RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt26117

Intelligent Squirrel Search-Optimized VGG16 with HOG Feature Fusion for High-Accuracy Lung Cancer Classification from DICOM Images

T.S. Sheeja [a, *], Arun Chokkalingam [b]

[a] Department of Biomedical Engineering, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, India

* Corresponding Author Email: sheebio123@outlook.com

DOI: https://doi.org/10.54392/irjmt26117

Received: 07-08-2025; Revised: 19-12-2025; Accepted: 14-01-2026; Published: 30-01-2026

Abstract: Accurate classification and early detection of lung cancer are crucial for effective treatment and improved patient outcomes. Although recent advances in medical imaging have improved diagnostic workflows, many traditional approaches remain limited in their ability to efficiently analyze large volumes of imaging data. Accordingly, this study proposes a deep learning framework for lung cancer classification using Digital Imaging and Communications in Medicine (DICOM) images. The proposed approach integrates an Intelligent Squirrel Search (ISS)–tuned VGG16 network (ISS-VGG16) to improve classification performance. A set of lung cancer DICOM images was obtained from an open-source dataset. Preprocessing steps, including image resizing and contrast enhancement, were applied to standardize the inputs for model training. ISS was used to optimize key hyperparameters of the VGG16 model, thereby improving classification performance. The model was implemented in Python. Experimental results indicate that the proposed ISS-VGG16 approach outperforms baseline methods, achieving a recall of 0.92, precision of 0.96, F1-score of 0.95, and overall accuracy of 0.97 on the evaluated lung cancer image dataset. These results demonstrate reliable classification performance across the defined lung cancer classes. Overall, the proposed framework provides an effective and dependable approach for lung cancer image classification.

Keywords: Lung cancer classification, Deep learning, Intelligent Squirrel Search–tuned VGG16 (ISS-VGG16), DICOM images.

1. Introduction

Globally, most common cause of mortality from the illnesses is lung cancer. Every day, there are more and more people passing away from cancer-related diseases. Diagnoses have to be made immediately. An unprecedented 1.61 million people die from lung cancer annually. Cancer cells develop in several body tissues or organs [1]. Computed tomography (CT) is one medical imaging filtering technique that uses attractive fields to take film images in the scope of medicine; medical image analysis is exceptionally superior, especially when it comes to noninvasive treatment and medical evaluation [2]. Symptoms of lung cancer frequently include weariness, weight loss, and haemorrhaging (coughing up blood). Furthermore, there is a correlation between lung cancer and several risk factors, including food, alcohol, smoking, and air quality. Differences in the shape and structure of the cancer cells tend to separate non-small-cell lung cancer (NSCLC) from small-cell lung cancer (SCLC). NSCLC is thought to account for 85% of occurrences of lung cancer, making it the most prevalent kind as opposed to SCLC, which accounts for 5% of all occurrences [3]. It has a big impact on bioinformatics, especially cancer identification. It is the most deadly disease that has ever struck humankind [4]. As a result, people especially those who drink and smoke can only have a regular check-up every six months. Lung cancer is the second-deadliest cancer globally.

Most of these individual diagnoses are made in stages 3 or 4 of lung cancer since early signs are usually disregarded [5]. When staging lung cancer, computed tomography (CT) imaging is yet one of the gold standards for diagnostics. Nonetheless, radiologists' training and experience have an impact on the standard assessment of radiological images, making it partly subjective and primarily qualitative. While radiological scans can help identify the location and size they are unfit for evaluating the clinical-pathological characteristics of a cancer data that is necessary to make a treatment decision [6]. Patient outcomes for improved early classification and accurate detection of lung cancer can lead to increase detection [7-8]. Cancer is currently a major global health hazard that kills millions

of people without their knowledge. Lung, breast, colorectal, liver, and stomach cancers are the most common types [9]. The unchecked proliferation of aberrant cells from the body's tissues or organs causes cancer. Furthermore, if an early diagnosis is not made, there is a chance that cancer cells will expand to two different organs. It can be argued that a careless diet contributes to the development of colon cancer, although smoking negatively affects the development of lung cancer [10].

Treatment for lung cancer is extremely difficult since the disease is often discovered at an advanced stage. Lung cancer is a major cause of cancer-related mortality, despite the challenge of early detection. Since the signs of cancer emerge in the late stages, when there is no prospect of recovery, it is challenging to prevent. To rectify these problems, this study utilizes an Intelligent Squirrel Search-tuned VGG16 (ISS-VGG16) to enhance lung cancers classification accuracy with DICOM imaging statistics.

2. Related Work

Table 1 summarizes representative prior studies relevant to lung cancer detection and classification, highlighting the objectives, methods, datasets, and key outcomes.

Table 1. Overview of Literature Survey

Author(s)	Year	Objective	Methodology	Dataset	Outcomes
Riquelme & Akhloufi [11]	2020	Improved CAD systems for CT-based lung cancer detection	Deep Learning (DL)-based CAD systems with two categories: false-positive reduction and nodule detection	Public CT lung datasets	Evaluated dataset efficacy; compared methods for better diagnosis support
Wang [12]	2022	Developed DL-based imaging technologies for early lung cancer detection	DL applied to CT and MRI imaging	Medical imaging datasets	Enabled reliable identification and classification of lung nodules across wide patient ranges
Javed et al. [13]	2024	Classified lung cancer using CNN	Convolutional Neural Network (CNN)	CT image datasets	CNN achieved highest accuracy, improving lung cancer categorization and diagnosis
Heuvelmans et al. [14]	2021	Proposed risk-specific management for lung cancer screening	Low-Dose CT (LDCT), LCRAT-CT risk model	18,129 LDCT scans (NLST data)	Personalized risk prediction framework; model may delay diagnosis for 30% of cancers but offers individualized tracking
Abdullah et al. [15]	2021	Compared classifiers for early lung cancer identification	KNN, CNN, and SVM	UCI lung cancer datasets	CNN showed higher accuracy among compared classifiers
Shafi et al. [16]	2022	Proposed DL-assisted SVM model for early pulmonary nodule detection	SVM-based DL model	Lung CT scans	Outperformed existing models; improved early identification support for radiologists

Yang et al. [17]	2021	Classify lung cancer subtypes and mimics from histology	ResNet-50 and EfficientNet-B5, threshold-based six-type classifier	Whole Slide Images (WSI)	AI classifier effectively predicted lung cancer subtypes through WSI analysis
Shin et al. [18]	2020	Examined exosome SERS signals for lung cancer detection	DL on exosome SERS spectral data	SERS data from normal and cancer cells	Achieved AUC of 0.912, showing correlation between molecular similarity and cancer stage
Li et al. [19]	2020	Analyzed DL-based lung cancer segmentation challenge	Single vs. Multi-model DL approaches	Segmentation dataset	Multi-model approach significantly outperformed single-model ones ($p < 0.01$)
Pradhan & Chawla [20]	2020	Reviewed ML algorithms for lung cancer prediction in IoT context	Comparative analysis of ML techniques	Literature-based survey	Identified limitations and suggested IoT-integrated ML for future development
Hatuwal & Thapa [21]	2020	Diagnosed lung cancer subtypes using CNN	CNN model for tissue classification	Image dataset containing cancerous and benign tissues	Achieved 96.11% training and 97.2% validation accuracy in subtype classification
Thakur et al. [22]	2020	Reviewed CAD-based CT methods for lung cancer detection	CAD systems with CT image analysis	Various CT scan datasets	Provided comprehensive review of CAD methods assisting radiologists
Huang et al. [23]	2023	Highlighted AI's role in early lung cancer detection	DL, RL, ML, NLP techniques	Multi-modal AI research	Demonstrated potential for AI-driven screening and personalized treatment
Quasar et al. [24]	2024	Developed ensemble model for lung cancer classification	Ensemble (CNN + DenseNet)	Chest CT-scan dataset	Achieved 98% prediction accuracy; ensemble models outperform single-model approaches
Alshmrani et al. [25]	2023	Created multi-class DL model for lung disease classification	CNN blocks with VGG19 pre-trained model	Medical image datasets (COVID-19, pneumonia, cancer)	Outperformed competing models; effective in lung disease recognition and rapid patient handling

3. Methodology

The methodology for the classification of lung cancer entails preprocessing of DICOM images through resizing and contrast adjustment, accompanied by feature extraction by the use of Histogram of Oriented Gradients (HOG). This study introduces the novel Intelligent Squirrel Search- tuned VGG16 (ISS-VGG16) method, which combines the Intelligent Squirrel Search (ISS) algorithm and optimizes the VGG16 model's hyper parameters. Figure 1 shows the overall flow of the methodology.

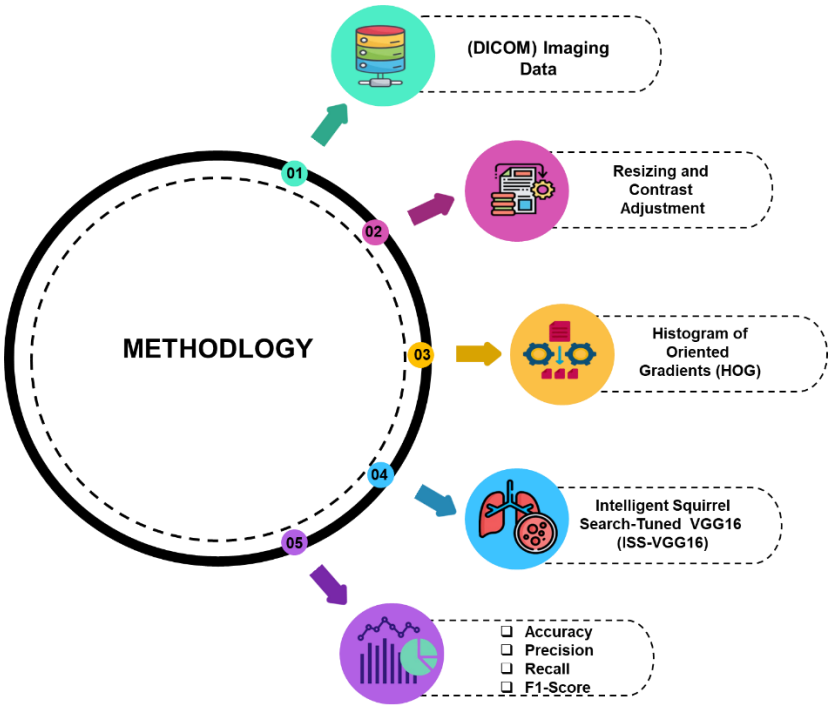

Figure 1. Overflow of methodology

3.1 Dataset

The study compiles lung cancer patients' DICOM images from the Kaggle open-source dataset, which has XML annotation files that use bounding boxes to show the location of the tumour [26]. The images were obtained retroactively from patients who underwent CT and a routine lung biopsy as part of their treatment for suspected lung cancer.

3.2 Pre-Processing of Images

The collected data needs to pre-process, to enhance the classification of the lung cancer images. For pre-processing the images, this study utilizes techniques such as resizing images and contrast adjustment.

3.3 Resizing of Images

The resizing involves zooming in to make the given image larger than the normal size, while the other involves shrinking the image to make the outsized image smaller to meet the desired scale. The pixel relationship method combined with an interpolation scheme that looks at the intensity values of nearby pixels is used to implement these two techniques. They can be enlarged or shrunk to a normal shape while keeping the aspect ratio of the image intact without sacrificing any important details. The image's size can be reduced using the inter-area approach, and its magnitude can be increased using the linear interpolation method.

3.4 Contrast Adjustment

It is a technique for extending the dynamic range of the image's intensity value, hence lengthening the contrast. The minimum and maximum values of the image must be ascertained to stretch it. The image's border will be established by these numbers, which are the minimum and maximum. The lowest limit of this suggested method is an image with an 8-bit gray level, while the higher limit is a range of values 0 to 255. The Equation (1) depicts the contrast adjustment,

$$h(w,z) = \frac{e(w,z)-min}{-max-min} W255 \qquad (1)$$

Where, $h(w,z) = \text{matrix of the resulting image}$ $e(w,z) = \text{original image matrix value}$ where the input is represented by $h(w,z)$ and the output by $e(w,z)$. The values of image intensity range from 0 (the lowest value) to 255 (the greatest value). By calculating the minimum and maximum values using stretching. After subtracting the maximum value from the image value (w,z), the resultant value of $h(w,z)$ as a new image is split by the outcomes of the minimum and maximum reduction. The pixel value will be calculated by multiplying the findings by 255.

3.5 Feature Extraction using Histogram of Oriented Gradients (HOG)

The HOG feature typically improves the classification process's superiority. To get higher accuracy, blocks of the smallest gradients on an image are utilized. The computed gradient indicates that the

persistence of the HOG feature's intensity and direction consistency is a benefit. The area intensities of the block are then standardized by summing the local histogram values with the remaining cells in the selected block.

In lighted and shaded environments, this standardization procedure yields superior outcomes. Equations (2-5) show the gradient $|H|$. Between a certainty kernel factor and an image pixel, the gradients' vertical and horizontal directions are computed analytically.

$$|H| = \sqrt{J_w^2 + J_z^2} \quad (2)$$

$$J_w = J * C_w; C_w = [-1 \quad 0 \quad 1] \quad (3)$$

$$J_z = J * C_z; C_z = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad (4)$$

$$\theta = \arctan\left(\frac{J_z}{J_w}\right) \quad (5)$$

Where, $|H| = gradient\ magnitude$

$J_w = horizontal\ gradient$

J = image pixel intensity

$J_z = vertical\ gradient$

$C_w = horizontal\ kernel\ factor$

$C_z = verticle\ kernel\ factor$

The image direction and magnitude components are subsequently separated into specific blocks to provide a bin containing the directions of the histogram structure. The bin is then arranged to form HOG features.

3.6 Intelligent Squirrel Search- tuned VGG16 (ISS-VGG16)

The integration of ISS with VGG16 reduces over fitting and complements the method generalization by choosing the most excellent weights and parameters that improve model robustness throughout datasets.

3.6.1 VGG16 model

For classifying the lung cancer classification, VGG16 is utilized in this study. The architecture for a deep CNN with 16 layers is called VGG16. A 5-layer convolution plus a 2-layer completely connected make up the 7-layer VGG 16 model's architecture (Figure 2).

Input layers

The raw image is captured in this input layer and sent to the following layer to extract the subsequent feature.

Convolutional Layer

The convolution processing starts with the following input layer. To discover new object features at this layer, a variety of filters will be applied to the objects. These features, or attributes, are then utilized to determine whether each test process is appropriate. The following Equation (6) illustrates the convolutional layer.

$$Output = \frac{m+2o-e}{t} + 1 \quad (6)$$

Where m depicts input height or length; e is kernel feature height or filter length; o is padding and t is strides.

Pooling

The pooling layer receives the feature extraction data. This layer helps shrink an image object's dimensions and properties so that it can hold more significant data. Its role is to keep each layer's maximum value constant. The current layer has no assigned weight. The polling process is split into two parts: global average pooling, which takes an average of all of the convolution results, and max polling, where the work computer takes the maximum value from the convolution results or computations (7).

$$Outputmaxpool = \frac{m-e}{t} + 1 \quad (7)$$

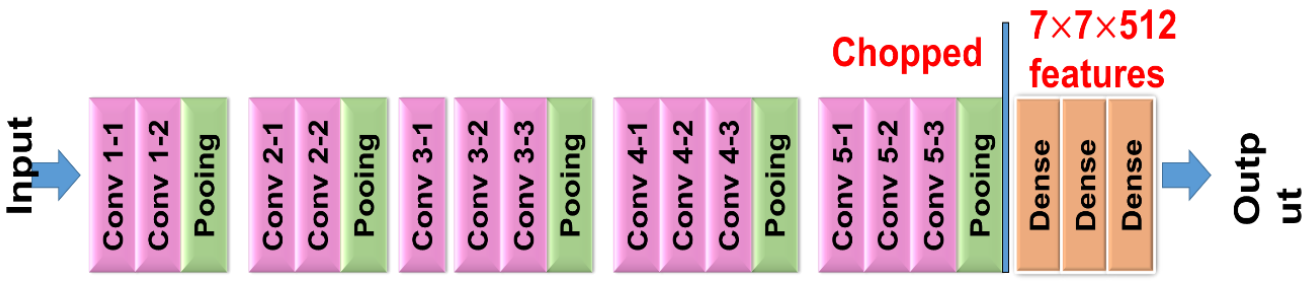

Figure 2. Structure of VGG16 model

Fully Connected Layer

The system operates by taking images that have been filtered at the highest level and translating them into categorical labels. This is the last layer or the completely connected layer. To determine which features are associated with other classes (correlated) in this layer, information will be taken from the outcomes of earlier processing Equation (8). This layer is used to bring all nodes together into a single dimension.

$$Y_i = \sum_{i=1}^{d} x_{j,i}^{S} W_j + a_i \quad (8)$$

Where:

Y_i = The output's value on the network

W_j = Value of the feature extraction input result

$X_{j,i}$ =represents the weight of the network, with a dimension of j x i

j = the number of input features.

i = the number of target classes

a_i = the network's bias

3.6.2 Intelligent Squirrel Search (ISS)

ISS is an optimization algorithm stimulated with the aid of the foraging behavior of squirrels for the classification of lung cancer. It correctly explores and exploits search spaces to locate top-quality answers to various computational problems. Additional movements that are included in the ISS, Movements such as horizontal, vertical, diagonal, and exponential are included in the process of looking for flying squirrels. The ISS technique, which is comparable to the fundamental ISS algorithm, assumes that the squirrels move between three distinct tree species: normal, oak, and hickory. The hickory and oak trees are the only ones that give their nuts sustenance; the other trees do not. To identify nutrients and food resources M_{et} for a specific number of flying squirrels ET, the ISS mathematically assumes that the squirrels are flying in directions to discover the optimal hickory tree as well as the next best possibilities, which are three oak trees. To better balance exploration and exploitation, the agents are split into smaller groups, and within each group, the agents are flexibly altered. 30% of people are in the exploitation group and 70% of people are in the exploration group overall. The individuals in the exploration and exploitation groups are altered to allow for a greater increase in the average fitness of all the individuals, hence raising the fitness level of each group. Mathematically, agents in the exploration group repeatedly look for the greatest option in terms of fitness among the various possibilities in the surrounding area to search for possible locations near their current position in the search space. The flying squirrels' locations and speeds are shown by the following Equations (9) and (10):

$$ET = \begin{bmatrix} ET_{1,1} & ET_{1,2} & ET_{1,3} & \cdots & ET_{1,c} \\ ET_{2,1} & ET_{2,2} & ET_{2,3} & \cdots & ET_{2,c} \\ ET_{3,1} & ET_{3,2} & ET_{3,3} & \cdots & ET_{3,c} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ ET_{m,1} & ET_{m,2} & ET_{m,3} & \cdots & ET_{m,c} \end{bmatrix} \quad (9)$$

$$U = \begin{bmatrix} U_{1,1} & U_{1,2} & U_{1,3} & \cdots & U_{1,c} \\ U_{2,1} & U_{2,2} & U_{2,3} & \cdots & U_{2,c} \\ U_{3,1} & U_{3,2} & U_{3,3} & \cdots & U_{3,c} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ U_{m,1} & U_{m,2} & U_{m,3} & \cdots & U_{m,c} \end{bmatrix} \quad (10)$$

Where j varies from 1 to n and i ranges from 1 to$c$, and the notation $ET_{j,i}$ indicates the position of the jth flying squirrel in the ith dimension. In the ith dimension, the velocity of the jth flying squirrel is represented by$U_{j,i}$. Within specified lower and higher bounds, a uniform distribution determines the beginning locations of$ET_{j,i}$. For every flying squirrel, the physical values $e = e_1, e_2\, e_3, \ldots, e_m$ are computed, with the ideal value denoting a hickory tree. These values are then arranged in ascending order. Regarding the hickory nut tree, ET_{gs} is regarded as the best option. The next three best solutions are then assessed using Equations (11-13). ET_{bs}On trees that bear acorn nuts. It is expected that the remaining solutions on normal trees areET_{ms}. Based on the following scenarios, each flying squirrel's location in the ISS is updated. A randomly generated variable p determines the particular case to be applied. If o equals or exceeds 0.5, the subsequent scenarios will be executed. ET_{bs} Location and relocation to the hickory nut tree:

$$ET_{bs}^{s+1} = \begin{cases} ET_{bs}^{s} + c_h \times H_d\left(ET_{gs}^{s} - ET_{gs}^{s}\right) if\ Q_1 \geq O_{co} \\ \text{Random location , otherwise.} \end{cases} \quad (11)$$

Where ET_{bs} is and how to get to the acorn nut trees:

$$ET_{bs}^{s+1} = \{ET_{bs}^{s} + c_h \times H_d(ET_{bs}^{s} - ET_{ms}^{s})\ if\ Q_2 \geq O_{co} \quad (12)$$

Where ET_{bs} s and how to get to the hickory nut tree:

$$ET_{bs}^{s+1} = \begin{cases} ET_{ms}^{s} + c_h \times H_d\left(ET_{gs}^{s} - ET_{ms}^{s}\right) if\ Q_3 \geq O_{co} \\ \text{Random location, otherwise.} \end{cases} \quad (13)$$

Where $Q_1, Q_2, and\ Q_3$ are random numbers selected during this process from the range of 0 to 1. The current iteration is represented by the variable s, while a random glide distance is represented by the variablec_h. To strike a balance between exploration and exploitation, the constant value of $\boldsymbol{H_d}$ is set to 1.9. Furthermore, in each of the three scenarios, the probability value of $\boldsymbol{O_{co}}$ is set to 0.1.

The following scenarios (14) will be applied by the algorithm if the random value o is less.

If the position of ET_{ms} and diagonal movement:

$$ET_{ms}^{s+1} = \begin{cases} ET_{bs}^{s} + U_{ms}^{s} + \\ d_1 q\left(ET_{gs}^{s} - ET_{ms}^{s}\right) + \\ d_2 q(ET_{bs}^{s} - ET_{ms}^{s}) \quad if\ O_b < b \\ \text{Random } ET_{rand}^{s} \in ET_{ms}^{s} \text{ otherwise,} \end{cases} \tag{14}$$

In these scenarios, random numbers d_1, d_2, q, O_b, and b are selected from the interval [0, 1]. The fitness values for ET^s rand and ET^snt will be computed to ascertain the direction of movement if a random agent, $ET^s \quad rand$, is chosen from the regular agents, $E_m(ETnt^s)$, $E_m\ (ET^srand)$ determines whether the movement is vertical or horizontal. If it is smaller than $E_m(ETnt^s)$, movement will be vertical (15) as follows:

If the position of ET_{ms} and traveling in a vertical or horizontal direction according to the fitness value $E_m(ET_{rand}^{s})$

$$ET_{ms}^{s+1} = \begin{cases} ET_{ms}^{s} + U_{ms}^{s} + \\ d_3 q\ (ET_{rand}^{s} - ET_{ms}^{s}) \\ if\ E_m(ET_{rand}^{s}) < E_m(ET_{ms}^{s}) \\ ET_{ms}^{s} + U_{ms}^{s} + \\ d_1 q\left(ET_{gs}^{s} - ET_{ms}^{s}\right) \text{otherwise} \end{cases} \tag{15}$$

d_3 is a random number that falls between 0 and 1. A second instance is applied in the last scenario in the event that the requirements for both horizontal and vertical movement are not fulfilled. The equation used as follows [16] and [17]. If the movement and location of ET_{ms} will be exponential:

$$ET_{ms}^{s+1} = ET_{ms}^{s} + |ET_{rand}^{s} - ET_{ms}^{s}|exp(as)cos(2\pi s) \tag{16}$$

Here a is a random number ∈ [0,1].

$$ET_{ms}^{new} = ET_{gs}^{s} + 2q((ET_{gs}^{s} - ET_{ms}^{s})\left(1 - \left(\frac{ET_{gs}^{s}+ET_{ms}^{s}}{ET_{ms}^{s}}\right)^2\right)) \tag{17}$$

ISS and VGG16 are integrated to improve the lung cancer classification. ISS-VGG16 develops lung cancer classification by integrating instance-unique scaling with the VGG16 structure, improving feature extraction and model accuracy, leading to more precise and reliable cancer detection. Through the integration of ISS with VGG16, the method reduces dimensionality and concentrates on important aspects of the lung cancer images, improving classification efficiency and accuracy. Pseudocode 1 illustrates the ISS-VGG16.

Pseudocode 1: Intelligent Squirrel Search- tuned VGG16 (ISS-VGG16)

Step 1: Initialize the VGG16 version with pre-skilled weights from ImageNet

Step 2: Remove the very last fully related layers from the VGG16 model.

Step 3: Add new fully connected layers in keeping with the specific task requirement.

Step 4: Define the parameters for Intelligent Squirrel Search (ISS).

Step 5: While (termination situation not met)

a. Evaluate the health of every flying squirrel (agent) in the populace.

b. Sort flying squirrels primarily based on health values:

c. Update the positions of flying squirrels using ISS regulations:

d. Update velocities of squirrels the use of:

Step 6: Check convergence criteria:

Step 7: Select the great-located hyperparameters for the VGG16 model.

Step 8: Fine-track the VGG16 version using decided on hyperparameters: Output the optimized VGG16 model and file the very last accuracy and different metrics.

4. Result

On Windows 11, tasks were completed using Python 3.12.6. There was 32 GB of RAM and an Intel Core i7 12th Gen CPU. A modern laptop setup was used for the test, which allowed for performance evaluations for classification tasks. The objective was to increase the classification accuracy of lung cancer using ISS-VGG16. To estimate the effectiveness of the recommended technique, the following metrics are evaluated: recall, F1 score, precision, and accuracy. For estimating the effectiveness, the proposed ISS-VGG16 is compared with other existing methods such as K-Nearest Neighbours (KNN) [26], gradient boosting [26], and Light Gradient Boosting Machine (LGBM) [26].

4.1 Accuracy

It represents the ratio of every result to precisely anticipated observations. The whole sample size is separated by the sum of the TP and TN values to regulate accuracy. The equation (18) is used to calculate it. The accuracy result of the approach is displayed in Figure 3 and Table 2.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FN+FP}} \tag{18}$$

Where TP =True Positive, TN = True Negative, FP = False Positive and FN= False Negative.

4.2 Precision

It indicates the proportion of properly predicted positive observations produced by the system to the total amount of predictable positive observations. This metric

is computed using Equation 19. Figure 4 and Table 3 show the precision results.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FN}} \quad (19)$$

Table 2. Result of Accuracy

Methods	Accuracy
KNN [26]	0.73
Gradient Boosting [26]	0.71
Light GBM Classifier [26]	0.91
ISS - VGG16 [Proposed]	0.97

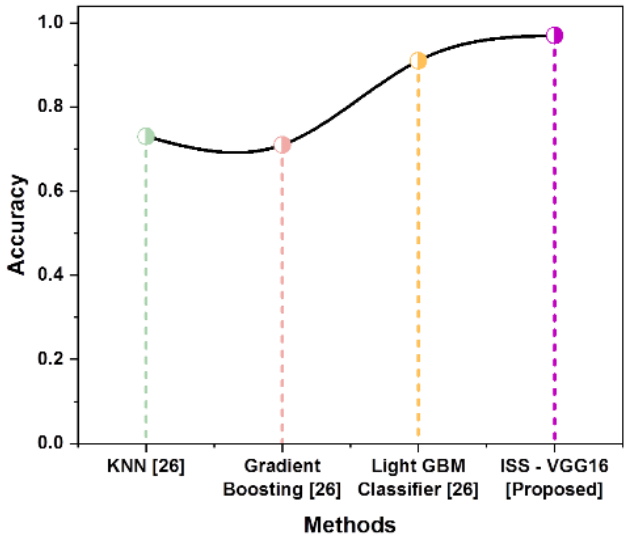

Figure 3. Output of Accuracy

Table 3. Result of precision

Methods	Precision
KNN [26]	0.86
Gradient Boosting [26]	0.90
Light GBM Classifier [26]	0.94
ISS - VGG16 [Proposed]	0.96

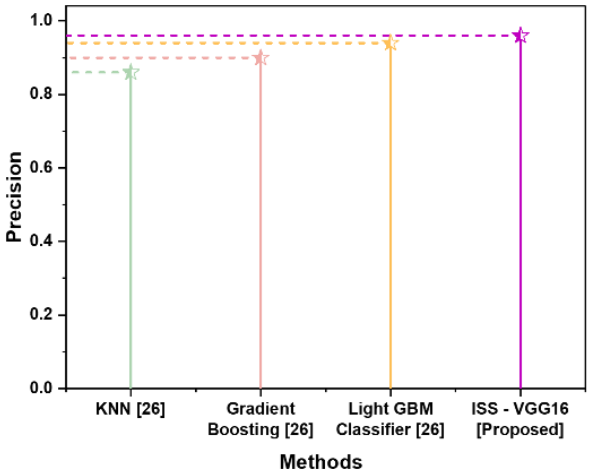

Figure 4. Result of precision

The evaluation of the proposed ISS-VGG16 model demonstrates a high precision rate (0.96), reflecting its capability to accurately distinguish various lung cancer types with minimum false positives. This precision underlines over KNN (0.86), gradient boosting (0.90), and LGBM classifier (0.94) by classification efficacy.

4.3. Recall

It assesses the proportion of all TP to the percentage of system-generated results that accurately anticipated positive observations. Another name for it is sensitivity. It is computed using Equation 20. Figure 5 and Table 4 show the recall results.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+Fp}} \quad (20)$$

Table 4. Result of Recall

Methods	Recall
KNN [26]	0.54
Gradient Boosting [26]	0.47
Light GBM Classifier [26]	0.86
ISS - VGG16 [Proposed]	0.92

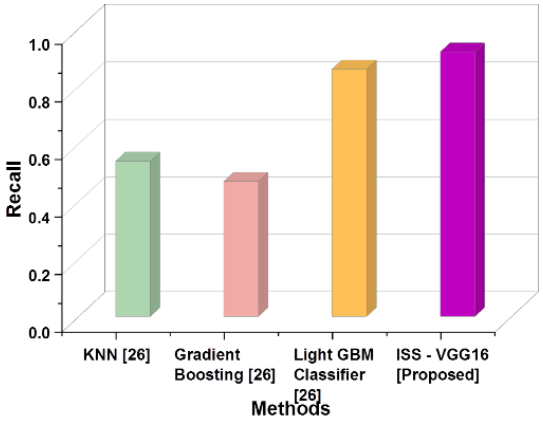

Figure 5. Outcome of recall

This evaluates that the ISS-VGG16 method achieved a greater recall rate of (0.92) for classifying lung cancer. It shows improved recall, demonstrating the model's ability to accurately identify other methods such as KNN (0.54), gradient boosting (0.47), and LGBM classifier (0.86) for lung cancer classification.

4.4 F1-score

The weighted average of the F1 score is determined by recall and precision. It stands for the coordination of precision and recall. F1-score accounts for both false FP and FN. It is computed using Equation 21. Figure 6 and Table 5 show the F1score results.

$$\text{F1score} = \frac{2*\text{Precision}*\text{recall}}{\text{precision}+\text{recall}} \quad [21]$$

Table 5. Result of F1-Score

Methods	F1-Score
KNN [26]	0.67
Gradient Boosting [26]	0.62
Light GBM Classifier [26]	0.90
ISS - VGG16 [Proposed]	0.95

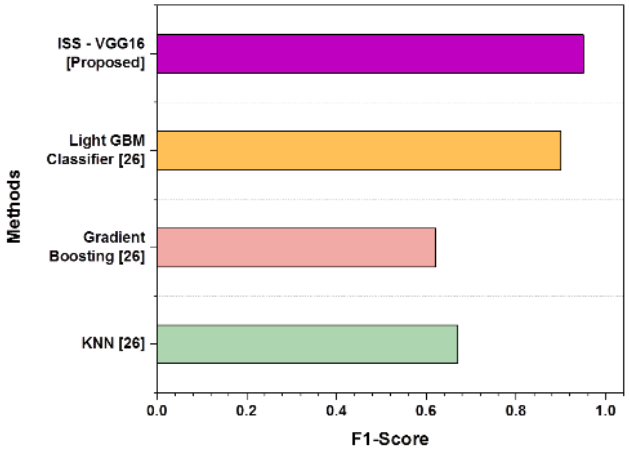

Figure 6. Performance of F1-score

The ISS-VGG16 model classified lung tumors with a high balance between precision and recall, as seen by its F1-score of 0.95. This overall performance demonstrates the proposed method achieved a greater F-1 score than other models such as KNN (0.67), gradient boosting (0.62), and LGBM classifier (0.90).

5. Discussion

In this research, ISS-VGG16 architecture is introduced to improve lung cancer classification accuracy using existing methods such as KNN, Gradient boosting and LGBM. KNN [26] demands for storing the whole dataset, which could cause high memory usage and elevated computational cost, especially with large datasets not unusual in scientific imaging. KNN is rather touchy to noise and outliers, which significantly affect class accuracy. Gradient Boosting [26] models, in particular, combined with huge datasets or complex features from medical imaging, require vast computational assets and time, making them much less green for actual-time or useful resource-restrained environments. The foremost drawback of the LGBM [26] in lung cancer type lies in its sensitivity to hyper parameter settings and characteristic engineering. LGBM [26] can over fit the facts, mainly when the dataset is small or imbalanced, which is regularly the case in clinical programs like lung cancers category. By leveraging VGG16's feature extraction capabilities, optimizing hyper parameters, accelerating convergence, and skilfully managing high-dimensional data, the ISS-VGG16 method improves lung cancer classification and decreases over fitting while increasing the accuracy of lung cancer diagnosis from imaging data.

In this research, a lung cancer classification system using artificial intelligence (AI) was created to develop a diagnostic system by using deep learning and modern feature extraction methods for better diagnosis [27]. The proposed Intelligent Squirrel Search--Tuned VGG16 (ISS--VGG16) model helps to provide significant improvement in cancer diagnosis based on medical images and provides a dependable, highly accurate alternative for the early detection of cancerous materials in the lungs. Automation of deep learning coupled with optimized feature representation makes for this one transformative step in precision oncology [26]. With further clinical validation, the potential for this framework also goes to greater patient survival rates, as well as more efficient organization of healthcare as a whole through earlier and more precise diagnoses. The architecture of ISS -- VGG16 network extends the original VGG16 network architecture with the Histogram of Oriented Gradients (HOG) generation to extract features from the images which enriches the representation of the texture and edge level image on DICOM lung imaging data. This hybridization allows the detection of malignant lung nodules with an extremely high accuracy and generalization capability [28]. Beyond being a way to be more accurate with diagnosis, the methodology holds a potential that includes in the design of personalized treatment strategies, where therapeutic strategies are combined according to the specific characteristics of present and identified tumor. By combining the power of advanced imaging techniques and deep learning algorithms, a potential to boost the survivability of many diseases can be dramatically increased, because the important early adoption of a patient to successful treatment and survival is achieved.

In order to confirm the performance of the Intelligent Squirrel Search (ISS) optimization algorithm, comparative experiments were conducted with four popular hyperparameter tuning methods, i.e., Grid Search (GS), Random Search (RS), Particle Swarm Optimization (PSO), and Genetic Algorithm (GA). The optimization process concentrated on significant model nature parameters that involved learning rate, batch size as well as different dropout rate to the amount of dense layers in the VGG16 model. Model evaluation was done using several metrics like classification accuracy and training efficiency (time) and stability of convergence [29].

The results showed a superior performance of accuracy from the classification with the performance of the ISS which is 1-4% higher than competing optimization algorithms. The reason for this improvement has been explained by the balanced search mechanism of ISS, which effectively combines

global exploration and local exploitation and as such avoids premature convergence [30]. In terms of computational efficiency, it was shown that ISS reached the optimal ones 20-50% faster than Grid Search and Random Search techniques, and 15-30% faster than PSO and GA techniques by an adaptive foraging approach, which reduces unnecessary evaluations. Furthermore, the highest stability, which is characterized by the lowest performance variance with accuracy determination of 0.5% in ten runs, was observed for ISS compared with PSO (1.2%) and GA (1.0%), which confirmed its robustness and reliability. Overall, the performance of the ISS-VGG16 model is captivating because it maintains a very good balance between accuracy, computational efficiency and stability by outperforming existing models and optimization strategies.

6. Conclusion

The proposed ISS-VGG16 version demonstrates improved type accuracy for lung cancer via efficiently utilizing DICOM imaging information and optimized hyper parameters. It outperformed conventional strategies in precision and robustness, ensuring extra dependable diagnoses. This advancement holds the ability to notably improve early detection and affected person effects. In lung cancer classification, our proposed (ISS-VGG16) method achieved precision (0.96), accuracy (0.97), F1-Score (0.95) and recall (0.92). The ISS-VGG16 model, at the same time as enhancing class accuracy, might also face challenges in computational complexity and generalization throughout various imaging datasets. Future work could focus on optimizing the model for quicker processing and extending it to contain multi-modal records for complete diagnostic insights. Additionally, refining the model to address information imbalances and uncommon lung cancer types is important for broader applicability.

References

[1] B.R. Pandit, A. Alsadoon, P.W.C. Prasad, S. Al Aloussi, T.A. Rashid, O.H. Alsadoon, O.D. Jerew, Deep learning neural network for lung cancer classification: enhanced optimization function, Multimedia Tools and Applications82(5) (2023) 6605–6624. https://doi.org/10.1007/s11042-022-13566-9.

[2] S.K. Lakshmanaprabu, S.N. Mohanty, K. Shankar, N. Arunkumar, G. Ramirez, Optimal deep learning model for classification of lung cancer on CT images, Future Generation Computer Systems, 92 (2019) 374–382. https://doi.org/10.1016/j.future.2018.10.009

[3] S.K.B. Sangeetha, S.K. Mathivanan, P. Karthikeyan, H. Rajadurai, B.D. Shivahare, S. Mallik, H. Qin, An enhanced multimodal fusion deep learning neural network for lung cancer classification, Systems and Soft Computing 6 (2024) 200068. https://doi.org/10.1016/j.sasc.2023.200068

[4] A.A. Shah, H.A.M. Malik, A. Muhammad, A. Alourani, Z.A. Butt, Deep learning ensemble 2D CNN approach towards the detection of lung cancer, Scientific Reports, 13 (2023) 2987. https://doi.org/10.1038/s41598-023-29656-z

[5] R. Mothkur, B.N. Veerappa, Classification of lung cancer using lightweight deep neural networks, Procedia Computer Science, 218 (2023) 1869–1877. https://doi.org/10.1016/j.procs.2023.01.164

[6] B. Dunn, M. Pierobon, Q. Wei, Automated classification of lung cancer subtypes using deep learning and CT-scan-based radiomic analysis, Bioengineering 10(6) (2023) 690. https://doi.org/10.3390/bioengineering10060690

[7] M.A. Thanoon, M.A. Zulkifley, M.A.A. Mohd Zainuri, S.R. Abdani, A review of deep learning techniques for lung cancer screening and diagnosis based on CT images, Diagnostics, 13(16) (2023) 2617. https://doi.org/10.3390/diagnostics13162617

[8] S. Huang, I. Arpaci, M. Al-Emran, S. Kılıçarslan, M.A. Al-Sharafi, A comparative analysis of classical machine learning and deep learning techniques for predicting lung cancer survivability, Multimedia Tools and Applications. 82(22) (2023) 34183–34198. https://doi.org/10.1007/s11042-023-16349-y

[9] R. Raza, F. Zulfiqar, M.O. Khan, M. Arif, A. Alvi, M.A. Iftikhar, T. Alam, Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images, Engineering Applications of Artificial Intelligence. 126 (2023) 106902. https://doi.org/10.1016/j.engappai.2023.106902

[10] M. Obayya, M.A. Arasi, N. Alruwais, R. Alsini, A. Mohamed, I. Yaseen, Biomedical image analysis for colon and lung cancer detection using tuna swarm algorithm with deep learning model, IEEE Access, 11 (2023) 94705–94712. https://doi.org/10.1109/ACCESS.2023.3309711

[11] D. Riquelme, M.A. Akhloufi, Deep learning for lung cancer nodules detection and classification in CT scans, AI 1(1) (2020) 28–67. https://doi.org/10.3390/ai1010003

[12] L. Wang, Deep learning techniques to diagnose lung cancer, Cancers, 14(22) (2022) 5569. https://doi.org/10.3390/cancers14225569

[13] R. Javed, T. Abbas, A.H. Khan, A. Daud, A. Bukhari, R. Alharbey, Deep learning for lung cancer detection: a review, Artificial Intelligence Review. 57(8) (2024) 197. https://doi.org/10.1007/s10462-024-10807-1

[14] M.A. Heuvelmans, P.M. van Ooijen, S. Ather, C.F. Silva, D. Han, C.P. Heussel, W. Hickes, H.U. Kauczor, P. Novotny, H. Peschl, M. Rook, Lung cancer prediction by deep learning to identify benign lung nodules, Lung Cancer, 154 (2021) 1–4. https://doi.org/10.1016/j.lungcan.2021.01.027

[15] D.M. Abdullah, A.M. Abdulazeez, A.B. Sallow, Lung cancer prediction and classification based on correlation selection method using machine learning techniques, Qubahan Academic Journal, 1(2) (2021) 141–149. https://doi.org/10.48161/qaj.v1n2a58

[16] I. Shafi, S. Din, A. Khan, I.D.L.T. Díez, R.D.J.P. Casanova, K.T. Pifarre, I. Ashraf, An effective method for lung cancer diagnosis from CT scan using a deep learning-based support vector network, Cancers, 14(21) (2022) 5457. https://doi.org/10.3390/cancers14215457

[17] H. Yang, L. Chen, Z. Cheng, M. Yang, J. Wang, C. Lin, Y. Wang, L. Huang, Y. Chen, S. Peng, K. Ke, Deep learning-based six-type classifier for lung cancer and mimics from histopathological whole slide images: a retrospective study, BMC Medicine, 19 (2021) 1–14. https://doi.org/10.1186/s12916-021-01953-2

[18] H. Shin, S. Oh, S. Hong, M. Kang, D. Kang, Y.G. Ji, B.H. Choi, K.W. Kang, H. Jeong, Y. Park, S. Hong, Early-stage lung cancer diagnosis by deep learning-based spectroscopic analysis of circulating exosomes, ACS Nano 14(5) (2020) 5435–5444. https://doi.org/10.1021/acsnano.9b09119

[19] Z. Li, J. Zhang, T. Tan, X. Teng, X. Sun, H. Zhao, L. Liu, Y. Xiao, B. Lee, Y. Li, Q. Zhang, Deep learning methods for lung cancer segmentation in whole-slide histopathology images—the ACDC@Lung challenge 2019, IEEE Journal of Biomedical and Health Informatics, 25(2) (2021) 429–440. https://doi.org/10.1109/JBHI.2020.3039741

[20] K. Pradhan, P. Chawla, Medical Internet of Things using machine learning algorithms for lung cancer detection, Journal of Management Analytics. 7(4) (2020) 591–623. https://doi.org/10.1080/23270012.2020.1811789

[21] B.K. Hatuwal, H.C. Thapa, Lung cancer detection using convolutional neural network on histopathological images, International Journal of Computer Trends and Technology. 68(10) (2020) 21–24. https://doi.org/10.14445/22312803/IJCTT-V68I10P104

[22] S.K. Thakur, D.P. Singh, J. Choudhary, Lung cancer identification: a review on detection and classification, Cancer and Metastasis Reviews. 39(3) (2020) 989–998. https://doi.org/10.1007/s10555-020-09901-x

[23] S. Huang, J. Yang, N. Shen, Q. Xu, Q. Zhao, Artificial intelligence in lung cancer diagnosis and prognosis: current application and future perspective, Seminars in Cancer Biology. 89 (2023) 30–37. https://doi.org/10.1016/j.semcancer.2023.01.006

[24] S.R. Quasar, R. Sharma, A. Mittal, M. Sharma, D. Agarwal, I. de La Torre Díez, Ensemble methods for computed tomography scan images to improve lung cancer detection and classification, Multimedia Tools and Applications. 83(17) (2024) 52867–52897. https://doi.org/10.1007/s11042-023-17616-8

[25] G.M.M. Alshmrani, Q. Ni, R. Jiang, H. Pervaiz, N.M. Elshennawy, A deep learning architecture for multi-class lung disease classification using chest X-ray (CXR) images, Alexandria Engineering Journal. 64 (2023) 923–935. https://doi.org/10.1016/j.aej.2022.10.053

[26] N. Devihosur, R.K. MG, Enhancing precision in lung cancer diagnosis through machine learning algorithms, International Journal of Advanced Computer Science and Applications. 14(8) (2023) 1069-1077.

[27] N. Ma, X. Fang, Y. Zhang, B. Xing, L. Duan, J. Lu, D. Ma, Enhancing the sensitivity of spin-exchange relaxation-free magnetometers using phase-modulated pump light with external Gaussian noise, Optics Express, 32(19) (2024) 33378–33390. https://doi.org/10.1364/OE.530764

[28] P. Kumari, L. Goel, Emerging computational intelligence based techniques for lung cancer diagnosis and classification on chest CT scan images: a comprehensive survey, Artificial Intelligence Review, 59 (2026) 44. https://doi.org/10.1007/s10462-025-11374-9

[29] A. Kumar, N. Mahendran, U-VQVAE-CTLesionNet: A generalized deep learning framework for multi-organ lesion detection and segmentation in medical imaging,

Computational Intelligence. 42(1) (2026) e70168. https://doi.org/10.1111/coin.70168

[30] S. Krishnamoorthy, S. Paulraj, S. Sriram, B. Periyasamy, S. Eswaran, R.A. Kumar, Hybrid GAN model with LSTM-combined ResNet discriminator for COVID-19 classification in CT images, in: Decision Sciences in Bioinformatics: Theory and Practice, CRC Press (2026) 58–78. https://doi.org/10.1201/9781003449058-4

Authors Contribution Statement

T.S. Sheeja: Conceptualization, Methodology, Performed experiments, Software, Visualization, Writing- Original draft preparation, Validation. Arun Chokkalingam: Supervision, Project Management, Validation. Both the authors read and approved the final version of the manuscript.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The dataset analyzed in this study is publicly available from the open-source repository cited in the Dataset section. Any derived data supporting the findings are available from the corresponding author upon reasonable request. The implementation code and model configuration used in this study are available from the corresponding author upon reasonable request.

Has this article screened for similarity?

Yes