, and fifty normal. The Kaggle repository makes the dataset publicly accessible (Anon, 2023b). Table 1 thoroughly analyzes the dataset broken down by class distribution.

Table 1. Class-wise distribution.

Class	Samples	Patients
Malignant	561	40
Benign	120	15
Normal	416	55
Total	1097	110

3.2. Data Preprocessing

The preprocessing steps that were performed on the dataset before the model was trained and tested are described in this section. To guarantee impartial instruction, each class's photos are first randomly shuffled. They are then divided into 80:20, meaning that 80% of the total images are part of the training set, and the remaining 20% are part of the testing set. Table 2 summarizes the classes' distribution within the train-test split. These undesirable regions were eliminated by cutting the original CT scan's extreme points off pictures since they had background noise, which could cause noisy training.

Table 2. 80:20 ratio for train-test-split.

Class	Split	Samples	Total
Malignant	Train	448	876
Benign			96
Normal			332
Malignant	Test	113	221
Benign			24

Normal	84

The images within both the training and testing sets are subsequently resized to a consistent resolution, aligning with the required input tensor shape for the pre-trained used models. This resizing process serves the purpose of maintaining the contextual information and features within an image while also matching the model's input specifications, thus assisting in reducing computational load during model training. Furthermore, this resizing operation aids in preserving the image's essential characteristics, alleviating computational burdens during model training. Lastly, the class labels within both the training and testing sets undergo label encoding, representing “malignant," "benign," and "normal" as 1, 0, and 2, respectively.

3.3. Proposed Hybrid Model

The primary variables utilized for classifying the images into distinct categories are the image features. Consequently, the extraction of crucial features, which exhibit diverse characteristics found in lung CT images, holds significant importance in augmenting the classification performance. Manual extraction of these features from lung CT images tends to consume a substantial amount of time. Its accuracy heavily relies on the considerable variability present in lung CT images. When the size of the dataset is restricted, it is beneficial to use a variety of pre-trained CNNs, including ShuffleNet, ResNet-50, DarkNet-53, DenseNet-201, NASNet-Large, Xception, NASNet-Mobile, EfficientNet-b0, DarkNet-19, GoogLeNet365, ResNet-101, GoogLeNet, Inception-ResNet-v2, ResNet-18, AlexNet, MobileNet-v2, and SqueezeNet are among the networks.

The concept of transfer learning is the use of pre-trained networks. These pre-trained networks are employed to extract deep features, especially when dealing with limited data resources [29,30]. The main concept of transfer learning is depicted in Fig 3. While pre-trained CNN architectures may exhibit varying numbers of layers, their core structure remains predominantly consistent. Among these models' crucial layers, the convolutional layer holds important significance as it generates feature maps based on its outputs. This layer operates by maneuvering filters across the image to generate input features. The quantity of filters in each convolutional and input layer can differ across models. Equation no.1 is formulated to elucidate the resulting image dimensions after the convolutional layer's filtering process. As shown in Figure 4, the total number of pictures in the lung cancer CT scan dataset and their corresponding feature matrices determine how many features are derived from these models

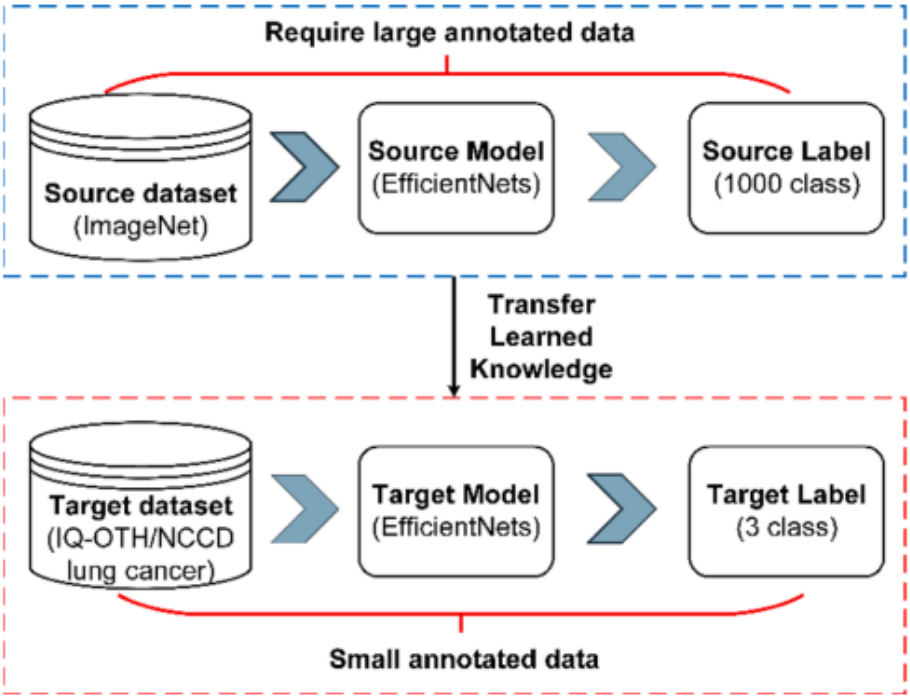


Figure 3. Transfer Learning Process

The feature matrix of ResNet101, is of size x 1000, whereas the feature matrix of SqueezeNet, another DL architecture, is of size x 1000. Consequently, the feature matrix becomes 1160 × 1000, and 1160 × 1000 for ResNet101 and SqueezeNet, respectively. The concatenated feature matrix size of 1160 × 2000 for each

image is the outcome of combining the retrieved features from these two architectures to create a hybrid model.Initially sized at 1160 × 2000, the feature matrix undergoes reduction using the mRMR, the optimal feature selection method. Although the count of valuable features remains unchanged, the overall feature quantity is minimized to 1000. Subsequently, the resulting feature matrix, sized at 1160 × 1000 and comprising the most crucial features, is independently used for classification employing ML classifiers like SVM and KNN. The complete proposed architecture is illustrated in Fig. 5 of the proposed model.

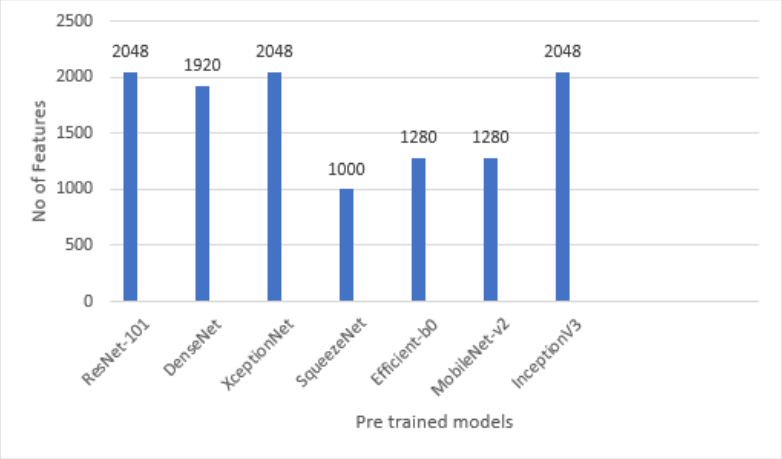


Figure 4. Total no of extracted features from each DL model.

$W = ((u - k) + 2p)/s + 1$ (1)

Where W, represents the output u represents input and k, the filter size. s and p represent step count and padding.

The convolution action is performed in Eq. (2):

$$\boldsymbol{O(m,n) = (I * K)(m,n) = \sum_p \sum_q I(p,q)K(m - p, n - q)} \quad (2)$$

Where O represents output, I represent input and k, kernel.

Most popular activation function ReLU also depicted in eq. (3):

$$\text{ReLU: } \boldsymbol{f(x) = \{0,\ x < 0\ x,\ x \geq 0\ ,f(x)' = \{0,\ x < 0\ 1,\ x \geq 0} \quad (3)$$

3.4. mRMR method

Extensive datasets are a result of technological advancements that create difficulties in extraction and classification of their attributes. The mRMR approach serves the purpose of selecting the most significant features from the array of available options (Tantisatirapong et al., 2014). This method aims to minimize computational costs, striving to achieve optimal correlation between variables and attributes.

$$\text{I(U; V)= } \boldsymbol{\sum_{m \epsilon X} \sum_{n \epsilon Y} p(m,n) log \frac{p(m,n)}{p(m)p(n)}} \quad (4)$$

The variables in the equation above are denoted by U and V, and the probability density function of these variables is shown by p(m, n).

$$\boldsymbol{Max\ D(S,c), D = \frac{1}{|s|} \sum_{xi\ \epsilon\ S} I(xi; c)} \quad (5)$$

$$\boldsymbol{Min\ R(S), R = \frac{1}{|s|^2} \sum_{xi\ ,xj\ \epsilon S} I(xi; xj)} \quad (6)$$

To simultaneously maximize D and R, utilize equations (7) and (8).

$$\max \Phi(D, R),\ \Phi = D - R \quad (7)$$

$$\max \Phi(D, R),\ \Phi = D/R \quad (8)$$

$$\text{Xj } \boldsymbol{\epsilon\ X - S_{m-1}} \left[\boldsymbol{I(xi; c/ \frac{1}{m-1} \sum_{xi, S_{m-1}} I(xj; xi)} \right] \quad (9)$$

$$\text{Xj } \boldsymbol{\epsilon\ X - S_{m-1}} \left[\boldsymbol{I(xi; c/ \frac{1}{m-1} \sum_{xi, S_{m-1}} I(xj; xj)} \right] \quad (10)$$

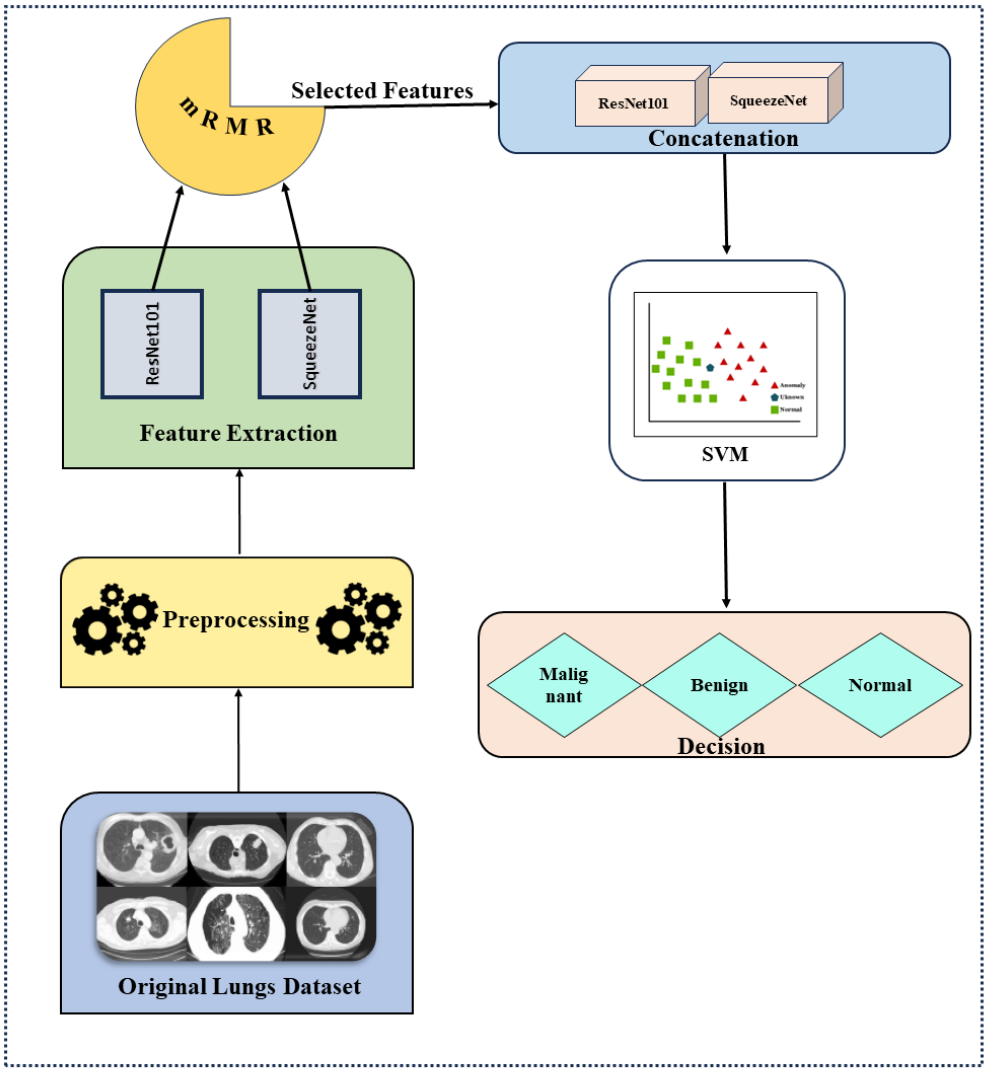


Figure 5. The Proposed Model

4. Experimental Setup and Results

Table 3 presents a comparison of the experimental accuracy results for pre-trained CNN architectures employing two different classifiers. The experimental results of the hybrid model, including confusion measures and performance indicators, are also thoroughly analyzed in this part. The table presents the findings of the categorization of seven CNN architectures, each of which was analyzed separately with the help of the two classifiers.

Table 3. Performance of seven deep learning architectures.

Models	SVM	Models	KNN
ResNet101	98.64	ResNet101	97.74
SqueezeNet	98.64	SqueezeNet	92.34
XceptionNet	97.64	XceptionNet	95.04
DenseNet-201	97.74	DenseNet-201	96.84
EfficientNet-B0	81.08	EfficientNet-B0	59.90
MobileNetV2	96.84	MobileNetV2	97.29
InceptionV3	97.74	InceptionV3	93.24
Proposed Hybrid Model	**99.09**	**Proposed Hybrid Model**	**98.64**

The following assessment measures were used in this investigation.

Accuracy: Accuracy in terms of true positives and false positives refers to the proportion of correctly identified instances (true positives and true negatives) out of the total instances evaluated. It measures how well a test or model correctly identifies both the presence (true positive) and absence (true negative) of a condition, while minimizing incorrect identifications (false positives and false negatives).

accuracy = (TP+TN)/(TP+TN+FP+FN).
Here, the terms "false positive," "false negative," "true positive," and "true negative" are used, respectively.

Precision: The ratio of exact positive outcomes is called precision, and it is calculated using equation 12.
Precion=TP/(TP+FP) (12)

Recall/Sensitivity: Recall signifies the ratio of correctly identified actual positive cases and is determined by Equation 13.
Recall/Sensitivity=TP/(TP+FN) (13)
Specificity: Specificity is a metric used in binary classification that measures the model's ability to correctly identify the true negative cases among all the actual negative cases in the dataset.
Specificity=TN/(TN+FP) (14)

F1-Score: The F1 score is an equilibrium-based harmonic mean involving precision and recall, derived from Equation 15.
F1-Score=2TP/(2TP+FP+FN) (15)

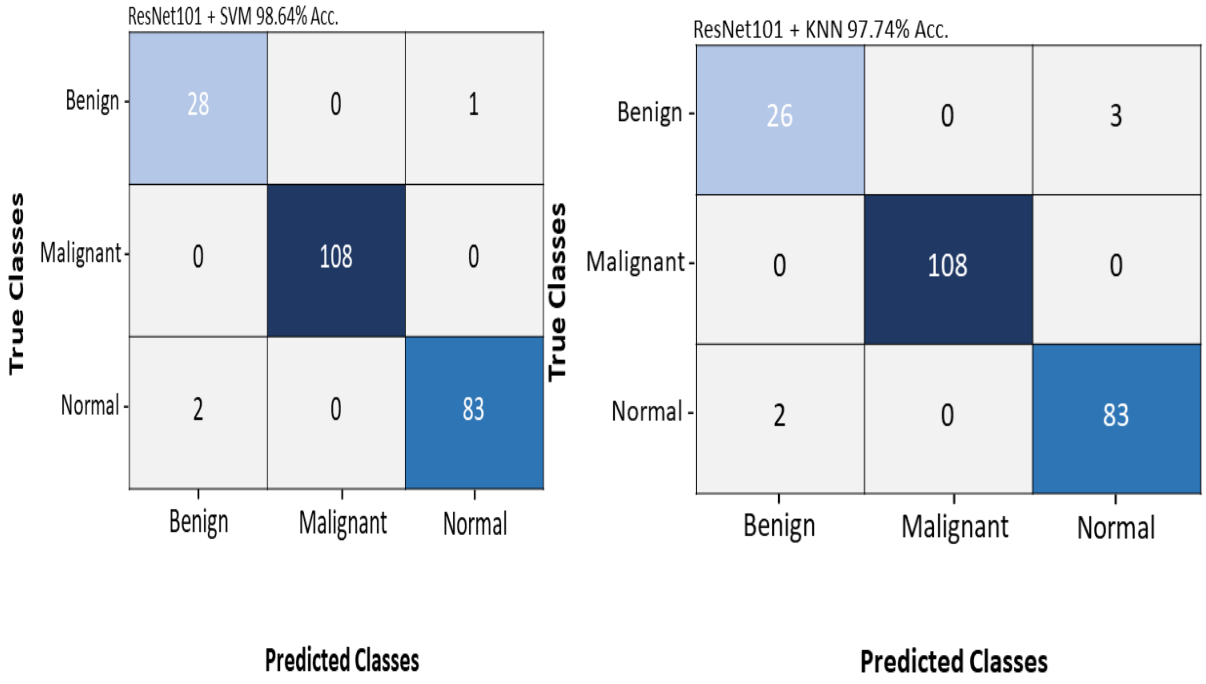
4.1. Experimental Setup

The Google Colab Pro framework is used to implement the hybrid model that this study suggests, providing accelerated training and model assessment, which proves beneficial for the development and testing phases. The experimental configuration utilized this study's outline is as follows: Model training makes use of TensorFlow as the backend, the Keras library, and Python programming. A Tesla T4 GPU with 25 GB of RAM that is available on Google Colab Pro is used for the experiments.

Performance investigation revealed that, as shown in Table 3, ResNet101 and SqueezeNet with SVM had the highest accuracies, reaching up to 98.64%. We used features, derived from ResNet101 and SqueezeNet architectures for our hybrid model and classified those using SVM and KNN ML algorithms. Our proposed hybrid method achieved 99.09% accuracy by using mRMR as the optimal feature selector and SVM as a classifier. Which is a better result than various methods available in recent research.

4.2. Performance Metrics

The suggested methodology is evaluated and compared in this study using five distinct performance evaluation metrics: accuracy, precision, sensitivity, specificity, and F1-measure. Figure 6 visually represents the confusion matrices generated by the SVM and KNN classifiers for the ResNet101, SqueezeNet and hybrid models.



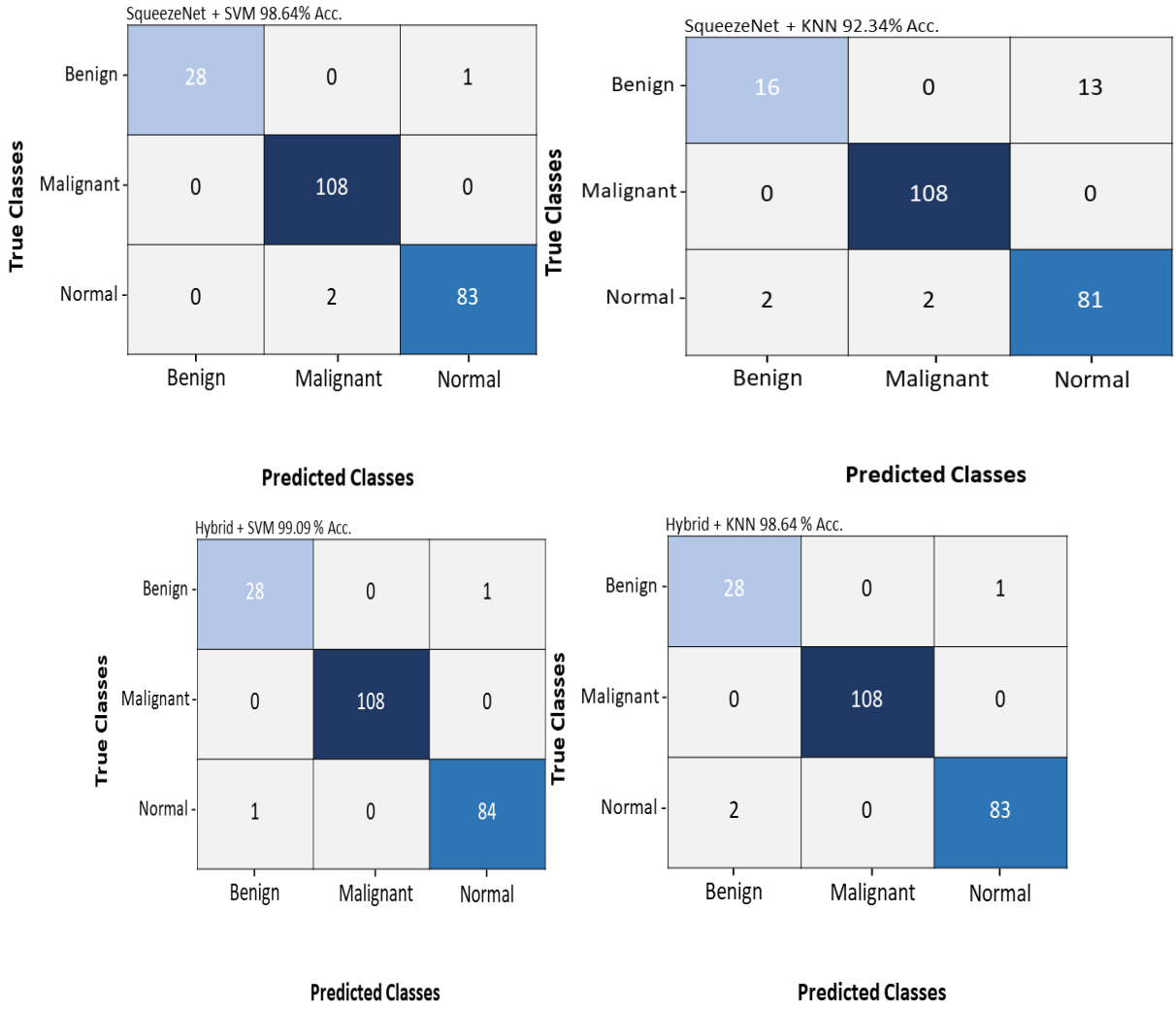


Figure 6. Confusion Matrices

Table 4. Performance of Hybrid model

Model	Classes	Accuracy	Sensitivity	Specificity	Precision	F1-Score
Proposed Hybrid Model	Benign	99.09	96.55	99.48	96.55	96.55
	Malignant	100	100	100	100	100
	Normal	99.09	98.82	99.27	98.82	98.82

Table 5. Comparative analysis of the proposed model with other existing studies.

Study	Approach	Year	Dataset	Accuracy
Narin and Onur	AlexNet and ResNet	2022	IQ-OTH/NCC	98.58%
Uğur Demiroğlu	Hybrid DarkNet-53 and DenseNet-201	2023	Chest CT-Scan from Kaggle	98.86%
Shah et al. [41]	Transfer Learning VGG16	2020	LUNA16	95.00%
Surekha Nigudgi1	Hybrid AlexNet, VGG and GoogleNet	2023	IQ-OTH/NCCD	97.00%
Dodia et al. [40]	SqueezeNet+ ResNe	2022	LUNA16	94.87%

Chen et al. [42]	CNN with embedded NLP	2021	CT scan images	88.00%
Guo et al. [43]	Feature-Based Optimized CNN	2021	Lung CT-Diagnosis	95.96%
Iftikhar Naseer et al.	Modified Alexnet with SVM	2023	LUNA16	97.64%
Proposed hybrid Model	**ResNet101+SqueezeNet**	**2024**	**IQ-OTH/NCCD**	**99.09%**

5. Conclusion and Future work

The study results were cross-referenced with various results reported in relevant literature. Upon comparison, the proposed method demonstrated significant success. These promising outcomes suggest the potential of this method to assist professionals in effectively categorizing lung cancer, potentially reducing their workload in this field. The accuracy achieved by this hybrid method 99.09% of lung cancer cases are classified and detected with mRMR. The exceptional performance demonstrated by deep learning algorithms like ResNet101 and SqueezeNet holds the promise of offering more accurate diagnoses, thereby enhancing patient outcomes. Furthermore, a wider array of datasets should be utilized to test the models across diverse scenarios. This broader exploration would contribute to a comprehensive evaluation of the models' effectiveness, consequently refining their approach.

Data Availability: The IQ-OTH/NCCD dataset used in this study is publicly available for research purposes. The dataset can be accessed through the [IQ-OTH/NCCD official repository] (https://data.mendeley.com/datasets/4drtyfjtfy/1), which hosts comprehensive information and download options. This repository includes detailed documentation on the data collection process, variable descriptions, and guidelines for proper usage. Researchers and practitioners are encouraged to utilize this resource to support reproducibility and facilitate further advancements in the field. If any additional information is required, the authors can be contacted for further assistance.

Conflicts of Interest: We explicitly stated that there are no conflicts of interest in this research. The work reported in this study could not have been influenced in any way by any personal, financial, or other contacts. The authors have approved the final text of the manuscript.

References

1. American Cancer Society, I., 2023. What is lung cancer. Available: http://www.cancer. org/cancer/what-is-cancer.html.
2. U.S.D.o.H.a.H. Services, N.I.o. Health, N.C. Institute, 2023. Cancer treatment. Available: http://www.cancer.gov/about-cancer/treatment.
3. WHO, 2023. Cancer. Available: http://www.who.int/health-topics/cancer#tab=tab_1.
4. S. Luo, J. Zhang, N. Xiao and Y. Qiang, "Das-net: A lung nodule segmentation method based on adaptive
5. dual-branch attention and shadow mapping," Applied Intelligence, vol. 52, no. 4, pp. 1–15, 2022.
6. A. Pezeshk, S. Hamidian, N. Petrick and B. Sahiner, "3-D convolutional neural networks for automatic detection of pulmonary nodules in chest CT," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 5, pp. 2080–2090, 2018.
7. Clark SB, Alsubait S (2020) Non small cell lung cancer. StatPearls [Internet] Kadir, T., Gleeson, F., 2018. Lung cancer prediction using machine learning and advanced imaging techniques. Transl. Lung Cancer Res. 7, 304.
8. Wang, Shudong, Dong, Liyuan, Wang, Xun and Wang, Xingguang. "Classification of pathological types of lung cancer from CT images by deep residual neural networks with transfer learning strategy" Open Medicine, vol. 15, no. 1, 2020, pp. 190-197. https://doi.org/10.1515/med-2020-0028
9. S. Zheng, J. Guo, X. Cui, R. N. J. Veldhuis, M. Oudkerk et al., "Automatic pulmonary nodule detection in CT scans using convolutional neural networks based on maximum intensity projection," IEEE Transaction on Medical Imaging, vol. 39, no. 3, pp. 797–805, 2020.
10. Thakur SK, Singh DP, Choudhary J (2020) Lung cancer identification: a review on detection and classification. Cancer Metastasis Rev 39:989–998
11. Lai Y-H, Chen WN, Hsu TC, Lin C, Tsao Y, Wu S (2020) Overall survival prediction of non-small cell lung cancer by integrating microarray and clinical data with deep learning. Sci Rep 10(1):1–11. https://doi. org/10.1038/s41598-020-61588-w
12. Chen, J., Ma, Q., Wang, W., 2021b. A Lung Cancer Detection System Based on Convolutional Neural Networks and Natural Language Processing. In: 2021 2nd International Seminar on Artificial Intelligence, Networking and Information Technology. AINIT, pp. 354–359
13. H. Jiang, F. Gao, X. Xu, F. Huang and S. Zhu, "Attentive and ensemble 3D dual path networks for pulmonary nodules classification," Neurocomputing, vol. 398, pp. 422–430, 2019.
14. Riquelme D, Akhloufi MA (2020) Deep learning for lung cancer nodules detection and classification in CT scans. Ai 1(1):28–67
15. Tekade, R., Rajeswari, K., 2018. Lung cancer detection and classification using deep learning. In: 2018 fourth international conference on computing communication control and automation. ICCUBEA, pp. 1–5.
16. S. R. Jena and S. T. George, "Morphological feature extraction and KNG-CNN classification of CT images for early lung cancer detection," International Journal of Imaging System and Technology, vol. 30, no. 4, pp. 1324–1337, 2020.
17. Chaunzwa TL, Hosny A, Xu Y, Shafer A, Diao N, Lanuti M, Aerts HJ (2021) Deep learning classification of lung cancer histology using CT images. Sci Rep 11(1):1–12
18. Blanc-Durand P, Campedel L, Mule S, Jegou S, Luciani A, Pigneur F, Itti E (2020) Prognostic value of anthropometric measures extracted from whole-body CT using deep learning in patients with non-small-cell lung cancer. Eur Radiol 30:1–10. https://doi.org/10.1007/s00330-019-06630-w
19. Raza, R., Zulfiqar, F., Tariq, S., Anwar, G.B., Sargano, A.B., Habib, Z., 2022. Melanoma classification from dermoscopy images using ensemble of convolutional neural networks. Mathematics 10, 26
20. S. Dodia, A. Basava and M. A. Padukudru, "A novel receptive field-regularized v-net and nodule classification network for lung nodule detection," International Journal of Imaging System and Technology, vol. 32, no. 1, pp. 88–101, 2022.
21. Sun R, Yuexin P (2022) Efficient lung cancer image classification and segmentation algorithm based on improved Swin transformer. arXiv preprint arXiv:2207.01527
22. Lyu, Lei, 2021. Lung cancer diagnosis based on convolutional neural networks ensemble model. In: 2021 2nd International Seminar on Artificial Intelligence, Networking and Information Technology. AINIT. IEEE.
23. Nanglia P, Kumar S, Mahajan AN, Singh P, Rathee D (2021) A hybrid algorithm for lung cancer classification using SVM and Neural Networks. ICT Express 7(3):335–341
24. Huang W, Wang J, Wang H, Zhang Y, Zhao F, Li K, Su L, Kang F, Cao X (2022) PET/CT based EGFR mutation status classification of NSCLC using deep learning features and radiomics features. Front Pharmacol 13:1504. https://doi.org/10.3389/fphar.2022.89852

25. Almas B, Rajasekaran S (2019) A deep analysis of google net and alex net for lung cancer detection. Int. J. Eng. Adv, Technol
26. Ma S, Ahn G, Hong H (2022) Chest CT image patch-based CNN classification and visualization for predicting recurrence of non-small cell lung Cancer patients. J Korea Comput Graph Soc 28(1):1–9
27. Malik PS, Raina V (2015) Lung cancer: prevalent trends and emerging concepts. Indian J Med Res 141(1):5–7
28. Surekha Nigudgi et al. Lung cancer CT image classification using hybrid-SVM transfer learning approach. https://doi.org/10.1007/s00500-023-08498-x
29. I. M. Baltruschat, H. Nickisch, M. Grass, T. Knopp and A. Saalbach, "A. comparison of deep learning approaches for multi-label chest X-ray classification," Scientific Reports, vol. 9, no. 1, pp. 6381, 2019.
30. J. Kang and J. Gwak, "Ensemble of instance segmentation models for polyp segmentation in colonoscopy images," IEEE Access, vol. 7, pp. 26440–26447, 2019.
31. Tantisatirapong, S., Davies, N.P., Rodriguez, D., Abernethy, L., Auer, D.P., Clark, C.A., et al.: Magnetic resonance texture analysis: optimal feature selection in classifying child brain tumors. In: XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013, pp. 309–312. Springer, Cham (2014)
32. Gao, W., Hu, L., Zhang, P.: Feature redundancy term variation for mutual information-based feature selection. Appl. Intell. 50(4), 1272–1288 (2020)
33. Zhao, J., Meng, Z., Wei, L., Sun, C., Zou, Q., Su, R.: Supervised brain tumor segmentation based on gradient and context-sensitive features. Front. Neurosci. 13, 144 (2019)
34. Huda, S., Yearwood, J., Jelinek, H.F., Hassan, M.M., Fortino, G., Buckland, M.: A hybrid feature selection with ensemble classification for imbalanced healthcare data: a case study for brain tumor diagnosis. IEEE Access 4, 9145–9154 (2016)
35. Khan, A. N., Nazarian, H., Golilarz, N. A., Addeh, A., Li, J. P., Khan, G. A.: Brain tumor classification using efficient deep features of MRI scans and support vector machine. In: 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP) Chengdu, China, pp. 314–318. IEEE (2020)
36. Imyanitov EN, Iyevleva AG, Levchenko EN (2020) Molecular testing and targeted therapy for non-small cell lung cancer: current status and perspectives. Crit Rev Oncol Hematol:103194. https://doi.org/10.1016/j. critrevonc.2020.103194
37. Hintze C, Dinkel J, Biederer J, Heußel CP, Puderbach M (2011) Staging des Lungenkarzinoms. Radiologe 51(2):135–144. https://doi.org/10.1007/s00117-010-2112-8
38. Song MJ et al (2020) Increased number of subclones in lung squamous cell carcinoma elicits overexpression of immune related genes. Transl Lung Cancer Res 9(3):659. https://doi.org/10.21037/tlcr-19-589
39. Narin, D., Onur, T.O., ¨ 2022. The effect of hyper parameters on the classification of lung cancer images using deep learning methods. Erzincan Univ. J. Sci. Technol. 15, 258–268.
40. S. Dodia, A. Basava and M. A. Padukudru, "A novel receptive field-regularized v-net and nodule classification network for lung nodule detection," International Journal of Imaging System and Technology, vol. 32, no. 1, pp. 88–101, 2022.
41. G. Shah, R. Thammasudjarit, A. Thakkinstian and T. Suwatanapongched, "Nodulenet: A lung nodule classification using deep learning," Ramathibodi Medical Journal, vol. 43, no. 4, pp. 11–19, 2020.
42. Chen, J., Ma, Q., Wang, W., 2021b. A Lung Cancer Detection System Based on Convolutional Neural Networks and Natural Language Processing. In: 2021 2nd International Seminar on Artificial Intelligence, Networking and Information Technology. AINIT, pp. 354–359.
43. Z. Guo, L. Xu, Y. Si and N. Razmjooy, "Novel computer-aided lung cancer detection based on convolutional neural network- based and feature-based classifiers using metaheuristics," International Journal of Imaging System and Technology, vol. 31, no. 4, pp. 1954–1969, 2021.
44. Javed, R., & Khan, A. H. (2022). A Systematic Analysis for Cardiovascular Disease Classification Using Deep Learning. Journal of Computing & Biomedical Informatics, 3(02), 32-41.