Automatic lung cancer detection using hybrid particle snake swarm optimization with optimized mask RCNN

R. Sudha[1] · K. M. Uma Maheswari[2]

Abstract

As a result of its aggressive nature and late identification at advanced stages, lung cancer is one of the leading causes of cancer-related deaths. Lung cancer early diagnosis is a serious and difficult challenge that is crucial to a person's survival. The first diagnosis of the malignant nodules is typically made using chest radiography (X-rays) and computed tomography (CT) scans; however, the potential presence of benign nodules results in incorrect conclusions. The early phases of both benign and malignant nodules exhibit striking similarities. In this paper, a novel deep learning-based model is proposed for the precise diagnosis of malignant nodules. The proposed approach consists of two stages namely, pre-processing and lung nodule detection. Initially, the Lung CT scan images are collected from the dataset. Then, to remove the noise present in the input image, we apply an adaptive median filter. Then, to enhance the image, Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied. After pre-processing, the image is given to the optimized mask RCNN classifier to detect the malignant and benign nodules. To enhance the performance of the Mask RCNN classifier, the hyper-parameters are optimally selected using hybrid particle snake swarm optimization (PS^2OA). The proposed PS^2OA is a hybridization of particle swarm optimization (PSO) and snake swarm optimization (SSO). The performance of the proposed approach is analyzed based on different metrics and effectiveness compared with state-of-the-art works. The proposed approach attained the maximum accuracy of 97.67%. This work aimed at assisting radiologists to detect and diagnose small-size pulmonary nodules more accurately.

Keywords Lung cancer · Benign nodules · Malignant nodules · Optimized mask RCNN · Snake swarm optimization · Particle swarm optimization

✉ K. M. Uma Maheswari
umamahek@srmist.edu.in

R. Sudha
sr8947@srmist.edu.in

[1] Department of Computer Science and Engineering, School of computing, College of engineering and technology, SRM Institute of Science and Technology, Kattankulathur, India

[2] Department of Computing Technologies, School of Computing, College of engineering and technology, SRM Institute of Science and Technology, Kattankulathur, India

1 Introduction

According to global cancer statistics for 2020, Lung cancer has been the most common and lethal oncological illness in the world for many years. According to clinical research, if late-stage patients had been diagnosed and treated earlier, their survival rate within five years would climb to 52% from the current 10% to 16% range. Lung nodules are a key indicator of early-stage lung cancer, which can be evident on CT scans as localized, spherical shaped like lung shadows, and the size is no larger than 3 cm wide [1]. Notwithstanding, the nano size of lung nodules, their morphology, brightness, and other features are close to those of the vascular system and other tissues in the pulmonary parenchyma; therefore, physicians must carefully examine and screen each nodule individually; this procedure is cumbersome and easily leads to exhaustion, thereby increasing the chances of diagnostic errors. Therefore, it is vital to build an automatic detection method to assist physicians in improving the performance and accuracy of lung nodule diagnosis [2]. On grounds of these details, the movement in medicine toward Computer-Aided Diagnosis (CAD) systems, which are subject to quantitative analysis of CT lung images, can improve Lung CT image understanding, disease diagnosis, and detection of small malignant nodules (which are difficult for a clinician to notice), and diagnostic time [3]. The latest generation of CAD systems also helps in the screening process to detect Lung Nodules differentiating between benign and malignant nodules. CAD uses the Artificial Intelligence (AI) algorithm of Deep Learning methods that efficiently leverage object detection. Deep learning is a robust technique in machine learning in which the object detector automatically learns the image characteristics required for computer vision tasks. There are several available algorithms for object detection using deep learning, including Faster R-CNN, you only look once (YOLO), and single shot detection (SSD). Mask R-CNN is an improved version of Faster R-CNN [4] and it simultaneously generates a high-quality segmentation mask for each detected object in an image. It incorporates the two-stage object detection techniques, in the first stage, RoI is predicted using the Region Proposal Network (RPN), and in the second stage, the class and box offset values are predicted in parallel, also producing a segmentation mask for each RoI.

Any CNN model's performance is impacted by various factors, including the size of the dataset, the number of classes, the model's weights, hypermeters, the optimizer, and many others. Optimizing hyper-parameter plays a vital role during the training of Convolutional neural networks [5]. To what extent a convolutional neural network performs well relies on its architecture and the values of its hyper-parameters. CNN includes many hyper-parameters, based on the structure and training such as the number of convolution layers, the number of filters, the size of each filter, Batch size, Learning Rate, momentum, etc. [6, 7]. As not in the model parameter, hyper-parameter tuning can be done manually but this is a tedious and time-consuming process. We can use the automated tuning method to optimize the hyper-parameter [8] to overcome this. Nowadays Modern optimization techniques, namely the heuristic and metaheuristic algorithms are applied for optimizing objective functions. However, the heuristic methods have numerical inefficiency in the search process, like high dimension problems, which leads the complicity in the model [9]. To address this, meta-heuristics and Swarm Intelligence (SI) methodologies and variants were proposed to handle a variety of flexible real-world optimization tasks and address complex/large-scale optimization issues [10].

The key benefit of the SI optimization algorithms over the deterministic approach is the randomization introduced throughout the search phase can get stuck in circumstances with no global ideal solution. Therefore, obtaining the global best solution is practically significant in SI [11]. The Particle Swarm Optimization (PSO) algorithm is one of the variants of

SI proposed by James Kennedy, It is used in many real-world applications like integrating optimization of nonlinear functions and training of neural networks [12] The idea of flying potential solutions across hyperspace while speeding toward "better" solutions is exclusive to the concept of particle swarm optimization. In PSO, to achieve the optimal solution, it uses the global best it's possible to become stuck in the local optima [13], which has the problem of premature convergence and doesn't provide diversity, to avoid this we can use a trajectory-based technique called snake swarm optimization to avoid in stuck with the local optima. Snake swarm optimization works based on the behavior of snakes, which can be incorporated with MRCNN to attain a certain level of accuracy with a significant speedup in computing [14, 15]. So hybrid PS^2OA optimization technique can better seek the global optimal solution and successfully prevent particles from remaining at the local optimum [16]. The main contribution of the proposed approach is listed below;

- Initially, the Lung CT scan images are collected from the dataset and to remove the noise present in the input image, we apply an adaptive median filter. To enhance the image quality, CLAHE is applied.
- After pre-processing, the image is given to the optimized mask RCNN classifier to detect the malignant and benign nodules.
- To enhance the performance of the Mask RCNN classifier, the hyper-parameters are optimally selected using hybrid particle snake swarm optimization algorithm (PS^2OA). The proposed PS^2OA is a hybridization of particle swarm optimization (PSO) and snake swarm optimization (SSO).
- The performance of the proposed approach is analyzed based on different metrics and effectiveness compared with state-of-the-art works.

The rest of the paper is structured as follows: Sect. 2 discusses related work; Sect. 3 discusses the proposed model in detail; Sect. 4 discusses the experimental results and computational performance metrics, and Sect. 5 discusses the proposed work's conclusion and future directions.

2 Related work

In recent years, algorithms that use Deep Learning techniques to detect lung nodules have been used a lot in medical research. Sunyi Zheng et al.[17] developed a deep learning model to locate lung nodules by taking into account the sagittal, coronal, and axial slices of the CT images for the lung region is been evaluated. Here, the system is made up of two parts. First, a supervised encoder-decoder is trained to find the nodule by combining these slices. Multi-scale contextual information is extracted using 3D Dense CNN to get rid of the nodules. Reza Majidpourkhoei et.al. [18], proposed A novel deep-learning framework based on CNN to detect lung Nodules. This framework is designed using light-footed CNN based on the LeNet-5 model. The lung nodule images are processed and drawn on a patch basis. This model takes six hours to train, which is a time-consuming process. Ying Su et.al., designed the framework for detection of Lung Nodule detection using Faster R-CNN [19] and stated that optimizing the training parameters like learning rate and batch size improves the accuracy of the detection and also the dataset size is enhanced by including the medium lung nodule and by taking

into account the upper and lower nodules of the larger nodule that was identified on the CT slice. In this design, the parameters have to be tuned manually.

Menglu Liu et.al. [20], proposed segmentation of the lung nodule using Mask-R.CNN employs instance segmentation, this model is compared with the U-Net, and Mask-R-CNN outperforms in segmentation. Linqin Cai et al., [21] demonstrated pulmonary nodule detection based on the Mask-R-CNN and with a ray-casting volume rendering algorithm. where Mask-R-CNN helps to detect the pulmonary nodule by multiplying the mask matrices and sequences of raw medical images, and the ray-casting technique aids in visualizing the nodule in a 3D model. Here detecting the small nodule's accuracy has to be improved. These models are evaluated using LIDC-IDRI data sets, which is an open-source dataset of Lung Images, and widely for research purposes.

The Deep Learning model uses the CNN architecture for automatic feature extraction and diagnosis of Lung Images. CNN architecture consists of different hyper-parameters that must be optimized to improve performance. In recent days different metaheuristic optimization techniques have been proposed. Wei-Chang Yeh et.al. worked on Simplified Swarm Optimization (SSO) [22] combined with LeNet-5, where the author proposed the sequential Dynamic Variable Range (SDVR), in contrast to typical SSO the feasible range of the next variable, which is determined by the present variable's value. LeNet-SSO architecture has improved the quality of the solution by tuning parameters and items. This system is evaluated using MNIST, Fashion MNIST, and Cifar10 datasets. It outperforms when compared with the other metaheuristic algorithms with LeNet. Singh et.al. proposed a Hybrid MPSO algorithm, which uses multiple swarms in two levels to give a better solution for the objective function [23]. the architecture of the CNN and hyper-parameters are optimized at level 1 and level 2 respectively. To modify the exploration and exploitation characteristics of particles and prevent the PSO algorithm from prematurely converging into a local optimum solution, this technique employs sigmoid-like inertia weight. This system is evaluated using different benchmark datasets like Cifar10, Cifar-100, MNIST, Covexset, and MDRBI and it is outperformed when compared to randomly generated CNN.

Vijh et al., designed a hybrid bio-inspired algorithm [24] for automatic lung nodule detection. Here a novel variant of whale optimization and adaptive PSO (WOA-APSO) is used to optimize the feature selection and the selected features are grouped by employing the linear discriminant analysis, which aids the reduction of dimensions spaces. the lung images are enhanced using a wiener filter and by employing the different segmentation techniques RoI of the Lung region is obtained, this system is evaluated using LIDC datasets, the accuracy, sensitivity, and specificity are 97.18, 97, and 98.66 respectively. The PSO has the downside of easily falling into local optima in high-dimensional space and having a slow convergence rate in iterative processes, despite being well suited for non-linear complex problems. to avoid the pitfalls of premature optimization and to avoid being stuck in neighborhood searches. we adopt another meta-heuristic approach called snake swarm optimization for local search. Gunjan et.al., proposed work on analyzing different metaheuristic algorithms namely Simulated annealing [25], Tree-of-Parzen estimator, and Random search for optimizing the CNN structure hyper-parameter to classify the small pulmonary lung nodules. Here the system uses the LIDC datasets, and the results show that the SA performs well when compared to other metaheuristic algorithms. Sollini et al. [36] explained the lung lesions classification using a deep learning algorithm. The two main modules are the detection of lung nodules on CT scans and the classification of each nodule into benign and malignant types. Computer Aided Diagnostics (CADe) and Computer Aided Diagnostics (CADx) modules rely on deep learning techniques such as Retina U-Net and Convolutional Neural Networks.

3 Proposed lung nodule detection methodology

The main objective of the proposed methodology is to effectively detect the pulmonary nodule from the CT lung images. To achieve this objective, in this paper we proposed, an optimized MRCNN. The proposed MRCNN is enhanced by using a hybrid particle snake swarm optimization algorithm (PS^2OA). The hybrid PS^2OA algorithm is used to tune the hyper-parameter. The proposed approach consists of two main stages namely, pre-processing and detection. The pre-processing is done by adaptive median filter and CLAHE. After pre-processing the image is given to the Optimized MRCNN, in which the nodule is detected. The structure of the proposed methodology is given in Fig. 1.

3.1 Pre-processing

The original Lung CT images are pre-processed to remove the noise and enhance image contrast. For this, we apply Adaptive Right Median Filter (AMF) and Contrast Limited Adaptive Histogram equalization (CLAHE) methods.

3.1.1 Adaptive median filter (AMF)

The purpose of using the Adaptive median filter is to reduce the distortion and preserve the image edge details [28]. The advantage of utilizing an adaptive median filter over a standard median filter is that the kernel size is adjustable in the area around the distorted image, as a result, we can obtain better output and in contrast to the median filter, it will not replace all of the pixel values with the median value. This algorithm works on two levels.
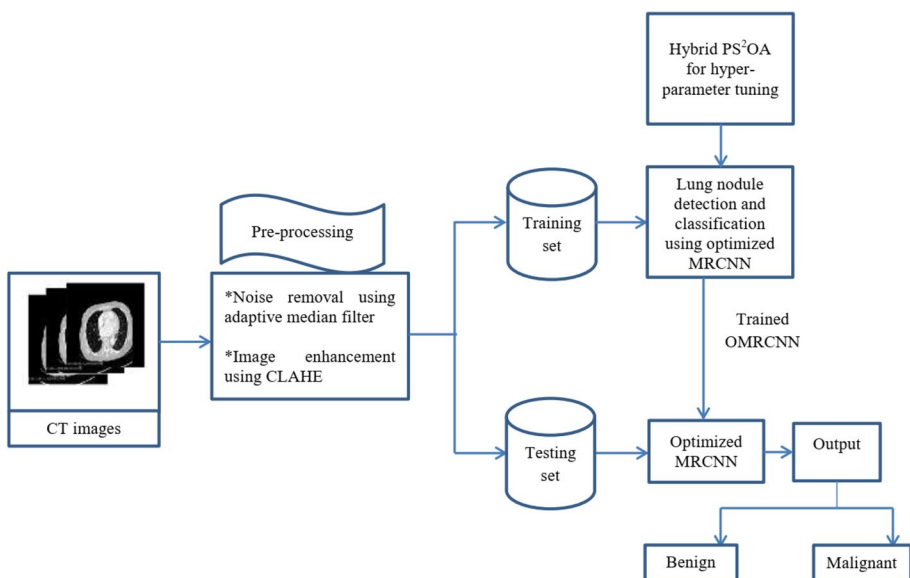

Fig. 1 Workflow of proposed methodology

Level: 1 The first level involves determining the kernel's median value.

$P_1 = Z_{min} - Z_{med}$

$P_2 = Z_{max} - Z_{med}$

If $P_1 > 0$ and $P_2 < 0$ go to Level 2

Else increase the kernel size

If kernel size $< = S_{max}$ iterate level 1

Else output Z_{xy}

S_{xy}—The local region of the gray level image at x,y.

Z_{min}, Z_{max}—Minimum and maximum gray level value in S_{xy}.

Z_{med}—Median gray level value in S_{xy}.

Z_{xy}—Gray level coordinates at x,y

S_{max}—the maximum allowed size of the region S_{xy}.

Level: 2 In level 2 determine whether the current pixel value is an impulse (salt and pepper noise) or not. If a pixel's value is corrupted, it either modifies it using the median or keeps the grayscale pixel value.

$Q_1 = Z_{med} - Z_{min}$

$Q_2 = Z_{med} - Z_{max}$

If $Q_1 > 0$ and $Q_2 < 0$ output Z_{xy}

Else return Z_{med}

Here the original CT Lung image of size 224×224 is shown in Fig. 2 (a) and the S_{max} maximum window size is assigned as 11. First, it is converted into a grayscale image as depicted in Fig. 2 (b). Then the AMF method is employed on that image for denoising, which is shown in Fig. 2 (c).

3.1.2 Contrast Limited Adaptive Histogram Equalization (CLAHE)

Histogram equalization is a method of processing images that modifies the intensity distribution of the histogram to change the contrast of an image. CLAHE is the variant of Adaptive Histogram Equalization (AHE). It reduces amplified noise by limiting contrast amplification. It performs this by evenly spreading the portion of the histogram that exceeds the clip limit across all histograms.

Histogram equalization (HE) is a technique that is often employed in image enhancement approaches; however, HE raises contrast globally, [29] whereas AHE is a technique that improves contrast in the local area. Unfortunately, AHE happens infrequently, which raises the contrast. The CLAHE method can handle this by providing a clip limit that specifies the maximum height of a histogram and region size. Here the denoised images are enhanced by having the clip limit as 0.02 and the tile size is assigned as 8X8. Clip limit gives the contextual region of the CLAHE [30]. The Rayleigh distribution method is used here to enhance the intensity values in every pixel. The bilinear interpolation method is
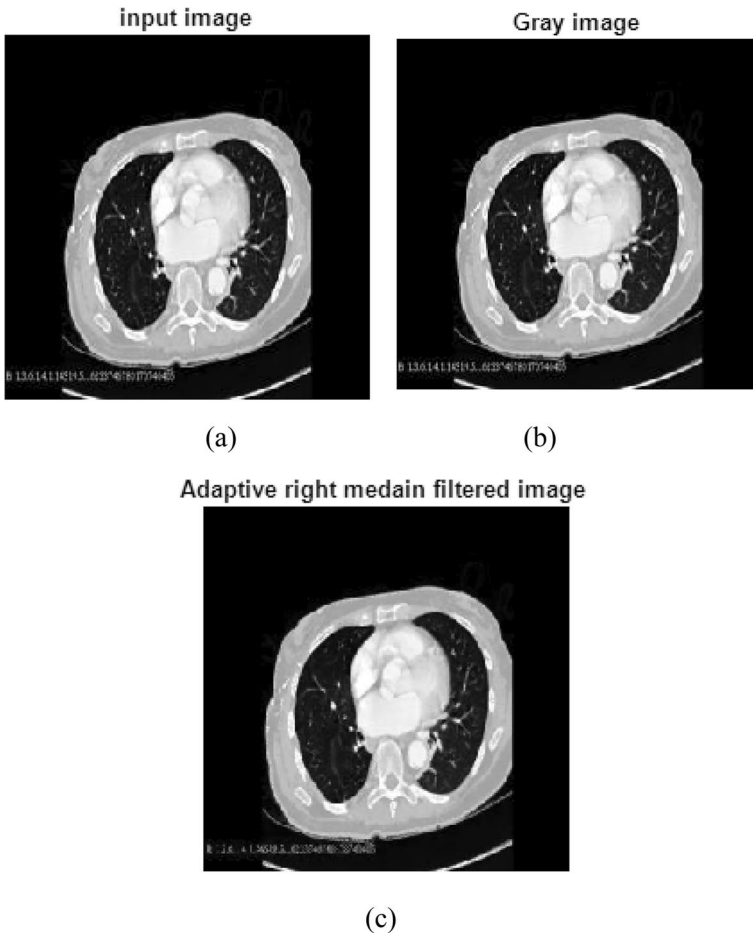
Fig. 2 Pre-processing output

used to remove the artifacts near the boundary of the tiles. HE is based on a transformation function, which is a combination of a probability distribution function (PDF) and a cumulative distribution function (CDF). The general histogram stretching is given by Eq. 1,

$$P_{out} = P_{in} - I_{miin}\left(\frac{Omax - Omin}{Imax - Imin}\right) + O_{min} \tag{1}$$

where P_{out} and P_{in} are the pixel value of the input image, I_{min}, I_{max} O_{min}, and O_{max} is the input and output images' respective minimum and maximum intensity levels.

$$PDF_{Rayleigh} = \frac{x}{a^2}e^{-\left(\frac{-x^2}{2a^2}\right)} \text{ for } x \geq 0, \ a \geq 0 \tag{2}$$

where x is the intensity value of the input image and α Rayleigh distribution parameter. Figure 3 depicts the Input CT lung image and Fig. 4 shows the image after applying the CLAHE technique. Figures 5 and 6 show the plot of Histogram Equalization and CLAHE.

Fig. 3 Input CT Lung Image

3.2 Lung nodule detection and classification using optimized Mask RCNN

After pre-processing, the pre-processed images are sent into an optimized Mask RCNN classifier to classify an image as malignant or benign**.** Mask R-CNN is a deep learning-based approach that is mainly used for object detection and image segmentation. In many computer vision tasks, including object detection, instance segmentation, and pose estimation, the Mask R-CNN algorithm has been extensively used. The mask RCNN generates the bounding box, segmentation mask, and corresponding class name. This works based on the Feature Pyramid Network (FPN) and a ResNet101 backbone. To enhance the performance of Mask RCNN, the hyper-parameter present in the Resnet is optimally selected. For the parameter selection process, a hybrid optimization algorithm is presented. For hybridization, particle swarm optimization and snake swarm optimization are presented. The structure of Mask RCNN is presented in Fig. 7.

   The Mask R-CNN model consists of three primary components which are the backbone, the Region Proposal Network (RPN), and RoIAlign. Backbone is a Feature Pyramid network-style deep neural network that can extract multi-level image features. The ResNet forms the backbone of the Mask R-CNN model. The CNN used here is ResNet 101. Further, it has 3.8×109 floating point operations. The RPN uses a sliding window to scan the input image and detects the infected regions in this study. The RoIAlign then examines the RoIs obtained from the RPN and extends the feature maps from the backbone at various locations. The RoIAlign is responsible for the formation of the precise segmentation masks on the images. The RoIPooling in Faster R-CNN is replaced by a more precise and accurate segmentation using the RoIAlign.
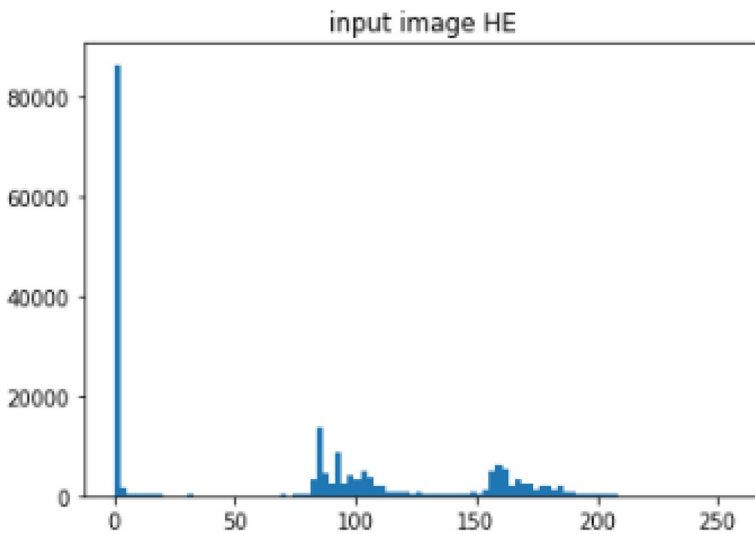
Fig. 4 CLAHE Image
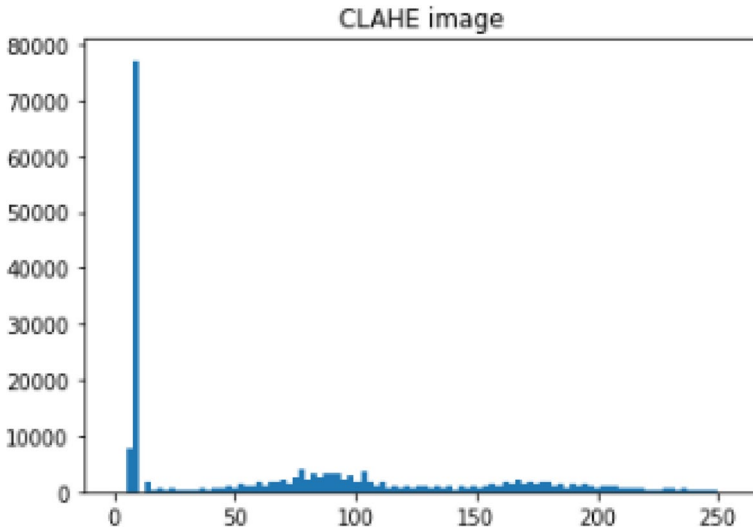

Fig. 5 HE plots of Input Lung CT

Fig. 6 Plot of CLAHE

3.2.1 ResNet 101 + FPN-based feature map generation

ResNet-FPN is a backbone architecture used for feature extraction in Mask R-CNN. ResNet is a deep convolutional neural network that is very effective for image classification tasks. While FPN builds an in-network feature pyramid out of a single-scale input, it uses a top-down architecture with lateral connections. In ResNet-FPN, FPN architecture is added on top of the ResNet backbone to create a more effective feature extractor. The FPN component allows for multi-scale feature maps to be generated from the input image, which can improve object detection accuracy [14].

ResNet sets up a series of convolution, polling, and activation FC layers one after the other. There are many types of ResNet architectures available, in this paper, ResNet 101 is utilized. ResNet 101 consists of 101 layers. The proposed ResNet 101 has lower complexity compared to VGG16 and VGG19 nets [11]. ResNet has three versions namely, ResNet Version 1, ResNet Version 2, and ResNeXt. Each version has different characteristics.

As shown in Fig. 8, the ResNet-101 has a bottom-up path, which reduces the resolution of the feature image. In contrast to ResNet-101, FPN improves the resolution of feature images from the top down. Lateral links between ResNet-101 and FPN combine features with the same resolution from ResNet-101 and FPN, respectively, to create new features in FPN [10]. In this, two features with the same resolution from ResNet101 and FPN combine to create a new feature in the path of FPN, and the ResNet101 backbone with FPN is used to train models for Lung CT images. The ResNet-101 is trained to optimize the following parameters;

- **ResNet Version:** ResNet 101 consists of a number of versions. The best version gives the proper output. So, we select the optimal version for the segmentation process.
- **Batch Size:** We can select any size of batch for processing. This will affect the performance. So, we choose the optimal batch size.
- **Pooling type:** Different types of pooling are available. So, we chose the optimal one.
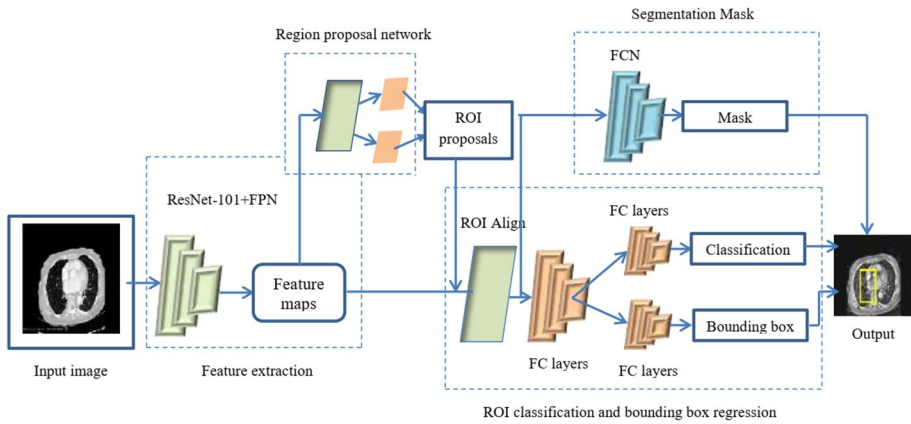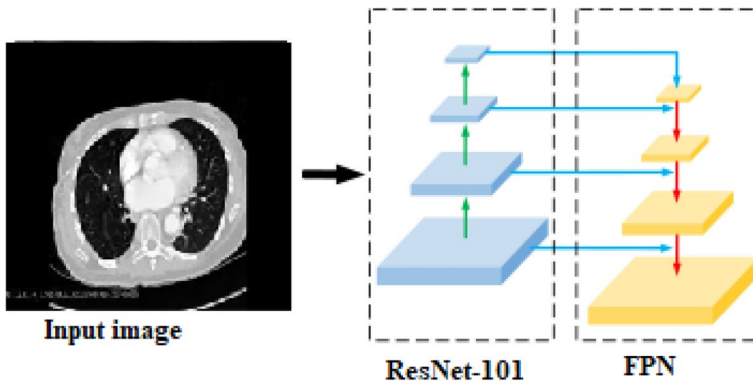
Fig. 7 Architecture of mask R-CNN

Fig. 8 ResNet-101 + FPN model

- **Learning rate:** For maximizing the final accuracy, the learning rate is another crucial factor. It can be difficult to determine the proper learning rate.
- **Optimizer:** The optimizer is used in the fully connected layer.

The above-mentioned five parameters are optimally selected by using the PS^{2}OA. The range of overlay parameter configurations used for the PS^{2}OA algorithm is shown in Table 1.

Step 1: Solution encoding: Solution initiation is an important factor in the PS^{2}OA, which is used to define the problem. In this paper, random initialization is used. Here, the parameters present in the ResNet101 namely, the version of ResNet, batch size, pooling type, learning rate, and optimizer are optimally selected by using the PS^{2}OA algorithm. Initially, these parameters are randomly initialized. In the PS^{2}OA, the solutions are called swarms, and the parameters are called particles. The initial solution format is given in Eq. (3).

$$S_i = \{S_1, S_2, \ldots, S_n\} \tag{3}$$

where S_n represent the nth swarm.

Table 1 Hyper-parameter range

Hyper-parameter	Range
Version	[v1, v2, v3]
Batch size	[32]
Pooling type	[average, maximum, minimum]
Learning rate	[0.1, 0.01, 0.001]
Optimizer	['adam', 'rmsprop', 'sgd']

Step 2: Fitness calculation: After the random creation, the fitness function is calculated for each swarm. In this paper, maximum accuracy is considered as the fitness function. The fitness function is given in Eq. (4).

$$Fitness = Max(Accuracy) \tag{4}$$

Step 3: Update the solution using hybrid Particle Snake Swarm Optimization: After the fitness calculation, we update the solution. For updating, in this paper PS^2OA algorithm is used. PS^2OA is a combination of PSO and SSO.

3.2.2 Particle Swarm Optimization

PSO is initialized based on fish schools and bird swarms in nature with a swarm intelligence method. Every velocity vector and position vector in PSO is defined as a particle. Each particle has its traverses and conducts a search space aligned with the best solution. Particles already know the optimal location the complete particle swarm has computed. The location and velocity vector updating process is formulated as follows,

$$v_i^{k+1} = v_i^k + c_1 r_1 \left(pbest_i^k - x_i^k\right) + c_2 r_2 \left(g_{best} - x_i^k\right) \tag{5}$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \tag{6}$$

3.2.3 Snake Swarm Optimization

Snake swarm optimization was developed in 2022 to reduce the mating characteristics of snakes [35]. Mating is achieved when food is available and at low temperatures.

Stage 1: Initialization Phase:

The random initial population of the SSOA is presented as follows,

$$N_{male} \approx \frac{n}{2} \tag{7}$$

$$N_f = n - N_{male} \tag{8}$$

Here, N_f is defined as female individuals, n is defined as the number of individuals and N_{male} is defined as male individuals. The random initialization of the SSOA is divided into two clusters such as male and female. In every iteration, the optimal individual candidate

solution is computed by validating every group for optimal female and optimal male. The food quantity and temperature are described as follows,

$$t = exp\left(\frac{-g}{t}\right) \tag{9}$$

$$fq = c_1 exp\left(\frac{g-t}{t}\right) \tag{10}$$

Here, c_1 is defined as constant to 0.5, t is defined as the total number of iterations, g is defined as the current iteration. When $fq < threshold$, the snakes search for food by choosing a random position and after that upgrade their position.

Stage 2: Male snake formulation:

To numerically design the exploration characteristics of the female and male snakes, it is used,

$$x_{i,j}(g+1) = x_{\left(rand\epsilon\left[1,\frac{n}{2}\right],j\right)}(g) \pm c_2 \times a_{i,male}(ub-lb) \times rand_{\epsilon U(0,1)} + lb, Here a_{i,male}$$
$$= exp\left(-\frac{f_{rand,male}}{f_{i,male}}\right) \tag{11}$$

Here, $\pm$ is a flag direction operator, $f_{i,male}$ is defined as the fitness of the male in the group, $f_{rand,male}$ is defined as the fitness of the earlier chosen random male snakes, $a_{i,male}$ is defined as the capability to compute the food by male, $rand$ is defined as the random number between 0 and 1, $x_{\left(rand\epsilon\left[1,\frac{n}{2}\right],j\right)}$ is defined as the position of a random male snake and $x_{i,j}$ is defined as the male snake position.

Stage 3: Female snake formulation

$$x_{i,j} = x_{\left(rand\epsilon\left[1,\frac{n}{2}\right],j\right)}(g+1) \pm c_2 \times a_{i,female}(ub-lb) \times rand_{\epsilon U(0,1)} + lb, Here a_{i,female}$$
$$= exp\left(-\frac{f_{rand,female}}{f_{i,female}}\right) \tag{12}$$

Here, $\pm$ is a flag direction operator, $f_{i,male}$ is defined as the fitness of the female in the group, $f_{rand,female}$ is defined as the fitness of the earlier chosen random female snakes, $a_{i,female}$ is defined as the capability to compute the food by female, $rand$ is defined as the random number between 0 and 1, $x_{\left(rand\epsilon\left[1,\frac{n}{2}\right],j\right)}$ is defined as the position of a random female snake and $x_{i,j}$ is defined as the male snake position. Female snake formulation is given as

Stage 4: Exploration phase:

In the exploitation phase, two scenarios are considered for computing optimal solutions. This condition is developed based on threshold parameters.

4 Condition 1:

$FQ < threshold$, it is updated by the below equation

$$x_{i,j}(g+1) = x_{food} \pm c_3 \times t \times rand \times \left(x_{food} - x_{i,j}(g)\right) \tag{13}$$

Here, c_3 is equivalent to 2, x_{food} is defined as the position of the optimal individuals and $x_{i,j}$ is defined as the position of individuals.

5 Condition 2:

$FQ > threshold$, it is updated based on the fighting and mating process.

Fighting Process.

The fighting capability of a female snake is formulated as follows,

$$x_{i,j}(g+1) = x_{i,j}(g) \pm c_3 \times f_{i,female} \times rand \times \left(x_{best,male} - x_{i,f}(g+1)\right) where\, f_{i,female} = exp\left(\frac{-f_{best,male}}{f_i}\right)$$

(14)

Here, $f_{i,female}$ is defined as the fighting capability of the female snake, $x_{best,male}$ is defined as the position of the best individual in the male group and $x_{i,j}$ is defined as the female position. The fighting capability of the male snake is formulated as follows,

$$x_{i,j}(g+1) = x_{i,j}(g) \pm c_3 \times f_{i,male} \times rand \times \left(x_{best,female} - x_{i,male}(g)\right) where\, f_{i,male} = exp\left(\frac{-f_{best,f}}{f_i}\right)$$

(15)

Here, $f_{i,male}$ is defined as the fighting capability of the male snake, $x_{best,female}$ is defined as the position of the best individual in the female group and $x_{i,j}$ is defined as the male position.

6 Mating process

In this phase, the female and male groups can upgrade their position,

$$x_{i,female}(g+1) = x_{i,f}(g) \pm MM_{i,female} \times rand \times \left(q \times x_{i,male} - x_{i,female}(g+1)\right), where\, MM_{i,female} = exp\left(\frac{-f_{i,male}}{f_{i,female}}\right)$$

(16)

$$x_{i,male}(g+1) = x_{i,m}(g) \pm MM_{i,male} \times rand \times \left(q \times x_{i,female} - x_{i,male}(g+1)\right), where\, MM_{i,male} = exp\left(\frac{-f_{i,female}}{f_{i,male}}\right)$$

(17)

Here, $MM_{i,female}$ is defined as the mating capability of a female, $MM_{i,male}$ is defined as the mating capability of male, $x_{i,m}(g)$ is defined as the position of male and $x_{i,f}(g)$ is defined as the position of female agents.

7 Proposed PS²OA

The major motive of the hybrid algorithm is to enhance the method's ability to utilize PSO while also exploring SSOA to achieve the optimization strength of both. The exploration and exploitation of the SSOA were managed by the inertia constant in the hybrid algorithm. Compared with the conventional computations, the primary agent's location in the hunting location is optimally upgraded. This is presented as follows,

$$Female\,(mating) = x_{i,f}(g) \pm MM_{i,female} \times rand \times \left(q \times x_{i,male} - x_{i,female}(g+1)\right), where\,MM_{i,female} = exp\left(\frac{-f_{i,male}}{f_{i,female}}\right)$$

(18)

$$Male\,(mating) = x_{i,m}(g) \pm MM_{i,male} \times rand \times \left(q \times x_{i,female} - x_{i,male}(g+1)\right), where\,MM_{i,male} = exp\left(\frac{-f_{i,female}}{f_{i,male}}\right)$$

(19)

The location and velocity are adjusted to combined PSO and SSO variations and are presented as follows,

$$v_i^{k+1} = w.(v_i^k + c_1 r_1\left(x_{i,female} - x_i^k\right)) + c_2 r_2\left(x_{i,male} - x_i^k\right)$$

(20)

$$x_i^{k+1} = x_i^k + v_i^{k+1}$$

(21)

Step 4: Termination criteria

The procedure is continued until the best hyper-parameter values are selected. The selected value is given to the lung cancer detection process.

Algorithm 1 Pseudocode of the proposed hybrid algorithm

Start

    Initialize the random parameters

    Describe the number of iterations

    Compute fitness function equation (3)

    Update velocity and position vector by considering the SSOA updating process

If $FQ < threshold$

        $x_{i,j}(g+1) = x_{food} \pm c_3 \times t \times rand \times \left(x_{food} - x_{i,j}(g)\right)$ (Updating process)

End if

If $FQ < threshold$

      Fighting mode

        Female and male snakes are upgraded by using equations (14) and (15)

      Mating mode

        Female and male snakes are upgraded by using equations (16) and (17)

End if

      Velocity and location are adjusted by using equations (18) and (19)

      Compute adjusted velocity and position vector by using equations (20) and (21)

      Check the termination condition

      Save the optimal results

End

---

8 Region proposal network (RPN)

The RPN network utilizes the features extracted by ResNet101+FPN as input to generate Regions of Interest (ROIs). In scenarios where the aspect ratios of objects differ, RPN can predict both the foreground and background of an image. To efficiently generate candidate regions, the image box is positioned on the network, delineating the border box in the expected feature image. A 3×3 convolutional layer scans the image, generating anchors distributed across the image in different sizes. These anchors serve as starting points for proposing potential regions of interest, facilitating subsequent object detection processes.

The network adapts its scale based on input images and utilizes a predefined set of anchor boxes [15]. Each anchor corresponds to a unique bounding box and ground-truth class, allowing for the recognition of defects of diverse sizes and shapes. Default bounding boxes encompass a range of sizes and aspect ratios to accommodate various object characteristics. With overlapping bounding boxes, determining the highest confidence score for detecting multiple Regions of Interest (ROIs) becomes more straightforward [20]. This assessment is facilitated by the Intersection over Union factor (IoU), calculated using equation (22), aiding in the accurate identification of RoIs.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ union} \tag{22}$$

9 ROI align model

RoIAlign processes a set of rectangular region proposals, extracting features from the feature map corresponding to each proposal. In RCNN networks, pixel accuracy and the ability to distinguish individual branches within the same pixel target are crucial for mask branch detection. After pooling and convolution of the original image, the image size undergoes changes, followed by segmentation. Direct pixel-level segmentation techniques often fail to produce accurate segmentation output. Hence, this paper proposes Mask RCNN, an enhanced version of Faster RCNN. Additionally, CNN's pooling layer is replaced with RoIAlign, which utilizes linear interpolation to preserve spatial details in the feature map. RoIAlign serves as a neural network layer employed in object detection and instance segmentation algorithms, such as Mask R-CNN.

In Fig. 9, the green dotted lines are referred to as the 5X5 feature diagram, which is derived after the convolution layer, and the feature corresponding to the ROI in the solid line feature diagram is smaller, and RoIAlign maintains a floating-point number boundary without scale processing. Initially, the feature's small volume was separated into 2X2 units (each unit boundary was not measured) and then each unit was separated into four smaller units; the center point is illustrated as a four-coordinate position blue dot in the figure. After that, two linear interpolations are performed to calculate the values of the four levels, followed by maximum pooling or average voting to generate a 2×2 scale feature map.

10 Loss function

The loss function in Mask R-CNN, a popular instance segmentation model, is a composite function that combines classification, bounding box regression, and mask segmentation losses. It serves to optimize the model parameters by minimizing the discrepancy between
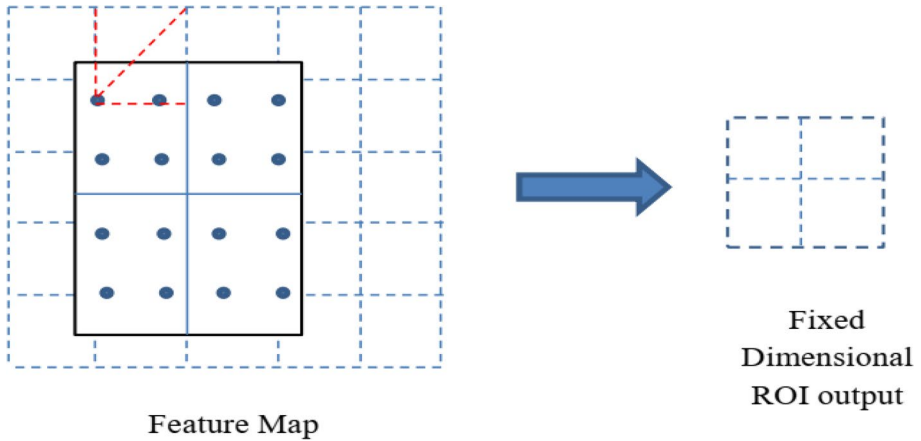
Fig. 9 Schematic diagram of the RoIAlign algorithm

predicted and ground-truth values for object classification, localization, and pixel-wise segmentation simultaneously. This loss function plays a pivotal role in training Mask R-CNN to accurately identify object instances and their corresponding masks in images. The loss function of the proposed model is given in Eq. (23).

$$L(OMRCNN) = Loss_{Class} + Loss_{Box} + Loss_{Mask} \tag{23}$$

where; prediction loss of the presented class label is represented as $Loss_{class}$, the loss of bounding box is represented as $Loss_{Box}$, and the presented segmentation mask loss is represented as $Loss_{mask}$. The $Loss_{class}$ is calculated based on the normal image and affected image. The mathematical expression of L^{Class} is given in equation (24).

$$Loss_{class}\left(A_i, A_i^*\right) = -\log\left[A_i A_i^* + \left(1 - A_i\right)\left(1 - A_i^*\right)\right] \tag{24}$$

where, A_i represents the candidate anchor i target prediction probability of having a disease and is the ground-truth label which is 1 for the positive anchor, otherwise 0. The below equation describes the regression loss of the bounding box function;

$$Loss_{Box}\left(B_i, B_i^*\right) = \sum_{i \in \{x,y,w,h\}} Smooth_{L_1}\left(B_i - B_i^*\right) \tag{25}$$

where;

$$Smooth_{L1}(X) = \begin{cases} 0.5x^2, & if\,|x| < 1 \\ |x| - 0.5, & otherwise \end{cases} \tag{26}$$

predicted bounding box is defined as B_i, the GT based positive anchor is defined as B_i^*. The loss function of the mask is calculated as below;

$$LossMask = -\frac{1}{n^2} \sum_{1 \le i,j \le n} \left[X_{ij} \log O_{ij}^t + \left(1 - X_{ij}\right)\log\left(1 - O_{ij}^t\right)\right] \tag{27}$$

where X_{ij} represent the value of a pixel (i,j) in a ground-truth mask of size n x n and O_{ij}^t is the predicted value of the same pixel in the mask learned for class ($t = 1$ *for Malignant and* 0 *for Benign*).

11 Results and discussion

The experimental results obtained by the proposed approach are presented in this section. The proposed method is executed in TensorFlow and performance is analyzed. The analysis was executed on Google colab in "Keras 2.3.1" with the "TensorFlow 1. Upon which system the experiment was conducted. "Windows 10" and had a "Random-Access Memory (RAM) of 8 GB" and "Graphics Processing Units (GPUs)" are used in this experiment. The performance of the proposed approach is analyzed based on different metrics namely, accuracy, precision, recall, and F-Measure.

11.1 Dataset description

Seven research institutions and eight private medical imaging businesses have collaborated to create a public dataset called the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) [26]. Ten hundred and eighty computed tomography (CT) scans were included in the database, with slice thicknesses ranging from 0.6 mm to 5.0 mm. Four radiologists read through these scans in two separate reading sessions. First, radiologists identified potentially malignant lesions and divided them into three categories based on their size (nodules > = 3 mm, nodules < 3 mm, and non-nodules). The results from all four radiologists were compiled, and then each radiologist unblinded and rechecked every annotation. In practical practice, detecting lung nodules requires scans with thin slices. Therefore, scans with slice thickness greater than 2 mm were not considered. So the author included a total of 888 images included in for the analysis, after excluding those with inconsistent slice spacing [27]. According to NLST screening criteria, nodules larger than 3 mm were judged to be significant lesions.

12 Experimental results

In this section, we presented the visual representation of the proposed experimental results.
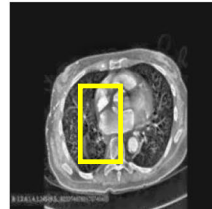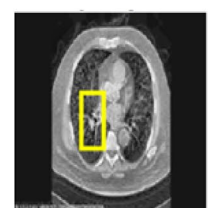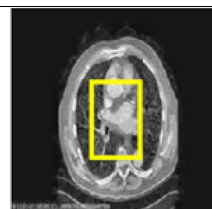
Table 2 represent the visual representation of the detection output. In Fig. 10, we analyze the accuracy performance by varying epochs, and in Fig. 11, we analyze the loss s by varying epochs. According to Fig. 15, we understand that as the number of epochs increases, the loss value decreases.

12.1 Comparative analysis results

In this section, we compare our proposed work performance with different detection models namely, Faster RCNN (FRCNN), Single Shot MultiBox Detector (SSD), YoLo model and SVM-based lung nodule detection.

In Fig. 12, the performance of the proposed approach is analyzed based on an accuracy measure. When analyzing Fig. 12, the proposed method attained a maximum accuracy of

Table 2 Visual representation of detection output, column (a) represent the input image, (b)represent the Gray image, (c) represent the adaptive median filtered image and (d) detected output

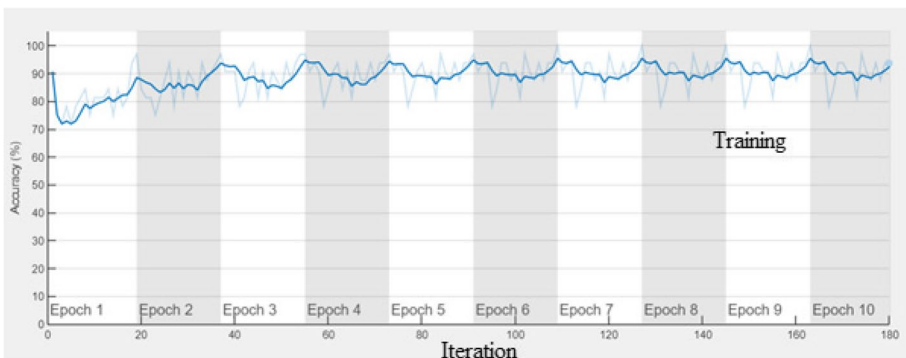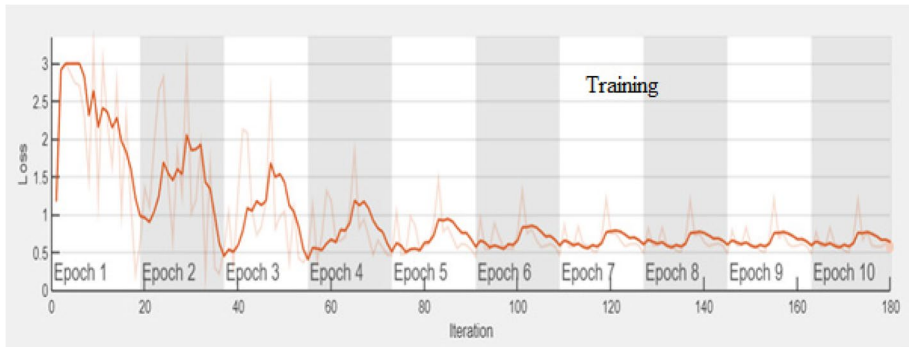| Input image	Gray image	Adaptive median filtered image	Detected output

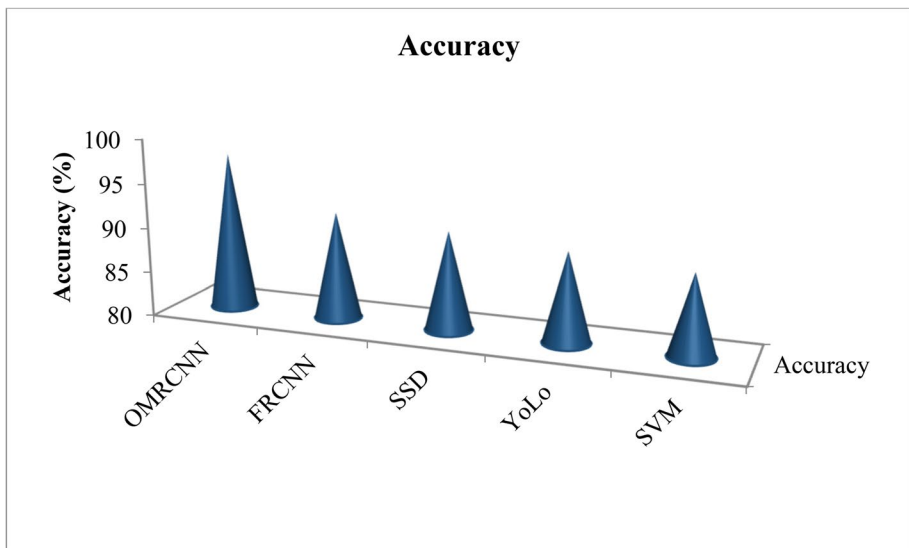**Fig. 10** Epoch vs accuracy

Fig. 11 Epoch Vs loss

Fig. 12 Performance analysis based on accuracy

97.67% and ANN-based lung nodule detection attained an accuracy of 89%. Compared to five existing classifiers, SVM-based classification attained the worst results. Due to hyper-parameter optimization in MRCNN, our proposed method attained better results compared to the existing techniques. In Fig. 13, the performance of the proposed approach is analyzed based on precision. A good classification should have the maximum precision value. When analyzing Fig. 13, the proposed method attained the maximum precision of 95.7% which is 2.6% better than FRCNN-based lung nodule classification, 4.5% better than SSD-based lung nodule classification, 6.2% better than YoLo-based classification and 8.2% better than SVM based lung nodule classification. The performance of the presented technique is analyzed based on recall is given in given in Fig. 14. As per Fig. 14, we understand that ORCNN-based lung nodule classification attained the maximum recall value compared to the existing techniques. Similarly, we attained the maximum F-score value shown in Fig. 15. From the results, we can understand proposed approach attained the maximum output compared to the existing techniques.
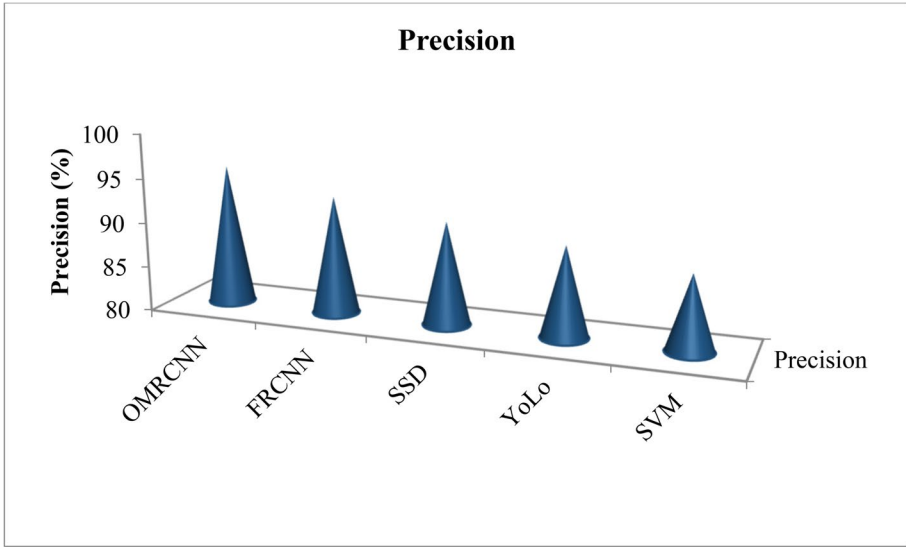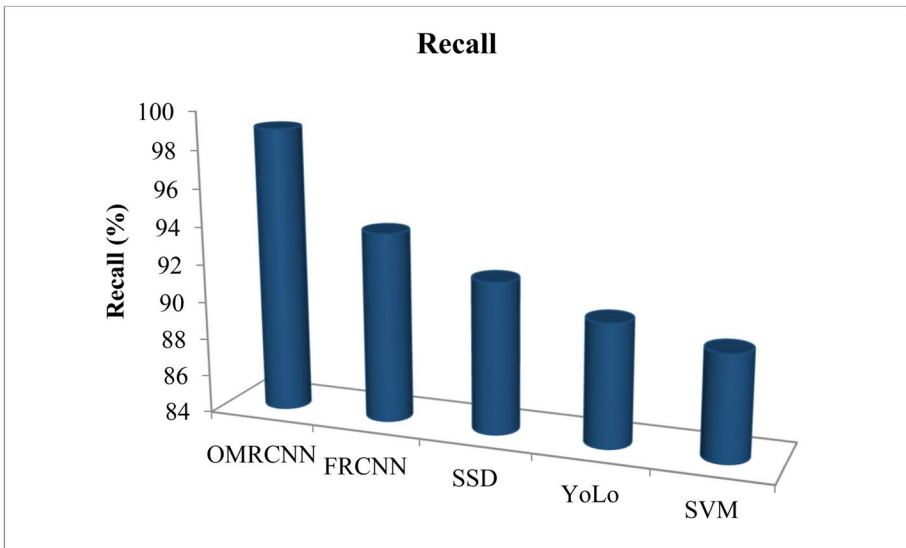
Fig. 13 Performance analysis based on precision

Fig. 14 Performance analysis based on Recall

12.2 Comparative analysis with published work

To prove the efficiency of our proposed approach, we compare our work with already published research works. For comparative analysis, we considered four research works namely 3DCNN [31], CNN [32], texture CNN [33], and BCNN [34]. These four techniques are deeply explained lung nodule classification. So, we compare our research work with these papers.
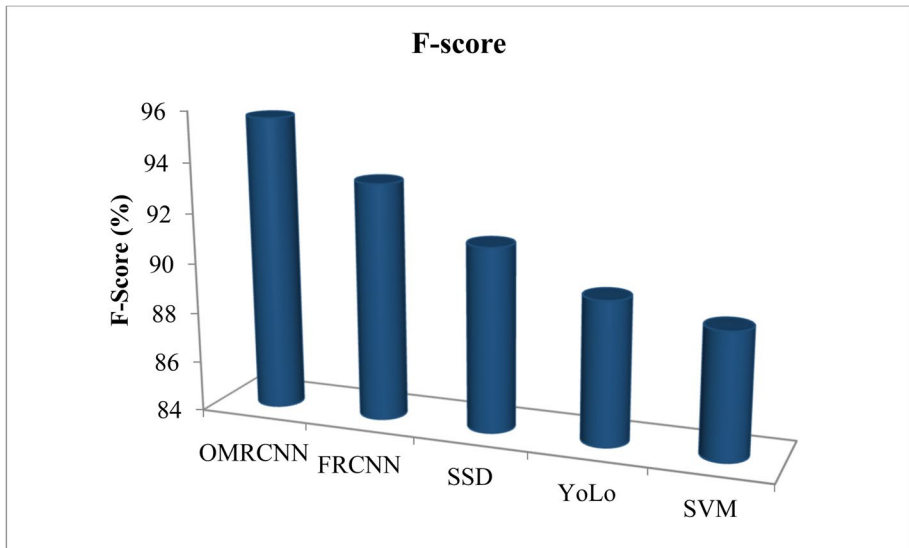
Fig. 15 Performance analysis based on F-score

Table 3 Comparative analysis results

References	Accuracy	Recall	precision	F-score	Dataset
MMEL-3DCNN [31]	90.6	83.7	-	-	LIDC-IDRI
CNN[32]	87.26	81	87.8	-	LIDC- IDRI
Texture CNN [33]	90.91	91.39	90.46	94.14	MNIST dataset
BCNN [34]	91.46	91.94	-	93.35	LUNA16
Proposed	97.67	99	95.7	95.67	LIDC-IDRI

The comparative analysis result is presented in Table 3. In this paper, for lung nodule classification, we optimized MRCNN. To improve the performance MRCNN classifier, the hyper-parameters are optimally selected using a hybrid PS^2OA algorithm. To prove the efficiency, we compare our work with [31–34]. When analyzing Table 3, our proposed approach attained the maximum accuracy of 97.67% which is 90.6% for [31], 87.26% for [32], 90.91% for [33], and 91.46% for [34]. Due to MRCNN and hyper-parameter optimization, our method produces superior classification outcomes.

13 Conclusion

In this section, the proposed methodology is effectively detecting the pulmonary nodule from the CT lung images. To achieve this objective, in this paper we proposed, an optimized MRCNN. The proposed MRCNN is enhanced by using a hybrid PS^2OA. The hybrid PS^2OA algorithm is used to tune the hyper-parameter. The proposed approach consists of two main stages namely, pre-processing and detection. The pre-processing is done by adaptive median filter and CLAHE. After pre-processing the image is given to the Optimized MRCNN, in which the nodules are detected. The performance of the proposed approach is analyzed based on different metrics and effectiveness compared with state-of-the-art works.

The performance analysis of accuracy is 94.67%, the recall value is 99%, the precision value of the proposed method is 95.7% and the f-score value of the proposed method is 95.67%. Overall, this integrated approach holds significant promise in enhancing the efficiency and accuracy of lung cancer detection, thereby contributing to improved patient outcomes and advancing the field of computer-aided diagnosis systems.

Declarations

Ethical Approval Not Applicable.

Information consent Not Applicable.

Conflicts of interest The authors have no financial or proprietary interests in any material discussed in this article. The authors declare that they have no conflict of interest.

References

1. Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians, 71(3), 209–249, doi: https://doi.org/10.3322/caac.21660.
2. Liu K (2022) Stbi-yolo: A real-time object detection method for lung nodule recognition. IEEE Access 10:75385–75394. https://doi.org/10.1109/ACCESS.2022.3192034
3. Shariaty F, Mousavi M (2019) Application of CAD systems for the automatic detection of lung nodules. Informatics in Medicine Unlocked 15:100173. https://doi.org/10.1016/j.imu.2019.100173
4. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961–2969). https://doi.org/10.48550/arXiv.1703.06870
5. Poojary, R., & Pai, A. (2019, November). Comparative study of model optimization techniques in fine-tuned CNN models. In 2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA) (pp. 1–4). IEEE. DOI: https://doi.org/10.1109/ICECTA48151.2019.8959681
6. Cui H, Bai J (2019) A new hyperparameters optimization method for convolutional neural networks. Pattern Recogn Lett 125:828–834. https://doi.org/10.1016/j.patrec.2019.02.009
7. Loussaief, S., & Abdelkrim, A. (2018). Convolutional neural network hyper-parameters optimization based on genetic algorithms. International Journal of Advanced Computer Science and Applications, 9(10), doi: https://doi.org/10.14569/IJACSA.2018.091031
8. https://neptune.ai/blog/hyper-parameter-tuning-in-python-complete-guide
9. Gaspar, A., Oliva, D., Cuevas, E., Zaldívar, D., Pérez, M., & Pajares, G. (2021). Hyperparameter optimization in a convolutional neural network using metaheuristic algorithms. In Metaheuristics in Machine Learning: Theory and Applications (pp. 37–59). Cham: Springer International Publishing, doi: https://doi.org/10.1007/978-3-030-70542-8_2
10. Da Silva GLF, Valente TLA, Silva AC, De Paiva AC, Gattass M (2018) Convolutional neural network-based PSO for lung nodule false positive reduction on CT images. Comput Methods Programs Biomed 162:109–118. https://doi.org/10.1016/j.cmpb.2018.05.006
11. Xue J, Shen B (2020) A novel swarm intelligence optimization approach: sparrow search algorithm. Systems science & control engineering 8(1):22–34. https://doi.org/10.1080/21642583.2019.1708830
12. Clerc, M. (2010). Particle swarm optimization (Vol. 93). John Wiley & Sons. https://kamenpenkov.files.wordpress.com/2016/01/pso-m-clerc-2006.pdf
13. http://www.ijcse.net/docs/IJCSE13-02-05-017.pdf

14. Rere LR, Fanany MI, Arymurthy AM (2015) Simulated annealing algorithm for deep learning. Procedia Computer Science 72:137–144. https://doi.org/10.1016/j.procs.2015.12.114
15. Kathpal S, Vohra R, Singh J, Sawhney RS (2012) Hybrid PSO–SA algorithm for achieving partitioning optimization in various network applications. Procedia engineering 38:1728–1734. https://doi.org/10.1016/j.proeng.2012.06.210
16. Chang, J., Li, Y., & Zheng, H. (2021). Research on key algorithms of the lung cad system based on cascade feature and hybrid swarm intelligence optimization for mkl-svm. Computational Intelligence and Neuroscience, 2021, DOI: https://doi.org/10.1155/2021/5491017
17. Zheng S, Cornelissen LJ, Cui X, Jing X, Veldhuis RNJ, Oudkerk M, van Ooijen PMA (2021) Deep convolutional neural networks for multiplanar lung nodule detection: Improvement in small nodule identification. Med Phys 48(2):733–744. https://doi.org/10.1002/mp.14648
18. Majidpourkhoei R, Alilou M, Majidzadeh K, Babazadehsangar A (2021) A novel deep learning framework for lung nodule detection in 3d CT images. Multimedia Tools and Applications 80(20):30539–30555. https://doi.org/10.1007/s11042-021-11066-w
19. Su, Y., Li, D., & Chen, X. (2021). Lung Nodule Detection based on Faster R-CNN Framework. Computer Methods and Programs in Biomedicine, 200. https://doi.org/10.1016/j.cmpb.2020.105866
20. Liu, M., Dong, J., Dong, X., Yu, H., & Qi, L. (2018, September). Segmentation of lung nodule in CT images based on mask R-CNN. In 2018 9th International Conference on Awareness Science and Technology (iCAST) (pp. 1–6). IEEE, DOI: https://doi.org/10.1109/ICAwST.2018.8517248
21. Cai L, Long T, Dai Y, Huang Y (2020) Mask R-CNN-Based Detection and Segmentation for Pulmonary Nodule 3D Visualization Diagnosis. IEEE Access 8:44400–44409. https://doi.org/10.1109/ACCESS.2020.2976432
22. Yeh, W. C., Lin, Y. P., Liang, Y. C., & Lai, C. M. (2021). Convolution neural network hyperparameter optimization using simplified swarm optimization. arXiv preprint arXiv:2103.03995, https://doi.org/10.48550/arXiv.2103.03995
23. Singh P, Chaudhury S, Panigrahi BK (2021) Hybrid MPSO-CNN: Multi-level particle swarm optimized hyperparameters of convolutional neural network. Swarm Evol Comput 63:100863. https://doi.org/10.1016/j.swevo.2021.100863
24. Vijh, S., Gaurav, P., & Pandey, H. M. (2020). Hybrid bio-inspired algorithm and convolutional neural network for automatic lung tumor detection. Neural Computing and Applications, 1–14, https://doi.org/10.1007/s00521-020-05362-z
25. Gunjan VK, Singh N, Shaik F, Roy S (2022) Detection of lung cancer in CT scans using grey wolf optimization algorithm and recurrent neural network. Heal Technol 12(6):1197–1210. https://doi.org/10.1007/s12553-022-00700-8
26. Armato, S. G., McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., ... & Clarke, L. P. (2011). The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. Medical physics, 38(2), 915–931, doi: https://doi.org/10.1118/1.3528204
27. National Lung Screening Trial Research Team. (2011). Reduced lung-cancer mortality with low-dose computed tomographic screening. New England Journal of Medicine, 365(5), 395–409, https://www.nejm.org/doi/full/https://doi.org/10.1056/nejmoa1102873
28. Verma K, Singh BK, Thoke AS (2015) An enhancement in adaptive median filter for edge preservation. Procedia Computer Science 48:29–36. https://doi.org/10.1016/j.procs.2015.04.106
29. Hana, F. M., & Maulida, I. D. (2021, February). Analysis of contrast limited adaptive histogram equalization (CLAHE) parameters on finger knuckle print identification. In Journal of Physics: Conference Series (Vol. 1764, No. 1, p. 012049). IOP Publishing, DOI https://doi.org/10.1088/1742-6596/1764/1/012049
30. Ghani ASA, Isa NAM (2014) Underwater image quality enhancement through Rayleigh-stretching and averaging image planes. International Journal of Naval Architecture and Ocean Engineering 6(4):840–866. https://doi.org/10.1186/2193-1801-3-757
31. Liu H, Cao H, Song E, Ma G, Xu X, Jin R, Liu C, Hung C-C (2020) Multi-model Ensemble Learning Architecture Based on 3D CNN for Lung Nodule Malignancy Suspiciousness Classification. J Digit Imaging 33:1242–1256. https://doi.org/10.1007/s10278-020-00372-8
32. Agnes, S. A., & Anitha, J. (2020, January). Automatic 2D lung nodule patch classification using deep neural networks. In 2020 Fourth International Conference on Inventive Systems and Control (ICISC) (pp. 500–504). IEEE, doi:https://doi.org/10.1109/ICISC47916.2020.9171183
33. Ali I, Muzammil M, Haq IU, Khaliq AA, Abdullah S (2020) Efficient lung nodule classification using transferable texture convolutional neural network. Ieee Access 8:175859–175870. https://doi.org/10.1109/ACCESS.2020.3026080

34. Mastouri R, Khlifa N, Neji H, Hantous-Zannad S (2021) A bilinear convolutional neural network for lung nodules classification on CT images. Int J Comput Assist Radiol Surg 16:91–101. https://doi.org/10.1007/s11548-020-02283-z
35. Hashim FA, Hussien AG (2022) Snake Optimizer: A novel meta-heuristic optimization algorithm. Knowl-Based Syst 242:108320. https://doi.org/10.1016/j.knosys.2022.108320
36. Sollini, M., Kirienko, M., Gozzi, N., Bruno, A., Torrisi, C., Balzarini, L., ... & Chiti, A. (2023). The Development of an Intelligent Agent to Detect and Non-Invasively Characterize Lung Lesions on CT Scans: Ready for the "Real World"?. Cancers, 15(2), 357, doi: https://doi.org/10.3390/cancers15020357