**ORIGINAL PAPER** 



# Lung Nodule Detection via Optimized Convolutional Neural Network: Impact of Improved Moth Flame Algorithm

Anuja Eliza Sebastian<sup>1</sup> · Disha Dua<sup>2</sup>

Received: 12 November 2021 / Revised: 30 September 2022 / Accepted: 2 November 2022 / Published online: 13 March 2023 © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

# Abstract

Lung cancer is a high-risk disease that affects people all over the world, and lung nodules are the most common sign of early lung cancer. Since early identification of lung cancer can considerably improve a lung scanner patient's chances of survival, an accurate and efficient nodule detection system can be essential. Automatic lung nodule recognition decreases radiologists' effort, as well as the risk of misdiagnosis and missed diagnoses. Hence, this article developed a new lung nodule detection model with four stages like "Image pre-processing, segmentation, feature extraction and classification". In this processes, pre-processing is the first step, in which the input image is subjected to a series of operations. Then, the "Otsu Thresholding model" is used to segment the pre-processed pictures. Then in the third stage, the LBP features are retrieved that is then classified via optimized Convolutional Neural Network (CNN). In this, the activation function and convolutional layer count of CNN is optimally tuned via a proposed algorithm known as Improved Moth Flame Optimization (IMFO). At the end, the betterment of the scheme is validated by carrying out analysis in terms of certain measures. Especially, the accuracy of the proposed work is 6.85%, 2.91%, 1.75%, 0.73%, 1.83%, as well as 4.05% superior to the extant SVM, KNN, CNN, MFO, WTEEB as well as GWO+FRVM methods respectively.

Keywords Lung cancer  $\cdot$  Pre-processing  $\cdot$  Segmentation  $\cdot$  LBP features  $\cdot$  Optimal CNN  $\cdot$  Measures

 Anuja Eliza Sebastian anujasebastiangdj@gmail.com
 Disha Dua disha.dua3@gmail.com

<sup>1</sup> T. John Institute of Technology, Bangalore, India

<sup>&</sup>lt;sup>2</sup> Indira Gandhi Delhi Technical University for Women, Delhi, Delhi, India

Abbreviations	
CAD	Computer-aided detection
CNN	Convolutional neural network
СТ	Computed tomography
DCNN	Deep CNN
FDR	False discovery rate
FMS	F-measure
FNR	False negative rate
FPR	False positive rate
GGO	Ground glass opacity
GWO+FRVM	Grey wolf optimizer + fuzzy relevance vector machine
LBP	Local binary pattern
KNN	K nearest neighbour
MCC	Matthews correlation coefficient
MFO	Moth flame optimization
NPV	Net predictive value
PNN	Probabilistic neural network
SVM	Support vector machine
TP	Training percentage
US	United States
WTEEB	Whale with tri-level enhanced encircling behavior

# **1** Introduction

Recently, lung cancer and COVID-19 [1] are two drastic pulmonary diseases, which cause millions of death globally each year. Lung cancer is said to be the 2nd most widespread form of cancer in both women as well as men and it is the primary cause of deaths occurring due to cancer in US [2–4]. The finest possibilities of survival emerge from earlier detection and diagnosis that could be aided by enhanced automated malignant nodule recognition techniques. A lung nodule will be round and it is a smaller growth of tissue found in the cavity of the chest. Nodules are usually below 30 mm in size, and outsized growths are termed as masses and are assumed to be malignant [5, 6]. Nodules among 5–30 mm might be malignant or benign, with the probability of malignancy rising with size. Spiculated or lobulated nodule edges might specify malignancy whereas Smooth nodules with indications of calcifications are expected to be benign.

There are two most important chest imaging methods, fundamental X-ray imaging and CT [7, 8]. Radiographs or chest X-ray images offer a single outlook on the chest cavity. Poster anterior analysis, where the X-ray beam passes over the chest of the patient from back to front is general. CT scans are 3-D images generated by means of X-ray images obtained from several orientations and it could offer an entire view of the internals parts of the chest and can, therefore, be exploited for easily detecting the shapes, sizes, locations, and densities of lung nodules [9, 10]. Nevertheless, CT scan equipment is highly-priced and is often not obtainable in rural areas or minor hospitals. Moreover, radiographs are comparatively fast and cheap, and the patients are exposed to minute radiation, hence they are typically the initial diagnostic step for identifying any abnormalities in the chest [11, 12].

CAD methods were deployed to identify the lung nodules more precise and quicker. Nodule recognition approaches [13–15] are modeled by conventional image processing schemes to discover areas of the chest radiograph, which encloses a brighter object of the expected texture, shape, and size of a lung nodule [16]. With current enhancements in CNNs, certain researchers have aimed at exploiting these techniques to categorize lung nodules. Unluckily, the accessible datasets are comparatively low in medical imaging [17].

The literature on detecting and diagnosing lung nodules is extensive. To date, the general technique for lung nodule diagnosis in all existing CAD systems has been to utilize a candidate identification stage. While some of these researches use low-level appearance-based variables to drive this identification task [18], others use shape and size information. Ypsilantis et al. proposed using recurrent neural networks in a patch-based strategy to improve nodule detection [19], which is related to deep learning-based methodologies. A 2D multi-step segmentation approach was presented by Krishnamurthy et al. to discover candidates [20]. There have also been in-depth studies of high-level discriminatory information extraction employing deep networks to improve FP minimization. Setio et al. employed a fusion technique to conduct FP reduction after training 9 independent 2D convolutional neural networks on 9 different perspectives of candidates [21]. For candidate detection, another study used a modified version of Faster R-CNN, which was the state-of-the-art object detector at the time, and a patch-based 3D CNN for the FP reduction step [22]. All of these approaches, however, are computationally ineffective (e.g., exhaustive deploy of sliding windows over feature maps).

Thus, the contributions of the research work are as follows:

- In this study, an Otsu Thresholding based segmentation process is introduced for detecting the lung nodules.
- An optimized CNN is used for classification, where the activation functions and convolutional layers are fine-tuned with the aid of a new IMFO algorithm.
- To enhance the global exploration capability of classic MFO algorithm, an improved MFO (IMFO) algorithm is proposed.

The structure of the article is given as below: The existing schemes on lung cancer are reviewed in Sect. 2. Section 3 depicts a development of novel lung nodule detection model and it discusses the segmentation and feature extraction stages. Section 4 briefs the optimal CNN-oriented classification for precise lung nodule detection and it clearly portrays the process of proposed IMFO algorithm. In addition, Sects. 5 and 6 discusses the results as well as conclusion.

### 2 Literature Review

In the existing works, an image processing approach to identify solid nodules, part solid, as well as GGO within CT of chest was established in [23]. The major procedures are, "lung segmentation, candidate detection, nodule enhancement, reduction of false positives and image pre-processing". Furthermore, deep learning approach based lung nodule identification was presented in [24] and in this the characteristics were retrieved via utilizing patch-based multi-resolution convolutional networks as well as 4 fusion approaches are employed for classification. Gu et al. [25] designed a unique CAD structure via employing 3D DCNN combined with multi-scale prediction scenario. Additionally, for the identification of lung nodule in high & low resolution CT pictures, a DeepLN scheme was provided in [26].

A unique categorization model for lung carcinomas is introduce in [27] and a probabilistic neural network was deployed as a classifier for discriminating the true nodules. In [28], a unique CAD system with CNN was utilized in CT image segmentation scheme. Moreover, a 4-channel CNN model was introduced for identifying the lung nodules on the basis of multi-group patches of lung images [29]. A multi-resolution CNN based lung nodule patient's categorization is established in [30] as well as the scheme was splitted into 3 stages through the knowledge transfer usage.

Table 1 reviews lung nodule recognition systems. The major drawbacks of the existing lung nodule detection system are listed as follows: In SVM, choosing the appropriate kernel function is tricky [23], CNN in [24] requires on optimization concept for further efficiency, DCNN has to concern more on FP reduction [25], robustness needs to be enhanced with more attention [26], more memory space was needed [27], insufficient detection sensitivity [28], and higher computational cost [29]. To tackle these limitations, a new lung nodule detection approach is introduced. The detailed explanation of the process is provided in underneath passages.

### 3 Development of Novel Lung Nodule Detection Model

The framework of the lung nodule recognition scheme is specified in Fig. 1. It encompasses "(1) pre-processing, (2) segmentation, (3) feature extraction and (4) classification". At first, pre-processing is conducted. Then, segmentation is done utilizing Otsu Thresholding model. After that feature extraction performed via LBP model. The LBP features are then classified via optimized CNN model, wherein, activation function and convolutional layer counts are tuned via IMFO scheme.

#### 3.1 Pre-processing

At first, the input images are subjected to the pre-processing phase and this process includes four steps. The detailed explanations are given below:

(a) **Read image:** At first, storing the path to image dataset in a variable and then constructed a function to load folders containing images into arrays.

lable I The adv	antages and challenges of the e	xtant lung nodule detection models	
Citation	Developed model	Advantages	Challenges
[23]	SVM	Improved execution time	Choosing appropriate kernel function is tricky
		Good rate of detection	
[24]	CNN	Greater sensitivity	Needs the optimization idea in CNN for further efficiency
[25]	Deep CNN	Greater detection of tiny lung nodules	Future works is on the concern of FP reduction
		Better sensitivity	
[26]	DeepLN	Increase the effectiveness of training process	Robustness needs to be enhanced with more attention
		Better prediction performance	
[27]	PNN	Good rate of classification	Needs mode memory space
		Detection of low-contrast nodules is possible	
[28]	CNN	Diminishes the system complexity	Detection sensitivity is not satisfactory
[29]	CNN	Minimizes the false positives extremely	Need some expert knowledge
		Greater detection precision	
[30]	CNN	High accuracy	High computational cost

Input image



Fig. 1 Structural design of the adopted lung nodule detection scheme

- (b) **Resize image:** Some times, the images from the dataset are varying in size, so a base size for all images is established.
- (c) **Noise removal:** In this step, the unwanted noises are removed from the resized images. In this work, a Gaussian smoothing is utilized for noise removal opera-

tion. The outcome of blurring an image with a Gaussian function is Gaussian smoothing.

(d) **Morphological operation:** In this step, first segment the image to separate the background from the foreground objects, and then refine our segmentation with more noise removal.

#### 3.2 Otsu Thresholding Based Segmentation

Following pre-processing, the images are submitted to an Otsu thresholding [31] segmentation algorithm. As a result of this idea, Otsu sets the threshold value, which reduces overlapping. As a result, dark as well as light zones (*T*1,*T*0) are segmented. With *T*1 as well as *T*0, the background and object were integrated. The lowest value of pixel level has been measured after scanning all of the thresholds. Consider the probability of histogram with the noted gray value as  $hp(\tilde{i}), \tilde{i} = 1, ...., \tilde{l}$  as well as shown in Eq. (1).

$$hp(\overline{i}) = \frac{no\{(\overline{r},\overline{c}) | image(\overline{r},\overline{c}) = \overline{i}\}}{\overline{R}.\overline{C}}$$
(1)

Here, "index image for row and column is given by  $\overline{r}$  and  $\overline{c}$ , in that order, and the count of rows and columns of the image is expressed by  $\overline{R}$  and  $\overline{C}$ , respectively.

The weight, variance, and mean of class T0 with intensity value lies amongst 0 to th is indicated as  $\omega_{\overline{b}}(th)$ ,  $\sigma_{\overline{b}}(th)$  and  $\mu_{\overline{b}}(th)$ , respectively. The weight, variance and mean of class T1 with intensity value fall among th + 1 to 1 and are explicated by  $\omega_{\overline{f}}(th)$ ,  $\sigma_{\overline{f}}(th)$  and  $\mu_{\overline{f}}(th)$ , correspondingly. Weighed sum of cluster variances is referred to  $\sigma_{\omega}^{2^{\prime\prime}}$ .

The values with the lowest class variances are referred to as the best threshold value th \*. The class variance is depicted in Eq. (2).

$$\sigma_{\omega}^2 = \omega_{\overline{b}}(th) * \sigma_{\overline{b}}^2(th) + \omega_{\overline{f}}(th) * \sigma_{\overline{f}}^2(th)$$
(2)

$$\omega_{\overline{b}}(th) = \sum_{\overline{i}=1}^{th} hp(\overline{i}) \tag{3}$$

$$\omega_{\bar{f}}(th) = \sum_{\bar{i}=th+1}^{\bar{l}} hp(\bar{i}) \tag{4}$$

$$\mu_{\overline{b}}(th) = \frac{\sum_{\overline{i}=1}^{th} \overline{i} * hp(\overline{i})}{\omega_{\overline{b}}(th)}$$
(5)

$$\mu_{\bar{f}}(th) = \frac{\sum_{\bar{i}=th+1}^{\bar{i}} \bar{i} * hp(\bar{i})}{\omega_{\bar{f}}(th)}$$
(6)

$$\sigma_{\overline{b}}^2(th) = \frac{\sum_{\overline{i}=1}^{th} (\overline{i} - \mu_{\overline{b}}(th))^2 * hp(\overline{i})}{\omega_{\overline{b}}(th)}$$
(7)

$$\sigma_{\bar{f}}^{2}(th) = \frac{\sum_{\bar{i}=th+1}^{\bar{l}} (\bar{i} - \mu_{\bar{f}}(th))^{2} * hp(\bar{i})}{\omega_{\bar{f}}(th)}$$
(8)

#### 3.3 LBP Based Feature Extraction

*LBP* [32] LBP is used to extract typical characteristics, which are constantly encoded via "binomial concatenation" between the reference pixel  $K_c$  and its neighborhoods with radius  $\hat{R}$  as shown in Eq. (9). The neighbours of reference pixel are specified  $K_{G,\hat{R}}$  as well as the neighborhood index in a given local sub-region Q, which surrounds  $K_c$  is indicated as  $\hat{g}$ . The coding function  $s_1(', ')$  is calculated for discovering the labels. Equation (10) illustrates the threshold function, where the coding function output  $s_1(', ')$  denotes the binary gradient direction [32].

$$LBP_{G,\hat{R}}(K_{c}) = \left\{ s_{1}\begin{pmatrix} Q(K_{1,\hat{R}}), \\ Q(K_{c}) \end{pmatrix}, s_{2}\begin{pmatrix} Q(K_{2,\hat{R}}), \\ Q(K_{c}) \end{pmatrix}, \dots, s_{1}\begin{pmatrix} Q(K_{\hat{g},\hat{R}}), \\ Q(K_{c}) \end{pmatrix} \right\} \qquad \left| \hat{g} = 1, 2, \dots, G; \hat{R} = 1$$
(9)

$$s_1 \begin{pmatrix} (Q(K_{\hat{g},\hat{R}}), \\ Q(K_c) \end{pmatrix} = \begin{cases} 1 & if \left(1 - \left(K_{\hat{g},\hat{R}}\right) - Q\left(K_c\right)\right) \ge 0 \\ 0 & else \end{cases}$$
(10)

The features derived via LBP are signified as *fe* that is classified via CNN model.

## 4 Optimized CNN-oriented Classification for Precise Lung Nodule Detection

### 4.1 Optimized CNN for Classification

"CNN are a biologically inspired trainable architecture that can learn invariant features". In CNN [33], the input feature *fe* is modelled as revealed in Eq. (11), with a dimension of  $m_1 \times m_2$ 

$$\operatorname{Im}_{Se}: \{1, \dots, m_1\} \times \{1, \dots, m_2\} \to A \subseteq \Re, (i, j) \mapsto fe_{i, j}$$

$$\tag{11}$$

Consider filter  $L \in \Re^{2g_1+1 \times 2g_2+1}$  and "discrete convolution with filter" *L* is represented via Eq. (12), in which *L* is modelled as in Eq. (13).

$$\left(fe_{i} * L\right)_{p,r} := \sum_{\nu=-g_{1}}^{g_{1}} \sum_{u=-g_{2}}^{g_{2}} L_{\nu,u} fe_{p+\nu,r+u}$$
(12)

$$L = \begin{pmatrix} L_{-g_1, -g_2} & \dots & L_{-g_1, -g_2} \\ \vdots & L_{0,0} & \vdots \\ L_{g_1, -g_2} & \dots & L_{g_1, g_2} \end{pmatrix}$$
(13)

The "discrete Gaussian filter  $L_{H(\sigma)}$ " is exposed as per Eq. (14), in which,  $\sigma$  signifies "standard deviation of Gaussian distribution".

$$\left(L_{H(\sigma)}\right)_{p,r} = \left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{p^2+r^2}{2\sigma^2}\right)\right]$$
(14)

Consider the "layer *s* be assumed as a convolutional layer that includes  $n_1^{(s)}$  feature maps with  $n_2^{(s)} \times n_3^{(s)}$  dimension at its output. Equation (15) demonstrate  $i^{th}$  feature map in *s* layer and  $W_i^{(s)}$  points out bias matrix and  $L_{ij}^{(s)}$  indicates filter size of  $2g_1^{(s)} + 1 \times 2g_2^{(s)} + 1$  that links the  $i^{th}$  feature map in *s* layer with  $j^{th}$  feature map in the layer (s-1)".

$$X_{i}^{(s)} = W_{i}^{(s)} + \sum_{j=1}^{n_{1}^{s-1}} L_{i,j}^{(s)} * X_{j}^{(s-1)}$$
(15)

Accordingly, the output feature map has a specific size as revealed as per Eq. (16).

$$n_2^{(s-1)} - 2g_1^{(s)} = n_2^{(s)} \& n_3^{(s-1)} - 2g_2^{(s)} = n_3^{(s)}$$
(16)

The convolutional layer function is evaluated as per Eqs. (17) and (18). Subsequently, Eq. (18) shows output, accomplished via the calculation of unit at position (p, r). Here,  $L_{i,j}^{(s)}$  indicates trainable weights and  $W_i^{(s)}$  indicates bias matrix.

$$\left(X_{i}^{(s)}\right)_{p,r} = \left(W_{i}^{(s)}\right)_{p,r} + \sum_{j=1}^{n_{1}(s-1)} \left(L_{i,j}^{(s)} * X_{j}^{(s-1)}\right)_{p,r}$$
(17)

$$\left(W_{i}^{(s)}\right)_{p,r} + \sum_{j=1}^{n_{1}(s-1)} \sum_{\nu=-g_{1}^{s}}^{g_{1}^{s}} \sum_{u=-g_{2}^{s}}^{g_{2}^{s}} \left(L_{i,j}^{(s)}\right)_{\nu,u} \left(L_{i,j}^{(s)}\right)_{p,r} \left(X_{j}^{(s-1)}\right)_{p+\nu,r+u}$$
(18)

Equation (19) is formed if the layer s - 1 is not regarded as fully connected, where,  $We_{i,j,p,r}^s$  denotes the "weight that associates the unit at the position (g, h) in the *j*<sup>th</sup> feature map of layer s - 1 and the *i*<sup>th</sup> unit in *s*".



Fig. 2 General architecture of CNN model

$$x_{i}^{(s)} = \left[f(v_{i}^{s}) \text{ with } v_{i}^{s}\right] = \left[\sum_{j=1}^{n_{1}(s-1)} \sum_{p=-1}^{n_{2}(s-1)} \sum_{r=1}^{n_{3}(s-1)} We_{i,j,p,r}^{s} \left(X_{j}^{(s-1)}\right)_{p,r}\right]$$
(19)

The general diagram of the optimized CNN model is depicted in Fig. 2.

#### 4.2 Solution Encoding

This study focuses on the exact detection of lung nodules, for which optimization logic is used. In particular, the activation function  $\sigma$  is fine-tuned for choosing proper functions such as sigmoid, softmax, linear, relu, or tanh. Along with, the number of convolutional layers indicated as *L* is optimally described using a new IMFO approach. Figure 3 shows the input to the suggested method. Further, the objective of the research is given in Eq. (20), in which *acc* indicates accuracy.

$$obj = \max(acc)$$
 (20)



Fig. 3 Solution encoding

Deringer

#### 4.3 Proposed IMFO Algorithm

The extant MFO comprise a variety of advantage; it meets with drawbacks and to overcome those drawbacks, certain enhancements were made in extant MFO. Usually, self- enhancement is established to be capable in conventional optimization schemes [34–38]. MFO [39] include three phases.

#### 4.3.1 Generation of Initial Moth Populace

Every moths flies in hyper-dimensional spaces as identified as per Eq. (21), wherein, the moth count is indicated by n and d denotes the dimension count of solution spaces.

$$R = \begin{bmatrix} r_{1,1} & r_{1,2} & \dots & r_{1,d} \\ r_{2,1} & r_{2,2} & & r_{2,d} \\ \dots & & & \\ r_{n,1} & r_{n,2} & \dots & r_{n,d} \end{bmatrix}$$
(21)

Equation (22) illustrates the fitness values of moths.

$$OR = \begin{bmatrix} OR_1 \\ OR_2 \\ \dots \\ OR_n \end{bmatrix}$$
(22)

Equations (23) and (24) demonstrate the flame in D-dimensional spaces.

$$H = \begin{bmatrix} H_{1,1} & H_{1,2} & \dots & H_{1,d} \\ H_{2,1} & H_{2,2} & & H_{2,d} \\ \dots & & & \\ H_{n,1} & H_{n,2} & \dots & H_{n,d} \end{bmatrix}$$
(23)

$$OH = \begin{bmatrix} OH_1 \\ OH_2 \\ \dots \\ OH_n \end{bmatrix}$$
(24)

#### 4.3.2 Moth's Position Update

The MFO algorithm is a three-tuple that computes the global optima of optimization difficulties, as shown as per Eq. (25), wherein, " $I : \varphi \to \{R, OR\}$  specifies initial random location, the motion of moth in search space is depicted as  $\ddot{P} : R \to R$  and the concluding search process is stated as  $T : R \to \{true, false\}$ ".

$$MFO = (I, \ddot{P}, T) \tag{25}$$

The arbitrary distribution is represented by Eq. (26), wherein, *ub* and *lb* specifies upper and lower bounds.

$$R(i,j) = [(ub(i) - lb(j)) * rand() + lb(i)]$$
(26)

As per IMFO algorithm, a random number 'r' is generated and a threshold value ( $\delta$ ) of 0.5 is fixed. If r is less than the threshold  $\delta$ , (i.e. if  $r < \delta$ ), then the logarithmic spiral update takes place as given by Eq. (27), where  $D_i$  is computed as mean distance of  $i^{th}$  moth with  $j^{th}$  flame (i.e. $D_i = |H_j - R_i|$ ), b symbolize logarithmic spiral shape and t lies between [-1, 1].

$$S(R_i, H_j) = D_i e^{bt} \cos(2\pi t) + H_j$$
<sup>(27)</sup>

On the other hand, if  $r > \delta$ , a new logarithmic spiral with new distance evaluation takes place by Eq. (28). In Eq. (28), the distance parameter  $D_i$  is determined as per Eq. (29), where C = 2.a, *a* represents the random number, which lies among [0, 1] as well as  $H_j^*$  denotes the current position of *j*th moth and  $R_i$  denotes current position of *i*th flame.

$$S(R_i, H_i) = D_i e^{bt} \cos(2\pi t) \tag{28}$$

$$D_i = \left| H_j^* - R_i \right| \times C \tag{29}$$

#### 4.3.3 Update the Flames Count

Here, the exploitation stage has been is improved via lessening the count of flames as in Eq. (30), wherein, G depicted maximal flame count, l and  $G_I$  delineates current and maximum iteration counts.

$$flame count = round \left(G - l * \frac{G - l}{G_I}\right)$$
(30)

The flowchart of the adopted IMFO algorithm is given in Fig. 4 and the pseudocode is provided in algorithm 1.



Fig. 4 Flowchart of the proposed IMFO algorithm

Algorithm 1: IMFO scheme					
"Initiate the Moth-flame parameter					
Initiate position of Moth $R_i$ arbitrarily					
for $i = 1$ to $n$ do					
Estimate the fitness function $H_i$					
end for					
while $iter \leq \max_{iter}$					
Generate a random number 'r' and fix a threshold value ( $\delta$ ) of 0.5					
If $r < \delta$					
Update the logarithmic spiral using Eq. (27)					
else					
Update the logarithmic spiral with a new distance					
evaluation as per Eq. (28)					
Update the $R_i$ position					
Estimate the flame count as per Eq. (30)					
Estimate the fitness function $H_i$					
if iter=1					
H = sort(R)andOH = sort(OR)					
else					
$H = sort(R_{t-1}, R_t) and OH = sort(R_{t-1}, R_t)$					
end if					
for $i = 1$ to $n$ do					
for $j = 1$ to $d$ do					
Update the value of b and t					
Update $R(i, j)$ as per Eq. (26)					
end for					
end while					
Return the best solution"					

# **5** Results and Discussion

### 5.1 Simulation Settings

The presented lung nodule detection system employing the CNN+IMFO+LBP model was conducted in **Python**. Here, the performances of offered method was evaluated over SVM [1], KNN [26], CNN [2], MFO [32], WTEEB [41] and GWO+FRVM [42] models. Moreover, the performance analysis was performed

with respect to "accuracy, sensitivity, specificity, precision, FPR, FNR, FMS, NPV, MCC, and FDR respectively by varying the training percentages from 40, 50, 60, 70, 80 and 90". Based upon the confusion matrix, the performance metrics are defined and it is given as follows:

(a) Accuracy "It is the closeness of the measured value to a standard true value" and it is expressed as per Eq. (31).

$$Accuracy = \frac{True \ negative + \ True \ positive}{True \ positive + \ False \ positive + \ True \ negative + \ False \ negative}$$
(31)

(b) *Sensitivity* "Sensitivity is the proportion of true positives that are correctly predicted by the model" and it is evaluated as per Eq. (32).

Sensitivity = 
$$\frac{\text{True positive}}{\text{True positive} + \text{False negative}}$$
 (32)

(c) *Specificity* "specificity is the proportion of true negatives that are correctly predicted by the model" and it is expressed as Eq. (33).

Specificity = 
$$\frac{\text{True negative}}{\text{False positive} + \text{True negative}}$$
 (33)

(d) *Precision* "It is the fraction of relevant instances among the retrieved instances" and it is expressed as per Eq. (34).

$$Precision = \frac{True \text{ positive}}{True \text{ positive} + \text{ False positive}}$$
(34)

(e) *Recall* "It is the fraction of relevant instances that were retrieved" and it is expressed as per Eq. (35).

$$Recall = \frac{True \ positive}{True \ positive + \ False \ negative}$$
(35)

(f) *FMS* "It is the weighted harmonic mean of the precision and recall of the test" and it is expressed as follows.

$$FMS = \frac{2 \times Precision \times Recall}{Precision + Recall}$$
(36)

(g) *MCC* "MCC is used to measure the difference between the predicted values and actual values" and it is given as follows:

$$MCC = \frac{\text{True positive} \times \text{True negative} - \text{False positive} \times \text{False negative}}{\sqrt{(\text{True positive} + \text{False positive})(\text{True positive} + \text{False negative})}}$$
(2)

(37)

(h) *NPV* "The NPV is the probability that following a negative test result, that individual will truly not have that specific disease" and it is expressed as below:

$$NPV = \frac{True negative}{True negative + False negative}$$
(38)

(i) *FPR* "A false positive is an outcome where the model incorrectly predicts the positive class" and it is given in Eq. (39).

$$FPR = \frac{False \text{ positive}}{True \text{ negative} + False \text{ positive}}$$
(39)

(j) *FNR* "A false negative is an outcome where the model incorrectly predicts the negative class" and it is given in Eq. (40).

$$FNR = \frac{False negative}{False negative + True positive}$$
(40)

(k) *FDR* "It is a measure of accuracy when multiple hypotheses are being tested at once, for example when multiple metrics, or variations, are being measured in a single experiment" and it is expressed as per Eq. (41).

$$FDR = \frac{False \text{ positive}}{True \text{ positive} + False \text{ positive}}$$
(41)

For clear understanding,

*True positive* "The image results are those that detecting a disease when there is disease".

*True negative* "The image results are those in which no nodule is detected as well as no disease is diagnosed".

*False positive* "The images results are those that detect a disease even when it is not present".

*False negative* "The image results are those that detect a disease when it does not already exist".

### 5.2 Dataset Description

The dataset used in this study has been downloaded from [40]. The lung computed tomography (CT) images in the Lung Image Database Consortium image collection (LIDC-IDRI) are marked-up annotated lesions from diagnostic and lung cancer screening CT scans. It is a worldwide resource that is available online for the creation, instruction, and assessment of computer-assisted diagnostic (CAD) techniques for the early identification and diagnosis of lung cancer. This public–private partnership, which was started by the National Cancer Institute (NCI), advanced by the Foundation for the National Institutes of Health (FNIH), and supported by the Food and Drug Administration (FDA) through active participation, shows the effective-ness of a consortium built on a consensus-based method.

This data set, which includes 1018 instances, was produced in collaboration with seven academic institutions and eight medical imaging businesses. Images from a clinical lung CT scan are included for each individual, together with an attached XML file that contains the findings of a two-phase image annotation process carried out by four qualified thoracic radiologists. Each radiologist individually assessed each CT image during the initial blinded-read phase and labelled lesions as "nod-ule > or = 3 mm," "nodule 3 mm," or "non-nodule > or = 3 mm." Each radiologist independently assessed their marks and the anonymized marks of the three other radiologists to form a conclusion during the subsequent unblinded-read phase. This procedure was designed to identify every lung lesion in every CT scan as completely as feasible without resorting to forced consensus.

### 5.3 Performance Analysis: Varying Learning Percentage

Figure 5 demonstrates the performance of the CNN+IMFO+LBP model over the traditional approaches via different TP. The suggested model achieves better value than the extant ones. Due to the shortcomings of the state-of-art methods (for instance, the training time of SVM is too high, KNN is suffered from memory intensive and slow estimations, and overfitting issue in CNN), the performance of the related works are low. Compared to CNN [24], the adopted CNN model utilizes the IMFO algorithm and it provides better value than the other.

From Fig. 5a, the accuracy of the suggested CNN+IMFO+LBP model at 90<sup>th</sup> TP is 3.42%, 2.21%, 1.46%, 0.73%, 1.84%, and 3.34% better than SVM, KNN, CNN, MFO, WTEEB and GWO+FRVM models. On concerning Fig. 5c, the suggested approach specificity is 3.04%, 2.26%, 1.49%, 0.74%, 1.84% as well as 13.87%, superior to SVM, KNN, CNN, MFO, WTEEB as well as GWO+FRVM models at 40<sup>th</sup> TP. The recall of the suggested CNN+IMFO+LBP approach at 90<sup>th</sup> TP is 4.04%, 2.21%, 2.07%, 0.73%, 1.84%, & 3.33% superior to SVM, KNN, CNN, MFO, WTEEB & GWO+FRVM approaches. On analysing the MCC, the adopted CNN+IMFO+LBP scheme is 45.12%, 20.31%, 3.75%, 0.72%, 1.85%, and 20.98% better than SVM, KNN, CNN, MFO, WTEEB and GWO+FRVM models at 80<sup>th</sup> TP. Also, on examining the FDR of suggested method at 90<sup>th</sup> TP is 54.18%, 46.99%, 37.15%, 22.81%, 41.81%, and 50.79% better than SVM, KNN, CNN, MFO, WTEEB and GWO+FRVM models Thus, the supremacy of CNN+IMFO+LBP method was confirmed from outcomes.

### 5.4 Statistical Analysis

The statistical analysis of the suggested CNN + IMFO + LBP model is given by Table 2 in terms of accuracy. Hence, the statistical measure of the adopted model is should be higher than the existing works. The proposed method with respect to best performance is 7.35%, 2.99%, 1.78%, 0.73%, 1.85%, and 3.34% better than SVM, KNN, CNN, MFO, WTEEB and GWO+FRVM models. The mean performances of implemented model for accuracy is 15.79%, 6.82%, 1.73%, 0.74%, 1.85%, and 5.73% better than SVM, KNN, CNN, MFO, WTEEB and GWO+FRVM

Fig. 5 Performance analysis of the proposed and conventional models on concerning performance meas- $\mathbf{b}$  ures such as a accuracy b sensitivity c specificity d Precision e Recall f NPV g FMS h MCC i FPR j FNR and k FDR

approaches. Additionally, the median of the proposed method is 9.81%, 3.01%, 1.73%, 0.74%, 1.71%, and 5.72% better than SVM, KNN, CNN, MFO, WTEEB and GWO+FRVM models. Hence, the betterment of suggested model is established from the results.

### 5.5 Overall Performance Analysis

The overall performance analysis of the suggested as well as extant methods in terms of various measures is discussed in this section. From Table 3, the suggested scheme for accuracy is 7.35%, 2.99%, 1.78%, 0.73%, 1.84% and 4.05% better than SVM, KNN, CNN, MFO, WTEEB and GWO+FRVM models. On considering sensitivity, the adopted CNN+IMFO+LBP model is 10.96%, 2.99%, 1.78%, 0.73%, 2.26%, and 4.46% better than SVM, KNN, CNN, MFO, WTEEB and GWO+FRVM approaches. On considering recall, the adopted CNN+IMFO+LBP model is 10.96%, 2.99%, 1.78%, 0.73%, 1.84%, and 3.35% better than SVM, KNN, CNN, MFO, WTEEB and GWO+FRVM approaches. On considering the negative measures (FPR, FDR, FNR), the adopted model acquires less value than the existing ones. As a result, the total analysis shows that the CNN+IMFO+LBP design is superior.

### 5.6 Computational Time Analysis

The computation time of the adopted & extant works is determined as well as the values are tabulated in Table 4. The proposed CNN + IMFO + +LBP model achieves small execution time that is 134.5963254 s and it is 22.66%, 1.77%, 10.26%, 7.85%, 19.05%, and 5.48% better than the existing SVM, KNN, CNN, MFO, WTEEB, and GWOFRVM approaches respectively. Therefore, this analysis proves the superiority of the adopted model.

# 6 Conclusion

In this study, the pre-processed images were segmented utilizing the Otsu Thresholding approach, which resulted in a new lung cancer detection method. Further, the LBP characteristics were extracted from the segmented images. As the next process, classification takes place for which the CNN model was exploited. In addition, an enhanced version of the traditional MFO algorithm named as IMFO model was presented for optimizing the activation function and count of convolutional layers. Finally, an examination was done to evaluate the performance of the offered scheme. The adopted approach in terms of accuracy is 7.35%, 2.99%, 1.78%, 0.73%, 1.85%, and 3.34% better than SVM, KNN, CNN,







Fig. 5 (continued)

Metrics	SVM [23]	KNN [32]	CNN [24]	MFO [32]	WTEEB [41]	GWO+FRVM [42]	CNN + IMFO + LBP
Best	90.28	91.36	92.03	92.70	91.65	90.26	93.3768
Worst	62.78	68.73	89.38	90.72	89.71	77.68	91.3968
Mean	79.79	86.49	90.82	91.71	90.68	87.10	92.39
Median	84.14	89.69	90.90	91.71	90.68	88.64	92.39

Table 2 Statistical analysis: Proposed vs. traditional approaches

MFO, WTEEB and GWO+FRVM models. On analyzing the statistical analysis, the mean performances of implemented model for accuracy was 15.79%, 6.82%, 1.73%, 0.74%, 1.85%, and 5.73% better than SVM, KNN, CNN, MFO, WTEEB and GWO+FRVM approaches. Thus, the betterment of the suggested CNN+IMFO+LBP model was proven from outcomes. However, the computation time of the mode is not too much better. In the future, we want to comprehend the features retrieved by the networks for classification by deploying various visualization tools and see if they are consistent with features utilized by radiologists for diagnosis and also utilize the proposed system for other pulmonary diseases.

Measures	SVM [23]	KNN [32]	CNN [24]	MFO [32]	WTEEB [41]	GWO+FRVM [42]	CNN + IMFO + LBP
Accuracy	86.25	89.89	90.96	91.91	90.88	88.83	92.58
Sensitivity	83.44	89.89	90.96	91.91	90.49	88.45	92.58
Specificity	89.89	90.57	91.24	91.91	91.27	89.10	92.58
Recall	83.44	89.89	90.96	91.91	90.88	89.48	92.58
FPR	16.51	14.41	12.31	10.21	13.20	15.41	8.11
FNR	26.51	14.41	12.31	10.21	13.69	15.89	8.11
FMS	87.04	89.89	91.13	91.91	90.88	87.68	92.58
NPV	79.87	89.89	90.96	91.91	90.09	87.68	92.58
MCC	50	64.44	88.70	91.91	89.71	64.07	92.58
FDR	16.51	14.41	12.31	10.21	13.20	15.41	8.11

 Table 3
 Overall performances of suggested scheme over conventional schemes

 Table 4
 Computational time analysis

Methods	Computation time (Sec)
SVM[23]	174.025777454
KNN[32]	137.025478151
CNN[24]	149.985832543
MFO[32]	146.058992414
WTEEB [41]	166.265926267
GWO+FRVM [42]	142.39977763
CNN+IMFO+LBP	134.5963254

**Data Availability** The data that support the findings of this study are openly available in https://wiki. cancerimagingarchive.net/display/Public/LIDC-IDRI.

### References

- Imad, A., Malik, N. A., Hamida, B. A., Seng, G. H. H., & Khan, S. (2022). Acoustic photometry of biomedical parameters for association with diabetes and Covid-19. *Emerging Science Journal*, 6, 42–56. https://doi.org/10.28991/esj-2022-SPER-04
- Toes-Zoutendijk, E., Kooyker, A. I., Dekker, E., Spaander, M. C. W. (2019). Incidence of interval colorectal cancer after negative results from first-round fecal immunochemical screening tests, by cutoff value and participant sex and age, *Clinical Gastroenterology and Hepatology*, (In press, uncorrected proof, Available online 20 August 2019) https://doi.org/10.1016/j.cgh.2019.08.021
- Rombouts, A. J. M., Hugen, N., Elferink, M. A. G., Poortmans, P. M. P., & de Wilt, J. H. W. (2020). Increased risk for second primary rectal cancer after pelvic radiation therapy. *European Journal of Cancer, 124*, 142–151. https://doi.org/10.1016/j.ejca.2019.10.022
- Yin, Y., Sedlaczek, O., Müller, B., Warth, A., González-Vallinas, M., Lahrmann, B., Grabe, N., Kauczor, H. U., Breuhahn, K., Vignon-Clementel, I. E., & Drasdo, D. (2018). Tumor cell load and heterogeneity estimation from diffusion-weighted MRI calibrated with histological data: An example from lung cancer. *IEEE Transactions on Medical Imaging*, 37(1), 35–46. https://doi.org/10. 1109/TMI.2017.2698525
- Jiang, J., Hu, Y. C., Liu, C. J., Halpenny, D., Hellmann, M. D., Deasy, J. O., Mageras, G., & Veeraraghavan, H. (2019). Multiple resolution residually connected feature streams for automatic lung

tumor segmentation from CT images. *IEEE Transactions on Medical Imaging*, 38(1), 134–144. https://doi.org/10.1109/TMI.2018.2857800

- Petousis, P., Winter, A., Speier, W., Aberle, D. R., Hsu, W., & Bui, A. A. T. (2019). Using sequential decision making to improve lung cancer screening performance. *IEEE Access*, 7, 119403–119419. https://doi.org/10.1109/ACCESS.2019.2935763
- Alahmari, S. S., Cherezov, D., Goldgof, D. B., Hall, L. O., Gillies, R. J., & Schabath, M. B. (2018). Delta radiomics improves pulmonary nodule malignancy prediction in lung cancer screening. *IEEE Access*, 6, 77796–77806. https://doi.org/10.1109/ACCESS.2018.2884126
- Tremblay, A., Taghizadeh, N., MacGregor, J.-H., Armstrong, G., & Burrowes, P. (2019). Application of lung-screening reporting and data system versus pan-canadian early detection of lung cancer nodule risk calculation in the alberta lung cancer screening study. *Journal of the American College* of Radiology, 16(10), 1425–1432. https://doi.org/10.1016/j.jacr.2019.03.006
- Xie, Y., Zhang, J., & Xia, Y. (2019). Semi-supervised adversarial model for benign-malignant lung nodule classification on chest CT. *Medical Image Analysis*, 57, 237–248. https://doi.org/10.1016/j. media.2019.07.004
- Liu, H., Cao, H., Song, E., Ma, G., & Hung, C.-C. (2019). A cascaded dual-pathway residual network for lung nodule segmentation in CT images. *Physica Medica*, 63, 112–121. https://doi.org/10. 1016/j.ejmp.2019.06.003
- Venugopal, V. K., Vaidhya, K., Murugavel, M., Chunduru, A., Mahajan, H. (2019). Unboxing AI

   radiological insights into a deep neural network for lung nodule characterization, *Academic Radiology*, (In press, corrected proof, Available online 14 October 2019) https://doi.org/10.1016/j.acra. 2019.09.015
- Nomura, Y., Higaki, T., Fujita, M., Miki, S., & Awai, K. (2017). Effects of iterative reconstruction algorithms on computer-assisted detection (CAD) software for lung nodules in ultra-low-dose CT for lung cancer screening. *Academic Radiology*, 24(2), 124–130. https://doi.org/10.1016/j.acra. 2016.09.023
- Gong, J., Liu, J.-Y., Wang, L.-J., Zheng, B., & Nie, S.-D. (2016). Computer-aided detection of pulmonary nodules using dynamic self-adaptive template matching and a FLDA classifier. *Physica Medica*, 32(12), 1502–1509. https://doi.org/10.1016/j.ejmp.2016.11.001
- Bonavita, I., Rafael-Palou, X., Ceresa, M., Piella, G., & González Ballester, M. A. (2019). Integration of convolutional neural networks for pulmonary nodule malignancy assessment in a lung cancer classification pipeline. *Computer Methods and Programs in Biomedicine*. https://doi.org/10.1016/j. cmpb.2019.105172
- Tajbakhsh, N., & Suzuki, K. (2017). Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs. *Pattern Recognition*, 63, 476–486. https://doi.org/10.1016/j.patcog.2016.09.029
- Setio, A. A. A., Traverso, A., de Bel, T., Berens, M. S. N., & Jacobs, C. (2017). Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. *Medical Image Analysis*, 42, 1–13. https://doi.org/10. 1016/j.media.2017.06.015
- Toğaçar, M., Ergen, B. & Cömert, Z. (2019). Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks, *Biocybernetics and Biomedical Engineering*, (In press, uncorrected proof). https://doi.org/ 10.1016/j.bbe.2019.11.004
- Lopez Torres, E., Fiorina, E., Pennazio, F., Peroni, C., Saletta, M., Camarlinghi, N., Fantacci, M., & Cerello, P. (2015). Large scale validation of the m51 lung cad on heterogeneous ct datasets. *Medical Physics*, 42(4), 1477–1489. https://doi.org/10.1118/1.4907970
- 19. Ypsilantis, P.P. & Montana, G. (2016) Recurrent convolutional networks for pulmonary nodule detection in ct imaging. arXiv preprint arXiv: 1609.09143
- Krishnamurthy, S., Narasimhan, G., & Rengasamy, U. (2016). An automatic computerized model for cancerous lung nodule detection from computed tomography images with reduced false positives. In *International conference on recent trends in image processing and pattern recognition*. (pp. 343–355). Springer. doi: https://doi.org/10.1007/978-981-10-4859-3\_31
- Setio, A. A. A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., van Riel, S. J., Wille, M. M. W., Naqibullah, M., Sanchez, C. I., & van Ginneken, B. (2016). Pulmonary nodule detection in ct images: false positive reduction using multi-view convolutional networks. *IEEE Transactions on Medical Imaging*, 35(5), 1160–1169. https://doi.org/10.1109/TMI.2016.2536809

- Ding, J., Li, A., Hu, Z., & Wang, L. (2017) Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks. In *International conference on medical image computing and computer-assisted intervention*. (pp. 559–567). Springer. doi: https://doi. org/10.1007/978-3-319-66179-7\_64
- Kuo, C.-F.J., Huang, C.-C., Siao, J.-J., Hsieh, C.-W., & Hsu, H.-H. (2020). Automatic lung nodule detection system using image processing techniques in computed tomography. *Biomedical Signal Processing and Control*. https://doi.org/10.1016/j.bspc.2019.101659
- Li, X., Shen, L., Xie, X., Huang, S., Yu, J. (2019). Multi-resolution convolutional networks for chest X-ray radiograph based lung nodule detection, *Artificial Intelligence in Medicine*, (In press, corrected proof, Available online 28 October 2019). https://doi.org/10.1016/j.artmed.2019.101744
- Gu, Y., Lu, X., Yang, L., Zhang, B., & Zhou, T. (2018). Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs. *Computers in Biology and Medicine*, 103, 220–231. https://doi.org/10.1016/j.compbiomed.2018.10. 011
- Xu, X., Wang, C., Guo, J., Yang, L., & Yi, Z. (2019). DeepLN: A framework for automatic lung nodule detection using multi-resolution CT screening images, *Knowledge-Based Systems*, (In press, corrected proof, Available online 17 October 2019). https://doi.org/10.1016/j.knosys.2019.105128
- Woźniak, M., Połap, D., Capizzi, G., Lo Sciuto, G., & Frankiewicz, K. (2018). Small lung nodules detection based on local variance analysis and probabilistic neural network". *Computer Methods* and Programs in Biomedicine, 161, 173–180. https://doi.org/10.1016/j.cmpb.2018.04.025
- Wang, Q., Shen, F., Shen, L., Huang, J., & Sheng, W. (2019). Lung nodule detection in CT images using a raw patch-based convolutional neural network. *Journal of Digital Imaging*, 32(6), 971–979. https://doi.org/10.1007/s10278-019-00221-3
- Jiang, H., Ma, H., Qian, W., Gao, M., & Li, Y. (2018). An automatic detection system of lung nodule based on multigroup patch-based deep learning network. *IEEE Journal of Biomedical and Health Informatics*, 22(4), 1227–1237. https://doi.org/10.1109/JBHI.2017.2725903
- Zuo, W., Zhou, F., Li, Z., & Wang, L. (2019). Multi-resolution CNN and knowledge transfer for candidate classification in lung nodule detection. *IEEE Access*, 7, 32510–32521. https://doi.org/10. 1109/ACCESS.2019.2903587
- Gokulkumari, G. (2021). Metaheuristic-enabled artificial neural network framework for multimodal biometric recognition with local fusion visual features. *The Computer Journal*. https://doi.org/10. 1093/comjnl/bxab001
- Fan, K.-C., & Hung, T.-Y. (2014). A novel local pattern descriptor—local vector pattern in highorder derivative space for face recognition. *IEEE Transactions On Image Processing*, 23(7), 2877– 2889. https://doi.org/10.1109/TIP.2014.2321495
- LeCun, Y., Kavukvuoglu, K., Farabet, C. (2010) Convolutional networks and applications in vision, In *Circuits and systems, international symposium on*, (pp. 253–256). https://doi.org/10.1109/ISCAS. 2010.5537907
- Rajakumar, B. R. (2013). Static and adaptive mutation techniques for genetic algorithm: A systematic comparative analysis. *International Journal of Computational Science and Engineering*, 8(2), 180–193.
- Rajakumarm, B. R. (2013). Impact of static and adaptive mutation techniques on the performance of genetic algorithm. *International Journal of Hybrid Intelligent Systems*, 10(1), 11–22.
- Swamy, S. M., Rajakumar, B. R., & Valarmathi I. R. (2013) Design of hybrid wind and photovoltaic power system using opposition-based genetic algorithm with cauchy mutation, In: *IET chennai fourth international conference on sustainable energy and intelligent systems (SEISCON 2013)*, Chennai, India, doi: https://doi.org/10.1049/ic.2013.0361
- George, A. & Rajakumar, B. R. (2013) APOGA: an adaptive population pool size based genetic algorithm, In AASRI procedia - 2013 AASRI conference on intelligent systems and control (ISC 2013), (Vol. 4, pp. 288–296). doi: https://doi.org/10.1016/j.aasri.2013.10.043.
- Rajakumar, B. R. & George, A. (2012). A new adaptive mutation technique for genetic algorithm, In Proceedings of IEEE international conference on computational intelligence and computing research (ICCIC), (pp. 1–7), Coimbatore, India, doi: https://doi.org/10.1109/ICCIC.2012.6510293.
- Mirjalili, S. (2015). Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowledge-Based Systems*, 89, 228–249. https://doi.org/10.1016/j.knosys.2015.07.006
- 40. Dataset link: https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI

- Kumar, M. K., & Amalanathan, A. (2022). Automated lung nodule detection in CT images by optimized CNN: Impact of improved whale optimization algorithm. *Computer Assisted Methods in Engineering and Science.*, 29(1–2), 7–31. https://doi.org/10.24423/cames.372
- 42. Sathiya, T., Reenadevi, R., & Sathiyabhama, B. (2021). Lung nodule classification in CT images using Grey wolf optimization algorithm. *Annals of the Romanian Society for Cell Biology*, 25(6), 1495–1511.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.