JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 07 Issue 02
2024
Research Article
https://doi.org/10.56979/702/2024

# An Optimized VGG-19 Architecture Integrated with Support Vector Machines for Lung Cancer Detection and Classification

**Zeeshan Mubeen[1], Nazish Rasheed[2*], Mudassar Rehman[3], Sajid Iqbal[4], and Shehroz Zafar[2]**

[1]Riphah School of Computing and Innovation, Riphah International University, Lahore, 54000, Pakistan.
[2]Faculty of Computer Science and Information Technology, Superior University, Lahore, 54000, Pakistan.
[3]Department of Computer Science, Riphah International University, Sahiwal, 57000, Pakistan.
[4]Department of Computer Science, University of Lahore, Lahore, 54000, Pakistan.
[*]Corresponding Author: Nazish Rasheed Email: nazishrasheed299@gmail.com

**Received:** March 23, 2024 **Accepted:** August 19, 2024 **Published:** September 01, 2024

**Abstract:** Lung cancer is an identical serious and deadly disease, normally signaled by small growths in the lungs called nodules. It usually happens because cells in the lung start increasing uncontrollably. Finding these lung nodules is important for detecting lung cancer, often done using CT scans. Catching the disease early significantly improves the chances of effective treatment and survival. To recuperate lung cancer detection, this study introduces an automated method for finding nodules in CT images, called Enhanced Visual Geometry Group (EVGG-SVM). This method uses an improved version of a well-known neural network model (VGG19) combined with a Support Vector Machine (SVM) to classify nodules as either inoffensive or cancerous. The proposed model was evaluated using the well-known LUNA16 dataset and demonstrated high levels of accuracy, sensitivity, and specificity. In comparison with other current techniques, the EVGG-SVM model demonstrated remarkable performance.

**Keywords:** Lung Cancer Detection and Classification; Modified VGG-19; SVM; LUNA16; Deep Learning; Medical Imaging; computer-aided Diagnosis

## 1. Introduction

Lung cancer is a significant cause of death due to its high fatality rate. Early detection is vital, but it can be difficult because the initial symptoms are often subtle. Early detection of lung nodules greatly increases the chances of survival. Imaging tests like CT scans, MRIs, and X-rays are commonly used to detect these nodules. Out of all the options available, CT scans are widely used because they are cost-effective and produce high-quality images. However, examining these images to identify nodules requires a significant amount of effort, which can place additional strain on radiologists.

Radiologists use CAD to diagnose lung cancer more precisely. These technologies substantially reduce radiologists' effort and increase diagnostic accuracy. Example: CDAM, a decision-making and characteristic-representation system. Research shows that deep learning (DL) CAD systems can identify lung nodules. ML techniques like SVMs are used in picture classification to detect key traits. Testing a recommended CAD system using the LIDC/IDRI dataset and 10-fold cross-validation. Automation, natural language processing, agriculture, and health may benefit from deep learning. Manifold Regularized Classification Deep Neural Network identifies lung nodules using CT data. Deep neural networks like convolutional neural networks (CNNs) can automatically extract essential features for object detection. A CNN constructed on five convolutional layers (CNN-5CL) was proposed by Manickavasagam et al. [11] to improve the ability of the CAD system to classify lung nodules. Zheng et al. [13] investigated CNN application in CT scan images for lung nodule identification using maximum intensity projection imaging. Their sensitivity was 94.2% using the Lung nodule analysis 2016 (LUNA16) dataset. A unique multi-model ensemble architecture built on a 3D convolutional neural network was created by Liu et al. [14] in

order to detect abnormal or potentially cancerous lung nodules. In order to change nodules with low contrast into those with high contrast, they employed image enhancement techniques. For varying nodule sizes, multi-model network architecture is used, which reduces the amount of time the framework needs to train. Additionally, the lung nodule categorization is made more efficient by applying ensemble learning. They also proposed a CNN-based technique for evaluating lung nodule malignancy. They solved the imbalance problem by using a new loss function with cross-entropy and a multiple lightweight network ensemble technique for limited datasets. Al-Shabi et al. [16] classified lung nodules with 92.81% accuracy through using CNN's attention operation on the 3D axial attention approach. In order to diagnose lung cancer, advanced machine learning and deep learning algorithms have been developed for lung nodule identification. Still up for debate, though, are performance-related concerns with lung cancer early detection.

The primary contributions are as follows:

- To cut down on computation time, we preprocessed the LUNA16 dataset, which is publicly accessible.
- To enable effective classification, a modified version of the It is recommended to use the VGG-19 architecture to extract features from a large number of CT scan pictures.
- Efficient model training with three optimizers reduces computation time and maximizes resource consumption.

The document is organized as follows: Section 2 provides a review of recent research. Section 3 details the systematic steps. Section 4 presents the results of the proposed model, while Section 5 summarizes the key conclusions.

## 2. Related Work

This section examines current, pertinent research efforts that mostly concentrate on deep learning methodologies. The review uses Elsevier, Springer, Scopus, and Web of Science resources to diagnose and categorize lung nodules and lung cancer. Twenty relevant 2020–2024 research publications were selected. The review is explained below:

Spatial attention was employed to discover ROIs, to improve characteristics, and an ensemble to improve detection robustness in [17]. Included in the model are three convolutions, three pooling layers, and two fully linked layers. CLAHE was used to segment ROI during preprocessing. After morphological analysis, the KNG-CNN model classified the lung nodule. This model has 87.3% LIDC-IDRI accuracy. Dodia et al. [19] proposed RFR V-Net, a deep-learning network for lung nodule classification. RFR is used to extract features from the deconvolution decoder and convolution encoder blocks. Lung nodules are classified using SqueezeNet and ResNet. Vijh's team. Preprocessing included a Wiener filter to minimize CT image noise and global thresholding to distinguish cancerous and noncancerous pictures.

To distinguish benign from malignant CT scans, Deep Convolutional Neural Networks (DCNN) were suggested. The public LIDC-IDRI dataset was tested. The technique employed CNN-based and feature-based classifiers from Guo et al. [23]. The CNN-based classifier used the Harris Hawks optimizer, Harelick, and local binary patterns to produce results. The CNN classifier was employed first. Hesamian et al. [24] discussed how difficult it is to segment nodules, discern irregular shapes, and find little contrast on lung CT images. Researchers solved this problem via deep learning. A synthetic image was built using slices with varying color patterns. U-Net was enhanced to detect color patterns using DL-based segmentation of synthetic photos. This method classified and segmented lung nodules more accurately and efficiently. Guo et al. developed CNN-based lung cancer detection [25]. Use residual connection and hybrid attention. The end-to-end classification network employs convolutional neural networks. Feature extraction, higher-level categorization, and class integration were automated. Researchers solved this problem via deep learning. A synthetic image was built using slices with varying color patterns. DL segmented synthetic photos to develop a modified U-Net model that recognizes color patterns. This method classified and segmented lung nodules more accurately and efficiently. Guo et al. automated lung cancer detection using CNN [25]. Use residual connection and hybrid attention. The end-to-end classification network employs convolutional neural networks. Feature extraction, higher-level categorization, and class integration were automated.

Through ensemble learning, confidence probability was collected and utilized as data. Lung nodules were classified using Naik et al.'s fractal net-work technique [30]. The Fractalnet model was used to train

and validate the system's p performance using the LUNA16 dataset, yielding an accuracy rate of 94.7%. LUNA16 was used to train and evaluate the linear discriminant analysis (LDA), support vector machine (SVM), and AdaBoostM2 algorithms. The results showed 96.89% accuracy in the categorization of pulmonary nodules. Comparably, a different study [32] created a deep learning-based segmentation technique for the identification of lung cancer. Numerous more cutting-edge previous studies [33–36] covered a range of life domains and other practical issues. This work has been centered toward the identification of lung cancer. Tab. 1 summarizes the most recent research on the same issue.

**Table 1.** Synopsis of the most recent literature review

| Author | Dataset | Year | Method | Accuracy |
|---|---|---|---|---|
| Jena et al. [18] | LIDC-IDRI | 2020 | KNG-CNN | 87.3% |
| Shah et al. [26] | LUNA16 | 2020 | Transfer learning VGG16, VGG19 | 95.0% |
| Bansal et al. [28] | LUNA16 | 2020 | Deep3DSCcan | 92.7% |
| Jiang et al. [17] | LUNA16 | 2019 | 3D dual CNN | 90.24% |
| Naik et al. [30] | LUNA16 | 2021 | FractalNet + CNN | 94.7% |
| Muzammil et al. [31] | LUNA16 | 2021 | Ensemble learning based fusion | 96.89% |
| Iftikhar Naseer[32] | LUNA16 | 2023 | Modified AlexNet | 96.37% |

We have determined from Tab. 1 that the existing CNN-based methods for detecting lung nodules from CT scans are neither accurate enough nor efficient enough in terms of extracting features. Furthermore, the strategies have employed more time and money to attain an accuracy range of 87.3% to 97.3%. The research project included [17] [18] [28] and called for increasing the accuracy of detection.

Our goal in this research is to create a diagnostic system that offers a dependable method for classifying lung cancer. Hence, we propose an accurately and swiftly diagnosis lung cancer nodules while using minimal resources. In Section 3, the suggested EVGG-SVM model is shown.

## 3. Materials and Methods

In Fig. 1, CT scan images from the publicly accessible lung cancer dataset LUNA16 are used. In order to prepare the data for the modeling stage, these photos are supplied as input during the data preparation stage. During the data preparation step, noise is removed, and the input photographs are rearranged. Moreover, there are three sizes for the input patch, including 16 × 16, 32 and 48. Additionally, data is sent to the data splitting phase, with 80% of the data going to the training set and 20% being set for validation.

Crucial data is extracted from the input photos during the feature extraction procedure and sent to the following stage. Three optimizers the techniques of stochastic gradient descent (SGD), root mean square propagation (RMSprop), and adaptive moment estimation (Adam) are used using Hyperparameters like 200 epochs, 50 batch sizes, and 0.0001 learning rates.

In the performance layer, the suggested EVGG-SVM model was assessed for accuracy and miss classification rate. The core component of a CNN is the convolution layer, which is in charge of extracting significant characteristics from the input data. Convolutional layers carry out convolution operations, denoted by ∗, and applying the convolutional operation to an image necessitates the use of a filter. Feature maps or activation maps are the results of the convolutional operator. The convolutional procedure is shown in Eq. (1) [32].

$$Z(x, y) = (U * V)(x, y) = p\ q\ U(x + p, y + q)\ V(p, q) \quad (1)$$

Where U is the input matrix (image), V is the p × q filter size, and Z is the output feature map. After convolving the input U and filter V, the feature map Z is formed. More non-linearity is added to a network using an activation function.

The ReLU activation function is used in this study to calculate activation.

$$f(z) = \max(0, z) \quad (2)$$

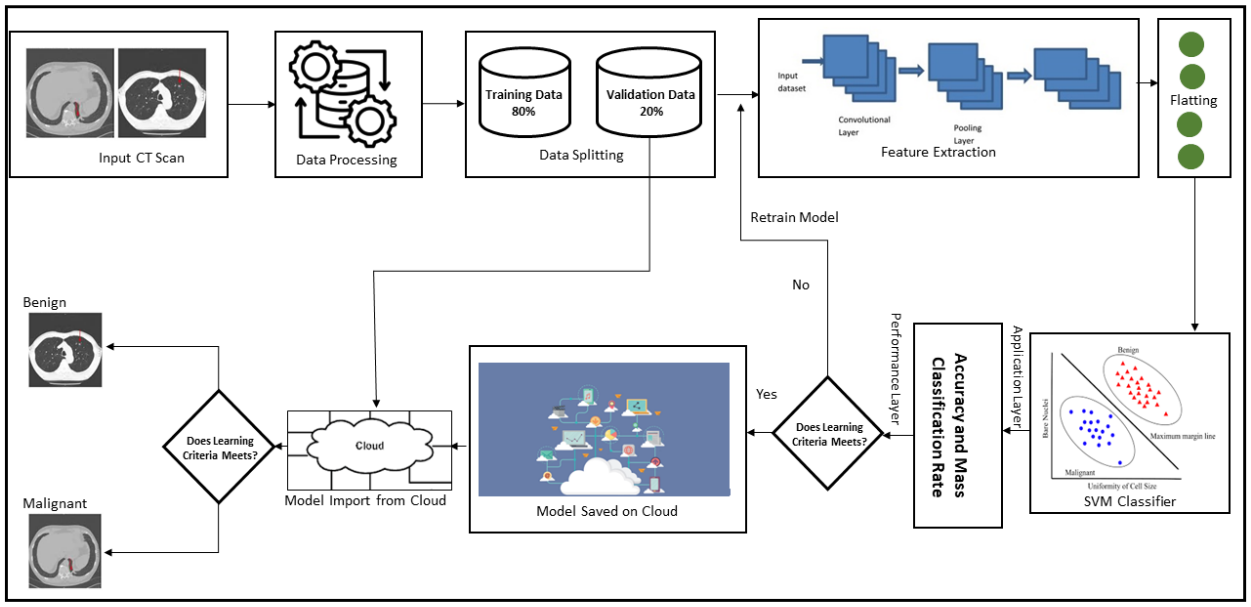


**Figure 1.** The proposed novel model

The Pooling layer, which comes after the Convolution layer, is used to lower the feature map's dimension while preserving its critical information. This technique is also known as down sampling. The pooling layer employs various strategies, including min, max, average, and sum pooling. While the pooling layer reduces the activation map's dimensions, it retains important information. In the convolutional layer, 80% of the preprocessed CT image data is used for convolutional operations. Twenty layers are used in the proposed EVGG-SVM, consisting of six pooling and fourteen convolutional layers. Figure 2 shows the suggested EVGG-SVM CNN network topology for lung nodule identification. Figure 2 demonstrates this with an image size of 32 × 32 × 1. Using the ReLU activation function, the suggested model's nonlinearity is eliminated. Convolutional layers are followed by a max-pooling layer.

The max-pooling layer uses a 2 × 2 filter with a stride of 2, reducing the image size to 16 × 16 × 32. Two subsequent convolutional layers, each with 64 filters, a 3 × 3 kernel, and ReLU activation, maintain the image size with the same padding. The second max-pooling layer, with a 2 × 2 kernel and stride. some additional convo layers, each with 128 filters and a 3 × 3 kernel, followed by another 2 × 2 max-pooling layer, decrease the image size to 4 × 4 × 128. Finally, the multidimensional output is flattened into a 1D vector of size 2048 × 1.

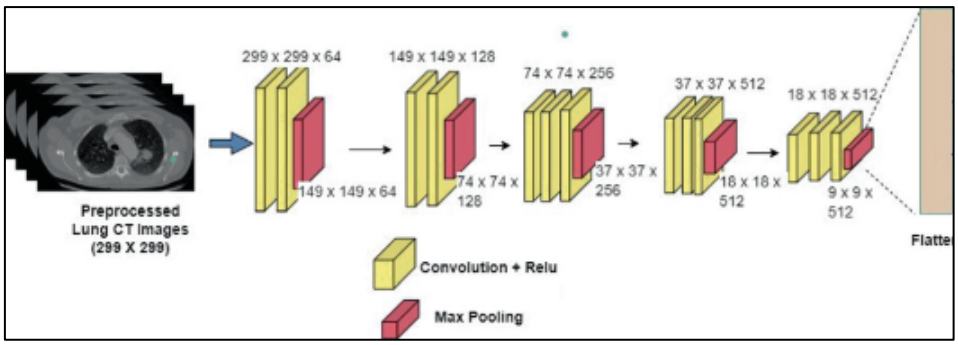


**Figure 2.** The updated VGG-19 architecture shown diagrammatically

Convolutional neural networks classify using a Fully Connected (FC) layer after features extraction [38]. In an FC layer of a traditional neural network, every neuron in the layer above is coupled to every other neuron in the layer below. The fully linked layer's output is sent to the AF for generates class scores. Two methods most frequently used to calculate classification purposes are soft-max and support vector machines. In order to achieve optimal precision in identifying lung cancer as benign or malignant, the model employs the SVM classifier.

Should the learning criteria fail to satisfy the specifications, the EVGG-SVM model that is recommended must be retrained. If the learning criteria does match the requirements, the model and

results are kept on the cloud for further use. The proposed EVGG-SVM training phase ends when it is ready for validation.

During the validation stage, the trained model receives 20% of the validation data, involves importing the trained model from the cloud and evaluating the suggested model. The trained model indicates malignant lung cancer if it finds a cancer nodule; the model suggests benign lung cancer if lung cancer does not reveal any malignant nodules.

**4. Results and Discussion**

This work uses the publicly accessible LUNA16 dataset as a baseline for use in the suggested EVGG-SVM model's training and validation [39]. There are 1018 CT scan images in all in this dataset. The dataset is further split into 20% for SVM model validation and 80% for training. The overall performance of the EVGG-SVM model is validated using metrics from the statistical performance evaluation. The standards are:

$$\text{Accuracy} = (TN + TP) / (TN + FN + FP + TP) \quad (3)$$

$$\text{Miss classification rate} = FN + FP / (TN + FP + FN + TP) \quad (4)$$

$$\text{Specificity} = TN / (TN + FP) \quad (5)$$

$$\text{Sensitivity} = TP / (TP + FN) \quad (6)$$

First, the LUNA16 dataset was used to test the efficacy of the three optimizers Adam, RMSprop, and SGD in the proposed EVGG-SVM model. The results were compared between the optimizers. Second, the effectiveness is assessed in comparison to other cutting-edge techniques for lung cancer detection. The comparative study of the recommended SVM model for training is shown in Tab. 2. Using Adam, RMSprop, and SGD optimizers, the anticipated EVGG-SVM model considers the three different input picture sizes: 16 × 16, 32 × 32, and 48 × 48, in its training phase.

**Table 2.** Evaluation of the suggested EVGG-SVM model in comparison (training)

| Input Image Size | Optimizer | Accuracy | Misclassification Rate | Sensitivity | Specificity |
|---|---|---|---|---|---|
| 16 × 16 | SGD | 97.75% | 2.52% | 97.16% | 96.40% |
| 32 × 32 | Adam | 97.83% | 2.17% | 99.67% | 97.29% |
| **48 × 48** | **Adam** | **98.98%** | **1.02%** | **97.71%** | **99.47%** |

With an input picture size of 16 x 16, the suggested EVGG-SVM model achieves the greatest accuracy of 97.75% during the training phase using an SGD optimizer. Its sensitivity, specificity, and miss classification rate are 2.25%. It is discovered that the suggested EVGG-SVM model, which uses an SGD, attains the greatest correctness of 98.83%. For 48 x 48, the suggested EVGG-SVM model with an SGD obtains the best accuracy of 98.98%.

**Table 3.** Evaluation and comparison of the suggested EVGG-SVM model

| Input Image Size | Optimizer | Accuracy | Misclassification | Sensitivity | Specificity |
|---|---|---|---|---|---|
| 16 × 16 | SGD | 94.64% | 5.36% | 95.16% | 98.73% |
| 32 × 32 | SGD | 98.50% | 1.50% | 98.25% | 98.74% |
| **48 × 48** | **SGD** | **98.98%** | **1.02%** | **98.47%x** | **99.47%** |

The proposed EVGG-SVM model was tested with three different input image sizes (16 × 16, 32 × 32, and 48 × 48) and three optimizers. Table 4 presents the confusion matrix. The model was trained using 1,860 samples, which were categorized into benign and malignant.

**Table 4.** Confusion matrix (training)

| Actual Class | Predicated Clas Benign | Malignant |
|---|---|---|
| Benign | 938 | 07 |
| Malignant | 12 | 903 |

In the benign scenario, the model is trained on a total of 945 samples; 938 samples are accurately predicted by the suggested EVGG-SVM model, whereas 7 samples are incorrectly predicted. The suggested

model, EVGG-SVM, was trained on a total of 915 samples in the malignant case. Of those samples, 903 were properly predicted, while 12 were incorrectly predicted.

The EVGG-SVM model's confusion matrix for the validation phase using the SGD optimizer that produced the best accuracy is shown in Tab. 5. To validate the suggested model, 466 samples in total are collected. A total of 466 samples were separated into malignant and normal categories.

**Table 5.** Confusion matrix (validation)

| **Actual Class** | **Predicated Class** | **Malignant** |
|---|---|---|
| | **Benign** | |
| Benign | 214 | 4 |
| Malignant | 08 | 240 |

Out of the 218 samples the model, the proposed EVGG-SVM model correctly predicted 214 and misclassified 4 samples as benign. For the malignant cases, 248 samples were used for validation, with the model correctly predicting 239 samples and incorrectly predicting 9 samples.

The outcomes of the benign and malignant predictions made using the suggested modified VGG-19 model for the identification of lung cancer are shown in Fig. 3. Since the first three images are actually benign and the suggested EVGG-SVM model accurately predicted them to be benign, they are displayed as true negatives. The next three photos are displayed as false positives; these are benign in reality, but the EVGG-SVM model incorrectly forecasted them to be cancerous. The first three photos are false negatives for malignancy, and the suggested EVGG-SVM model predicts them to be benign when in fact they are cancerous. The last three photos are true positives; as they are cancerous in real life, the suggested EVGG-SVM model predicts them to be malignant.

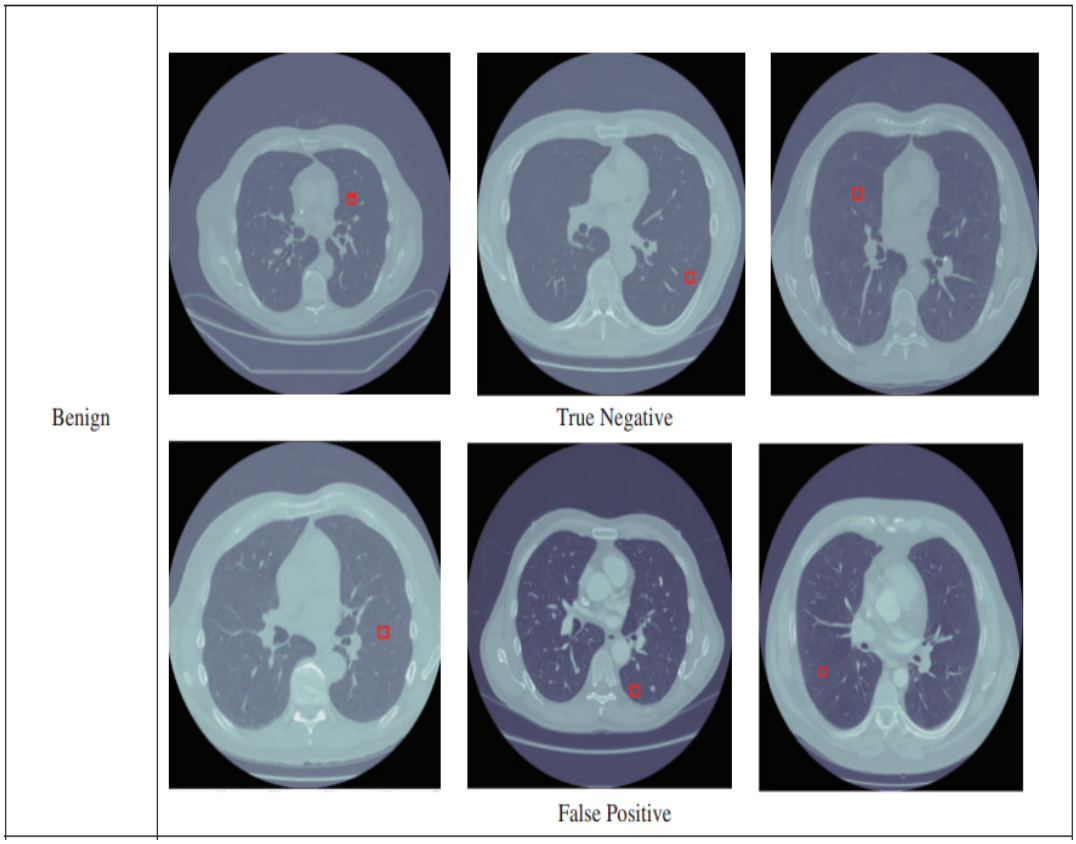


**Figure 3.** The benign detection

The comparison analysis between the suggested EVGG-SVM model and the most advanced techniques currently in use is shown in Tab. 6. The recommended EVGG-SVM model is more accurate and efficient, as shown by the accuracy, sensitivity, and specificity findings.

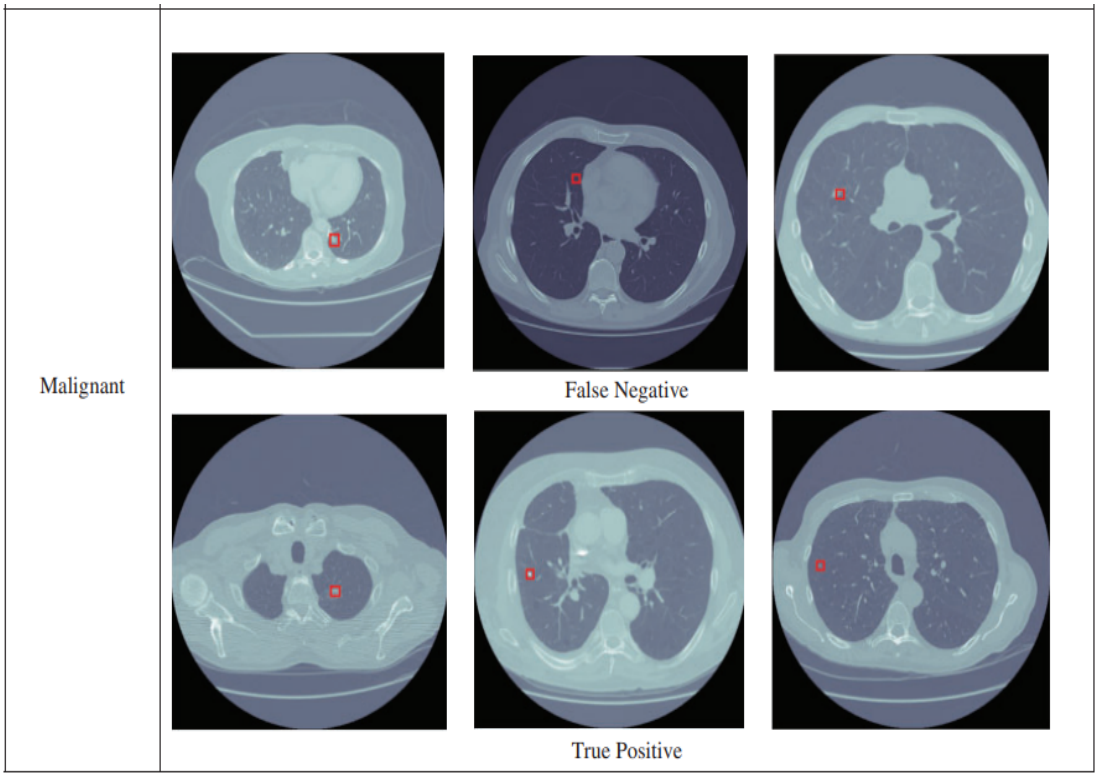


**Figure 4.** The malignant detection

**Table 6.** Comparative evaluation of EVGG-SVM

| Authors | Year | Dataset | Method | Accuracy |
|---|---|---|---|---|
| Jiang et al. [17] | 2019 | LIDC-IDRI | Ensemble 3D Dual Path | 90.24% |
| Rani et al. [21] | 2020 | LIDC-IDRI | KNG-CNN | 87.3% |
| Guo et al. [23] | 2021 | Lung CTDiagnosis | Feature-Based and Optimized CNN | 95.96% |
| Muzammil et al. [31] | 2021 | LUNA16 | Ensemble Learning | 96.89% |
| Bansal et al. [28] | 2020 | LUNA16 | FractalNet + CNN | 94.7% |
| Dodia and associates [19] | 2022 | LUNA16 | SqueezeNet + ResNet | 94.87% |
| Iftikhar Naseer[32] | 2023 | LUNA16 | Modified AlexNet | 96.37% |
| **PROPOSED EVGG-SVM** | **2024** | **LUNA16** | **MODIFID VGG-19** | **99.08%** |

## 5. Conclusions

The EVGG-SVM for the effective segmentation and classification of lung nodules from CT scans is shown in this work. An enhanced VGG-19 architecture is the basis for the suggested EVGG-SVM model. A support vector machine, three fully connected layers, three pooling layers, and fourteen convolutional layers that function as a classifier make up the suggested model. For both training and validation, we

considered three distinct input picture sizes: 16 × 16, 32 × 32, and 48 × 48. The suggested model is optimized for accuracy using three optimizers. The results show that for 48 × 48, the suggested EVGG-SVM with SGD generated the best exactness. The experimental analysis's Python implementation uses the LUNA16 dataset, which is available to the general public. The suggested model has been obtained with a sensitivity of 96.37%, a specificity of 98.62%. Comparing the suggested EVGG-SVM model to other cutting-edge research, it achieves the highest accuracy.

## 6. Limitations and Future Work

The two types of lung cancer identified by the proposed EVGG-SVM are benign and aggressive. The suggested EVGG-SVM, which is based on a modified version of the VGG-19 architecture, is utilized to improve the early detection system's performance for both cancer and non-cancerous conditions. The study demonstrates the limitations of the suggested model, including the fact that it was trained and validated using only one dataset, three different sizes of CT images, and three different optimizers. To improve lung cancer detection performance, the suggested model can be applied to additional publicly accessible datasets using different convolutional neural network topologies and optimizers.

**Data Availability Statement:** The LUNA16 dataset is accessible for non-commercial research, specifically to support developments in automated lung nodule detection. Researchers can obtain the dataset by registering on the official platform and accepting the terms of use. It contains annotated CT scans intended for the development and testing of machine learning models in medical imaging. Users are required to follow guidelines on data usage, including proper citation in research papers. Access to the dataset is granted after completing the registration and agreeing to the usage terms.

**Conflicts of Interest:** The authors confirm that there are no conflicts of interest in this work. The research was carried out independently, with no financial or personal connections that could have impacted the results. All conclusions and findings are based solely on the authors' analysis, without any influence from outside sources or competing interests.

**References**


1. Perez, S.I., Moral-Rubio, S., Criado, R., "A new approach to combine multiplex networks and time series attributes: Building intrusion detection systems (IDS) in cybersecurity", Chaos, Solutions and Fractals, Vol. 150, Pages 1-11, 2021.
2. Kanimozhi, V. and Jacob, T.P, "Artificial Intelligence based Network Intrusion Detection with hyper-parameter optimization tuning on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing", ICT Express, Vol. 5, Issue 3, Pages 211-214, 2019.
3. Saranya, T., Sridevi, S., Deisy, C., Chung, T.D., Ahamed, K.M., "Performance Analysis of Machine Learning Algorithms in Intrusion Detection System: A Review", Third IC on Computing and Network Communications (CoCoNet'19), Trivandrum, 2020.
4. Ghurab, M., Gaphari, G., Alshami, F., Alshamy, R., Othman, S., "A Detailed Analysis of Benchmark Datasets for Network Intrusion Detection System" Asian Journal of Research in Computer Science, Vol. 7, Issue 4, Pages 14-33, 2021.
5. Sharafaldin, I., Lashkari, A., Ghorbani, A., "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization", 4th International Conference on Information Systems Security and Privacy, Portugal, 2018.
6. Özekes, S. and Karakoç, E.N., "Makine Öğrenmesi Yöntemleriyle Anormal Ağ Trafiğinin Tespit Edilmesi", Düzce Üniversitesi Bilim ve Teknoloji Dergisi, Vol. 7, Issue 1, Pages 566-576, 2019.
7. Tama, B.A., Nkenyereye, L., Islam, S.R., Kwak, K.S., "An Enhanced Anomaly Detection in Web Traffic Using a Stack of Classifier Ensemble", IEEE Access, Vol. 8, Pages 24120 – 24134, 2020
8. Abdulrahman, A.A. and Ibrahem, M.K., "Toward Constructing a Balanced Intrusion Detection Dataset Based on CICIDS2017", Samarra Journal of Pure and Applied Science, Vol. 2, Issue 3, Pages 132-142, 2020.
9. Hosseini, S. and Seilani, H., "Anomaly process detection using negative selection algorithm and classification techniques", Evolving Systems, Vol. 12, Pages 769–778, 2021.
10. Hongle, D., Yan, Z., Gang, K., Lin, Z., Chen, Y.C., "Online ensemble learning algorithm for imbalanced data stream", Applied Soft Computing, Vol. 107, Pages 1-12, 2021.
11. Schapire, R.E., "The Boosting Approach to Machine Learning an Overview", In: Denison DD, Hansen MH, Holmes CC et al editors, Nonlinear Estimation and Classification. Lecture Notes in Statistics, Vol. 171, Springer, New York, Pages 1- 23, 2003.
12. Pham, X.T. and Ho, T.H., "Using boosting algorithms to predict bank failure: An untold story", International Review of Economics & Finance, Vol. 76, Pages 40-54, 2021.
13. Shahraki, A., Abbasi, M., Haugen, Q., "Boosting algorithms for network intrusion detection: A comparative evaluation of Real AdaBoost, Gentle AdaBoost and Modest AdaBoost", Engineering Applications of Artificial Intelligence, Vol. 94, Pages 1-14, 2020.
14. Li, Y., Shi, H., Duan, Z., Liu, H., "Smart wind speed forecasting approach using various boosting algorithms, big multi-step forecasting strategy", Renewable Energy, Vol. 135, Pages 540-553, 2019.
15. Ma, B., Meng, F., Yan, G., Yan, H., Chai, B., Song, F., "Diagnostic classification of cancers using extreme gradient boosting algorithm and multiomics data", Computers in Biology and Medicine, Vol. 121, Pages 1-10, 2020.
16. Abro, A.A, Taşcı, E., Uğur, A.A., "Stackingbased Ensemble Learning Method for Outlier Detection", Balkan Journal of Electrical & Computer Engineering, Vol. 8, Issue 2, Pages 191- 185, 2020.
17. Wen, L., Hughes, M., "Coastal Wetland Mapping Using Ensemble Learning Algorithms: A Comparative Study of Bagging, Boosting and Stacking Techniques", Remote Sensing, Vol. 12, Issue 10, Pages 1-18, 2020.
18. Xia, T., Zhuo, P., Xiao, L., Du, S., Wang, D., Lifeng, X. "Multi-stage fault diagnosis framework for rolling bearing based on OHF Elman AdaBoostBagging algorithm", Neurocomputing, Vol. 433, Pages 237-251, 2021.
19. Andiojaya, A. and Demirhan, H., "A bagging algorithm for the imputation of missing values in time series", Expert Systems with Applications, Vol. 129, Pages 10-26, 2019.
20. Yin, S., Liu, H., Duan, Z., "Hourly PM2.5 concentrations multi-step forecasting method based on extreme learning machine, boosting algorithm and error correction model", Digital Signal Processing, Vol. 118, Pages 1-21, 2021.
21. Freund, Y. and Schapire, R.E., "A decisiontheoretic generalization of on- line learning and an application to boosting", Journal of Computer and System Sciences, Vol. 55, Issue 1, Pages 119-139, 1997.
22. Chengsheng, T., Huacheng, L., Xu, B., "AdaBoost typical Algorithm and its application research", MATEC Web of Conferences, Vol. 139, Issue 2, France, 2017.

23. Qi, C., Wang, Y., Tian, W., Wang, Q., "Multiple kernel boosting framework based on information measure for classification", Chaos, Solutions and Fractals, Vol. 89, Pages 175-186, 2016.
24. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A., "CatBoost: unbiased boosting with categorical features", NeurIPS - 32nd Conference on Neural Information Processing Systems, Montreal, 2018.
25. Friedman J.H., "Greedy function approximation: a gradient boosting machine", Annals of statistics, Vol. 29, Issue 5, Page s1189-1232, 2001.
26. Kearns, M. and Valiant, L., "Cryptographic limitations on learning Boolean formulae and finite automata", Journal of the ACM, Vol. 41, Issue 1, Pages 67-95, 1994.
27. Friedman, J.H. "Stochastic gradient boosting", Computational Statistics & Data Analysis, Vol. 38, Issue 4, Page 367-378, 2002.
28. Dahiya, N., Saini, B., Chalak, H.D., "Gradient boosting-based regression modelling for estimating the time period of the irregular precast concrete structural system with cross bracing", Journal of King Saud University - Engineering Sciences, Pages 1-8, 2021.
29. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.Y., "LightGBM: a highly efficient gradient boosting decision tree", NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, Curran Associates Inc. California, 2017.
30. Shehadeh, A., Alshboul, O., Al Mamlook, R.E., Hamedat, O., "Machine learning models for predicting the residual value of heavy construction equipment: An evaluation of modified decision tree, LightGBM, and XGBoost regression", Automation in Construction, Vol. 129, Pages 1-16, 2021.
31. Chen, T. and Guestrin, C., "XGboost: A scalable tree boosting system", 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Pages 785–794, San Francisco, 2016.
32. Khan, M. I., Khan, Z. A., Imran, A., Khan, A. H., & Ahmed, S. (2022). Student performance prediction in secondary school education using machine learning. In 2022 8th International Conference on Information Technology Trends (ITT) (pp. 94-101). IEEE.
33. Ma, J., Zhongqi, Y., Qu, Y., Xu, J., Cao, Y., "Application of the XGBoost Machine Learning Method in PM2.5 Prediction: A Case Study of Shanghai", Aerosol and Air Quality Research, Vol. 20, Issue 1, Pages 128-138, 2019.
34. Sharma, N.V. and Yadav, N.S., "An optimal intrusion detection system using recursive feature elimination and ensemble of classifiers", Microprocessors and Microsystems, Vol. 85, Pages 1-11, 2021.
35. Liu, R., Ali, S., Bilal, S. F., Sakhawat, Z., Imran, A., Almuhaimeed, A., Alzahrani, A., & Sun, G. (2022). An intelligent hybrid scheme for customer churn prediction integrating clustering and classification algorithms. Applied Sciences, 12(18), 9355. MDPI.
36. Aksoy, B., Usta, U., Karadağ, G., Kaya, A.R., Ömür, M., "Classification of Environmental Sounds with Deep Learning", Advances in Artificial Intelligence Research, Vol. 2, Issue 1, Pages 20-28, 2022.
37. Aksoy, B. and Salman, O.K.M., "Detection of COVID-19 Disease in Chest X-Ray Images with capsul networks: application with cloud computing", Journal of Experimental & Theoretical Artificial Intelligence, Vol. 33, Issue 3, Pages 527-541, 202