Evolutionary Intelligence
https://doi.org/10.1007/s12065-020-00403-x

RESEARCH PAPER

# Automated mammogram breast cancer detection using the optimized combination of convolutional and recurrent neural network

**Rajeshwari S. Patil**[1] · **Nagashettappa Biradar**[2]

Received: 7 January 2020 / Revised: 20 February 2020 / Accepted: 31 March 2020


## Abstract

The objective of this study is to frame mammogram breast detection model using the optimized hybrid classifier. Image pre-processing, tumor segmentation, feature extraction, and detection are the functional phases of the proposed breast cancer detection. A median filter eliminates the noise of the input mammogram. Further, the optimized region growing segmentation is carried out for segmenting the tumor from the image and the optimized region growing depends on a hybrid meta-heuristic algorithm termed as firefly updated chicken based CSO (FC-CSO). To the next of tumor segmentation, feature extraction is done, which intends to extract the features like grey level co-occurrence matrix (GLCM), and gray level run-length matrix (GRLM). The two deep learning architectures termed as convolutional neural network (CNN), and recurrent neural network (RNN). Moreover, both GLCM and GLRM are considered as input to RNN, and the tumor segmented binary image is considered as input to CNN. The result of this study shows that the AND operation of two classifier output will tend to yield the overall diagnostic accuracy, which outperforms the conventional models.

**Keywords** Mammography · Breast cancer diagnosis · Optimized region growing · Deep hybrid learning · Firefly updated chick-based chicken swarm optimization

## Abbreviations

| | |
|---|---|
| CSO | Chicken swarm optimization |
| CAD | Computer-aided diagnosis |
| CNN | Convolutional neural network |
| FF | Firefly |
| FC-CSO | Firefly updated chicken-based CSO |
| FPR | False positive rate |
| GLRM | Gray-level run-length matrix |
| MCC | Mathews correlation coefficient |
| RNN | Recurrent neural network |
| MRI | Magnetic resonance imaging |
| ELM | Extreme learning machine |
| PDEs | Partial differential equations |
| SDEs | Simple differential equations |
| EINP | Estimation of intensity via non-parametric approach |
| IMBC | Incoherent motion in breast cancer |
| ADEWNN | Adaptive differential evolution wavelet neural network |
| MIAS | Mammographic image analysis society |
| MLP | Multi-layer perceptrons |
| PCET | Polar complex exponential transform |
| SVM | Support vector machines |
| IRMA | Image retrieval in medical applications |
| SURF | Speed-up robust features |
| DNN | Deep neural network |
| FNR | False-positive rate |
| MSVM | Multiclass support vector machine |
| LASSO | Least absolute shrinkage and selection operator |
| DL-CNN | Deep learning-convolution neural network |
| RBF | Radial basis functions |
| HGRE | High grey level run emphasis |
| ANN | Artificial neural network |
| LSTM | Long short-term memory |
| GRU | Gate recurrent unit |
| NPV | Negative predictive value |
| LGRE | Low grey level run emphasis |
| CSA | Crow search algorithm |
| FDR | False discovery rate |

✉ Rajeshwari S. Patil
rspatil1272014@gmail.com

[1] Department of Electronics and Communication, B.L.D.E.A, s V.P. Dr. P.G.Halakatti College of Engg. & Tech., (Affiliated toVisvesvaraya Technological University, Belagavi-590018), Vijayapur, Karnataka 586103, India

[2] Department of Electronics and Communication, Bheemanna Khandre Institute of Technology, Bhalki, Karnataka, India

Published online: 13 April 2020

| | |
|---|---|
| GLCM | Grey level co-occurrence matrix |
| IAP-CSA | Improved awareness probability-based chicken swarm optimization |
| WCO | World cup optimization |

## 1 Introduction

Breast cancer starts when cells begin to grow out of control and it usually form a tumor that can often seen on an x-ray or felt as a lump. Breast cancer occurs mostly in women. Breast cancer is very dangerous to women's health and life, and this cancer is ranked second and first in all-female diseases [1]. To decrease the death rate, lumps need to detect in an early stage [2]. Using the mammogram, the breast cancer has been detected as early as possible. Hence, the cost required for curing the breast cancer is less [3]. However, the accuracy is negatively influenced by several factors like the complexity of breast structure, subtle features of the early-stage disease, and radiologist fatigue and distraction [4, 5]. In the present scenario, the detection approaches implemented in recognition of breast cancer consisting of photoacoustic imaging, computed tomography, microwave imaging, mammography, MRI, (the Nomenclature table has been provided before the references section) etc., [5, 6]. Mammography is the best approach from all the methodologies to diagnose breast cancer [7]. In breast cancer, two symptoms, such as calcifications and breast masses, are considered as the early signs.

It is complicated for doctors to diagnose cancer based on mammogram images because of its complexity of premature breast cancer, combined with less brightness of mammogram images itself. Thus, it is essential for enhancing the detection efficiency of doctors via the CAD system of deep learning techniques [8]. A variety of CAD models have been introduced earlier for decreasing human involvement and assist the radiologist in automatic discrimination among malignant and benign tumors [9]. A CAD model is not only helpful in breast image and also benefit in other areas of medical diagnosis like respiratory-related chronic diseases [10], and diagnosis of pulmonary edema [11] for simple and early diagnosis. Moreover, the integration of expert's knowledge will enhance the diagnosis accuracy.

For segmenting the medical images, numerous approaches have been introduced. However, image segmentation is found to be a difficult job because each of these approaches involves their constraints. Level set [12] is an excellent approach applicable for segmenting the image, which was initially developed by Sethian and Osher [13]. This method segments the images precisely by considering the energy minimization problem. It might be prominent that region boundaries are very important in analyzing the mammogram images. The active contour is introduced as a zero level set [14]. The conventional level set techniques are categorized into three classes, region, edge, and combined approaches [15, 16]. However, these methods are generally very sensitive to weak boundaries and noise [17].

The key contributions of the paper are based on the following points.

- To develop an improved breast cancer detection model using mammogram images with the help of four phases like "pre-processing, tumor segmentation, feature extraction, and classification".
- To enhance the segmentation accuracy by developing the optimized region growing algorithm, in which the proposed FC-CSO algorithm optimizes the threshold or tolerance.
- To utilize the same FC-CSO algorithm to the hybrid classifier with the integration of CNN and RNN termed as CNN for optimizing the hidden layers, which in turn improves the detection accuracy.

The organization is designed as follows:

Literature review features and the challenges of state-of-the-art models are described in Sect. 2. Architectural view of a proposed breast cancer diagnosis is shown in Sect. 3. Section 4 depicts the optimized segmentation and classification for diagnosing breast cancer. The proposed and conventional models are shown in Sect. 5. Finally, the conclusion is determined in Sect. 6.

## 2 Literature review

### 2.1 Related works

In 2019, Wang et al. [18] have investigated the CAD approach based on feature fusion with the deep features of CNN. At first, a mass detection approach that was based on deep features of CNN and unsupervised ELM clustering. Next, a feature set was built by fusing morphological, texture, density, and deep features. To classify benign and malignant breast masses, an ELM classifier was introduced with the help of a fused feature set. Finally, the experiments have shown that the suggested model was outperformed n detection and classification. The CNN [18] has high classification and detection accuracy to extract the features. But it is computationally expensive.

In 2019, Geweid and Abdallah [19] utilized a non-parametric method for resolving the uncertainty in breast cancer detection. Moreover, it investigated the reliability of implementing the non-parametric method in mammography, to diagnose and provide treatment to breast cancer. The five approaches were tested using 156 mammogram images for acquiring accurate measurement of the efficiency of the detection procedure. Therefore, it was concluded that

the developed model was accurate in efficiently diagnosing the breast cancer. Non-parametric approach [19] is an accurate and reliable mechanism, and it demonstrates the affected region in an efficient borderline, and it is relatively not affected by mammogram noise. However, there are some conflicts such as it requires more computational time and needs more statistical power if the guess of an approximately equivalent parametric test is valid.

In 2016, Singh et al. [20] has introduced a CAD technique for decreasing the involvement of human and assisting the radiologist in diagnosing cancer using PCET moments. For feature extraction of a suspicious area, both phases of PCET moments and magnitude were utilized. For enhancing the classification accuracy of the developed CAD model, a new classifier ADEWNN was developed. For the experiment, the developed strategy was verified on mammographic images from the "MIAS database." Thus, results of the suggested model have obtained the best accuracy with less error rate. ADEWNN [20] has high accuracy, and it is more stable and efficient. Yet, it is not fully automatic because the region of interest is extracted manually.

In 2019, Li et al. [21] has introduced a modified DenseNet Neural Network approach and named the DenseNet-II Neural Network method for classifying cancerous mammography images accurately and effectively. Here, the "DenseNet-II Neural Network" approach was introduced. The simulation results shows that the DenseNet-II Neural Network approach has high classification performance when considered the tenfold cross-validation approach and also enhanced the strength and generalization. DenseNet Neural Network [21] enhanced the efficiency of the network model as well as calculation speed, and to merge different activations, this model can learn easily that leads to accurate classification. Still, it is highly expensive.

In 2018, Cruz et al. [22] have authenticated the implementation of classifiers on Neural Networks such as MLP, RBF, and SVM with various kernels for detecting the existence of breast lesions and categorizing those lesions into benign or malignant. Moreover, the IRMA database was utilized, which consists of 2796 patch images that include $128 \times 128$ pixels region of interest. Every image was indicated by texture patterns taken out from two-level decomposition using morphological wavelets. Also, multi-layer has two layers of successful technique in attaining high accuracy for breast cancer detection. MLP [22] has high accuracy when compared to existing models, and it is useful for regression and mapping. Though there are few defects like training huge dataset on vast problem consumes more time.

In 2019, Kaur et al. [23] have introduced a method on the mini-MIAS database concerning a pre-processing approach and create feature extraction with the help of K-mean clustering to SURF selection. Also, a layer was newly added in the classification level that gives a ratio of 70% training to 30% testing of DNN and MSVM. Therefore, the result was shown that the accuracy of the developed model by "K-mean clustering using MSVM" was outperformed than decision tree method. SVM [23] has high classification accuracy, and it works well when there is a clear margin of division among the classes. However, there are some challenges, such as it is not appropriate for huge databases, and if the count of features for every data point comes beyond the count of training samples, it doesn't work well.

In 2019, Hai et al. [24] have recommended an end-to-end learning method that relates to merge multi-level features. To extract high-level features, CNN was introduced. For optimizing the CNN, these extracted features were merged. Hence, the outcomes have demonstrated that the suggested model was producing the best results when compared to existing CNN approaches. S-DenseNet [24] has high performance and has the anti-over fitting capability. Though, it requires more amount of memory.

In 2017, Duraisamy and Emperumal [25] had developed a "deep learning-based framework" to categorize digital mammograms, and also verify existence of tumor tissues with level sets. The experimental results were performed on benchmark breast cancer database, which proved that the developed model had provided an improved performance than the conventional models. Fully complex-valued relaxation network [25] has high performance, and it doesn't require more effort while training. But, it is required to enhance classification accuracy.

In 2019, Razmjooy et al. [26] have analyzed the different evolutionary algorithms for optimizing neural networks in different ways for image segmentation purposes. In 2019, Guo et al. [27] have investigated a non-destructive test method called image processing for the detection of breast cancer. Image processing is generated due to different reasons, such as sampling to noise, initial digitalization, and intensity. The main purpose is to improve the Sobel filter based on interval analysis. The results are based on MIAS, which is an applicable database for breast cancer detection. This method provides efficient result in the presence of uncertainties. In image acquisition, the noise should be reduced. This method is not efficient for image processing. In 2013, Razmjooy et al. [28] have utilized the diagnosis system for malignant melanomas, which eliminates extra hairs using edge detection and deduces a color image that will segment by intensity thresholding. At last, the authors performed image analysis to measure a set of candidate attributes that contain enough information to differentiate malignant from benign melanomas. SMO did not require matrix memory and thus the training issues of SVM can be avoided. It needs a lot of time, so difficult to reach the real-time response.

In 2020, Razmjooy et al. [29] have introduced the importance of image processing for cancer detection. They

provided a literature review on the automatic skin cancer detection and discussed the various steps of skin cancer detection process based on the implantation capabilities. This method is computationally inexpensive, also it is fast and simple to implement. It provides extremely susceptible noise. Training time is long so that the overtraining should be avoided.

In 2018, Razmjooy et al. [30] have presented a method to detect cancer. The scales are eliminated using edge detection and smoothing, and then segmentation was done. Here, the morphological operations can be used to eliminate the extra information where the melanoma boundary exists. In this work, the WCO was used to optimize MLP neural network, which detects the malignancy in melanoma. This method shows good efficiency. It is difficult to show the problems in the network.

### 2.2 Review

Though there are various computer-aided detection approaches for diagnosing breast cancer and mammographic lesion, there are still some manually developed features so that a new mechanism needs to be introduced for acquiring the best performance in image classification. Table 1 describes the features and challenges of existing breast cancer detection approaches.

## 3 Architectural view of proposed breast cancer diagnosis

### 3.1 Proposed methodology

Various research contributions have proven to diagnose breast cancer using mammogram images to reduce mortality rate at a high level. The Deep Learning has risen the medical image processing, especially tumor detection. Moreover, it is being utilized for detecting breast cancer, which is also suggested by the radiologist to accomplish the cancer diagnosis as it is still improving the prediction outcomes effectively. Architectural model of the proposed breast cancer detection model is shown in Fig. 1.

The developed model involves four main phases: "pre-processing, tumor segmentation, Feature extraction, and classification." The median filtering is used to remove the noise from the image during pre-processing. Once mammogram image is pre-processed, it is subjected to tumor segmentation. Here, the optimized region growing approach is used for segmenting the tumor. As a modification to the conventional region growing algorithm, the tolerance is optimized or tuned by a new optimization algorithm termed as FC-CSO in such a way that the segmentation accuracy is maximum. $M_m$ represents input mammogram image.

The pre-processed image is indicated as $M_{m(pre)}$, and the tumor segmented binary image is represented as $M_{m(seg)}$. In addition, features like GLCM and GLRM extract the segmented binary image. In the detection side, hybrid classifiers with the integration of two well-performing clarifiers like CNN and RNN are used. In the hybrid classifier, the number of hidden layers is used by the proposed FC-CSO algorithm. The main objective of optimized classification is to maximize detection accuracy. Finally, the detection results in three classes namely normal, benign, and malignant categories corresponding the input images.

### 3.2 Image pre-processing

In this phase, median filter is for eradicating the irrelevant noise from the image restraining the low or high frequencies that enhance or recognize the edges in the image. Generally, median filtering [31] is a non-linear filter. The major intention of this filter is to substitute the noisy and neighbourhood pixels of an image, and then sorted based on the image's gray level. When the median filter is given to the input image $M_m$, the output of this phase is determined based on Eq. (1), which is denoted as $M_{m(pre)}$, in which the 2D-mask is indicated as $K$:

$$M_{m(pre)}(r,s) = med\{M_m(r-v, s-w)v, w \in K\} \tag{1}$$

Hence after filtering, the output $M_{m(pre)}$ is formed that could be applied to further processing.

### 3.3 Feature extraction

Two feature sets, such as GLCM and GLRM are extracted, and description is given below.

***GLCM*** [32]: It is an approach for measuring the spatial relationship of pixels. Moreover, it is used to evaluate the probability of pixels pairs by certain values that assist in investigating the spatial relationship in an image. The formulation for energy is denoted in Eq. (2):

$$ENR = \sum_u \sum_v B_{uv}^2 \tag{2}$$

In Eq. (2), the $(u,v)^{th}$ element in the normalized GLCM is specified by $B_{uv}$. The numerical formulation for entropy, contrast, and homogeneity is denoted in Eqs. (3), (4), and (5), respectively.

$$ENY = -\sum_u \sum_v B_{uv} \log_2 B_{uv} \tag{3}$$

$$CON = \sum_u \sum_v (u-v)^2 B_{uv} \tag{4}$$

**Table 1** Features and challenges of existing breast cancer detection models

| Authors [citations] | Methodology | Features | Challenges |
|---|---|---|---|
| Wang et al. [18] | CNN | It has high classification and detection accuracy<br>It is used to extract the uses from the region of interest | It is computationally expensive |
| Geweid and Abdallah [19] | Non-parametric approach | It is an accurate and reliable mechanism<br>It demonstrates the affected region in an efficient borderline, and it is relatively not affected by mammogram noise | It requires more computational time<br>Needs more statistical power if the guess of approximately equivalent parametric tests is valid |
| Singh et al. [20] | ADEWNN | It has high accuracy<br>It is more stable and efficient | It is not fully automatic because the region of interest is extracted manually |
| Li et al. [21] | DenseNet neural network | Enhanced the efficiency of the network model as well as calculation speed<br>To merge different activations, this model can learn easily that leads to accurate classification | It is highly expensive |
| Cruz et al. [22] | MLP | It has high accuracy when compared to other models<br>It is useful for regression and mapping | Training a huge dataset on vast problem consumes more time |
| Kaur et al. [23] | SVM | SVM has high classification accuracy<br>It works well when there is a clear margin of division among the classes | It is not suitable for huge databases<br>If the number of features for every data point exceeds the number of the training data sample, it doesn't work well |
| Hai et al. [24] | S-DenseNet | It has a high performance<br>It has an anti-over fitting capability | It requires more amount of memory |
| Duraisamy and Emperumal [25] | Fully complex-valued relaxation network | It has a high performance<br>It doesn't require more effort while training | It is required to enhance classification accuracy |
| Guo et al. [26] | Hukuhara difference, interval analysis, Sobel filtering | It has a necessary efficiency in presence of uncertainties | Noise should be reduced in the image acquisition<br>Image processing is not efficient |
| Razmjooy et al. [27] | Morphology-based technique | SMO did not need matrix memory, therefore SVM training issues can be avoided | It is time consuming<br>It is difficult to implement to reach the real-time response |
| Razmjooy et al. [28] | Image acquisition technique | It is computationally inexpensive<br>It is fast and simple to implement | The noise is extremely susceptible<br>Overtraining should be avoided |
| Razmjooy et al. [29] | Hybrid neural network- world cup optimization algorithm | It shows good efficiency | Difficulty of showing the problem to the network |

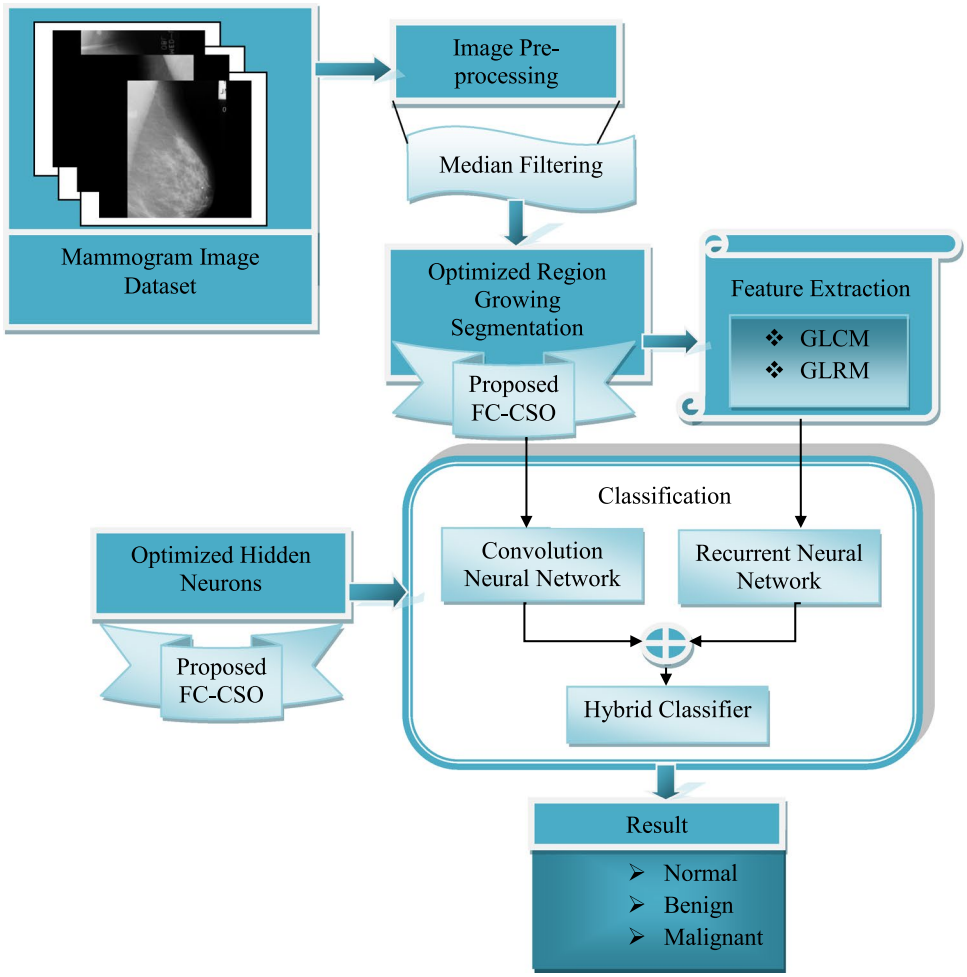


**Fig. 1** Block diagram of developed breast cancer detection model

$$HMG = \sum_{u}\sum_{v}\frac{1}{1+(u-v)^2}B_{uv} \tag{5}$$

Moreover, the variance, sum average, and correlation are given in Eqs. (6), (7), and (8), respectively. Here, the mean $B_{uv}$ is indicated by $\mu$ the count of grey level in image is given by $N_G$. The mean and standard deviation of $B_u$ and $B_v$ is denoted by $\mu_r$, $\mu_s$, and $\sigma_r$, $\sigma_s$, respectively.

$$VAR = \sum_{u}\sum_{v}(u-\mu)^2 B_{uv} \tag{6}$$

$$SAV = \sum_{u=2}^{2N_G-2} uB_{r+s}(u) \tag{7}$$

$$COR = \frac{\sum_{u}\sum_{v}(u\times v)B_{uv} - \mu_r\mu_s}{\sigma_r\sigma_s} \tag{8}$$

Also, the "sum variance, sum entropy, difference entropy, difference variance, information metrics of correlation, and maximum correlation coefficient" are expressed in Eqs. (9–14), respectively.

$$SVA = \sum_{u=2}^{2N_G}(u-SAV)^2 B_{r+s}(u) \tag{9}$$

$$SEN = \sum_{u=2}^{2N_G} B_{r+s}(u)\log\left\{B_{r+s}(u)\right\} \tag{10}$$

$$DEN = \sum_{u=0}^{N_{G-1}} B_{r-s}(u)\log\left\{B_{r-s}(u)\right\} \tag{11}$$

$$DVA = Variance\ of\ B_{r-s} \tag{12}$$

$$IMC1 = \frac{EPQ - EPQ1}{\max\{EP, EQ\}} \tag{13}$$

$$IMC2 = \sqrt{(1 - \exp[-2.0[EPQ2 - EPQ]])} \tag{14}$$

From Eqs. (13), and (14), $EPQ$, $EPQ1$, and $EPQ2$ is denoted in Eqs. (15), (16), and (17), respectively.

$$EPQ = -\sum_{u}\sum_{v} B_{uv} \log_2 B_{uv} \tag{15}$$

$$EPQ1 = -\sum_{u}\sum_{v} B_{uv} \log_2 \left\{B_r(u)B_s(v)\right\} \tag{16}$$

$$EPQ2 = -\sum_{u}\sum_{v} B_r(u)B_s(v) \log_2 \left\{B_r(u)B_s(v)\right\} \tag{17}$$

The mathematical formula for MCC is given in Eq. (18):

$$MCC = \sum_{w} \frac{B(u,w)B(v,w)}{B_r(u)B_s(w)} \tag{18}$$

***GLRM*** [33]: It is assumed as a matrix, where the texture analysis obtains the texture features. By increasing the image intensity, the higher-order statistical data is shown by GLRM. The calculation of short-run emphasis, long-run emphasis, grey level non-uniformity, run-length non-uniformity, run percentage, LGRE, HGRE is represented in Eqs. (19–25), respectively.

$$LR^{EM} = \frac{1}{n}\sum_{u,v} v^2 H(u,v) \tag{19}$$

$$SR^{EM} = \frac{1}{n}\sum_{u,v} \frac{H(u,v)}{v^2} \tag{20}$$

$$GL^{NU} = \frac{1}{n}\sum_{u}\left(\sum_{v} H(u,v)^2\right) \tag{21}$$

$$RL^{NU} = \frac{1}{n}\sum_{u}\left(\sum_{u} H(u,v)^2\right) \tag{22}$$

$$RN^{PER} = \sum_{u,v} \frac{n_t}{H(u,v)v} \tag{23}$$

$$LGR^{EM} = \frac{1}{n}\sum_{u,v} \frac{H(u,v)}{u^2} \tag{24}$$

$$HGR^{EM} = \frac{1}{n}\sum_{u,v} u^2 H(u,v) \tag{25}$$

From the above equations, the second-order joint conditional probability density function is denoted by $H(u, v)$. Therefore, the combination of GLCM and GLRM acquired, at last, is given as $R_z$, where $z = 1, 2, \cdots, N_Z$ and the whole count of features is denoted by $N_Z$.

## 4 Optimized segmentation and classification for diagnosing breast cancer

### 4.1 Solution encoding

This research work intends to solve the problem facing under tumor segmentation and classification of mammogram images. For maximizing the segmentation accuracy and classification accuracy, two sets of optimization are done. The proposed FC-CSO algorithm optimizes the threshold value of region growing and the hidden layer of hybrid classifier (CNN + RNN = CRNN). In region growing algorithm, $\delta$ indicates the threshold function. The number of hidden layers of CNN is denoted as $CE_L$ and RNN as $RE_L$. The bounding limit of the solution is between 5 and 25.

### 4.2 Objective model

The objective model of the developed breast cancer detection focuses on tumor segmentation and its categorization. The accuracy is determined using Eq. (26), where $Tr^P$, $Tr^N$, $Fa^P$, and $Fa^N$ is the true positive, true negative, false positive, and false negative elements. Moreover, maximized objective function is shown in Eq. (27):

$$ACC = \frac{Tr^P + Tr^N}{Tr^P + Tr^N + Fa^P + Fa^N} \tag{26}$$

$$OB = Max(ACC) \tag{27}$$

This objective function is attained by adopting the optimization strategy with a new modified algorithm called FC-CSO algorithm on tumor segmentation and its classification in the proposed breast cancer diagnosis framework.

### 4.3 Optimized region growing

The output of the pre-processing phase image i.e. $M_{m(pre)}$ is further processed for segmenting the tumor from an image.

The segmentation process is utilized by a single seed region growing algorithm named optimized region growing algorithm. Figure 2 shows the diagrammatic representation of the proposed region growing -based segmentation.

Seed region growing [34] is mostly used for different image segmentations. This region growing algorithm works based on the group of pixels into image regions. In this the image is done by a seed pixel and then attaches the homogeneous pixels until the segmentation is satisfied.

Assume $W$ as whole area of input image $M_{m(pre)}$, which is divided into $o$ sub-regions $W_1, W_2, \cdots, W_o$. The below-specified conditions must be fulfilled during the region growing algorithm.

1. $\bigcup_{u=1}^{o} W_o = W$;
2. To $W_1, W_2, \cdots, W_o, W$ indicates the connected region;
3. For any $u, v, u \neq v, W_u \cap W_v = \emptyset$;
4. $J(W_u) = true$;
5. $W(W_u \cup W_v) = false$;

From the above conditions, the grey level value of the set $W_u$ is denoted by $J(W_u)$, and the whole sum of pixels of an image is given by $o$. As discussed before, the starting step of the seeded region growing algorithm is the selection of seed points. Every phase of the algorithm includes that the above set is added by one pixel, while the value of pixels meets the condition given in Eq. (29). The average of the image $t$ is denoted in Eq. (28):

$$t = \frac{1}{o} \sum_{(r,s) \in W} V(r,s) \tag{28}$$

$$|V(r,s) - t|_{(r,s) \in W} < \delta \tag{29}$$

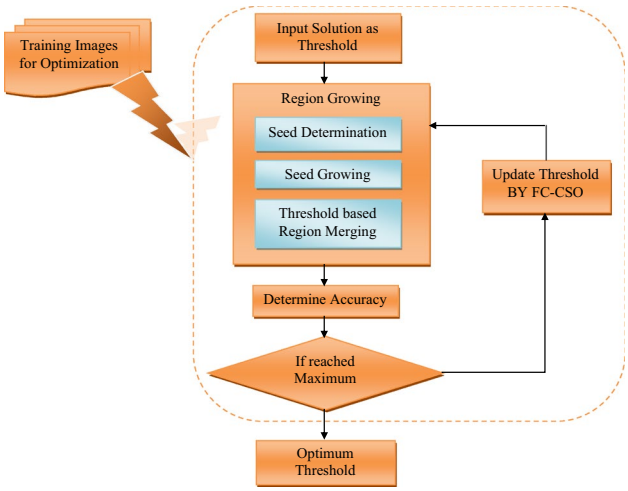


**Fig. 2** Optimized region growing for tumor segmentation

From Eqs. (39) and (40), $V(r,s)$ denotes the value at the coordinate $(r,s)$. In addition, the term $\delta$ is determined as tolerance or threshold that must be very close to the average of an image. Accordingly, the working condition of this model is the selection of seed and threshold plays a key role in improving the segmentation performance. The traditional seeded region growing algorithm requires human involvement for choosing these parameters that initiate the subjective elements, and the accurate tuning of parameters are demanded. For attaining accurate automatic segmentation, the threshold of the seeded region growing algorithm is optimized by the help of suggested FC-CSO algorithm, which must be in the range of [0, 256].

### 4.4 Existing chicken swarm optimization

The classic CSO [35] is inspired by mimicking the hierarchical order of chicken swarms. CSO is capable to extract the chickens swarm to optimize the efficient problems. The behaviors of the chicken are idealized by the rules that are as follows:

In a chicken swarm, many groups exist in a dominant rooster, a couple of hens and the chicks in each group. The chicken swarms are divided according to the fitness values of chickens. The worst fitness values are termed as chicks whereas the best fitness values are termed as roosters, and others are denoted as hens. The mother–child relationship among the chicks and hens are demonstrated randomly, which remains unchanged. The status of the dominance and mother–child relationship are updated at multiple time steps $tst$. Chickens follow the group-mate rooster while searching for food, chickens follow their as they are avoided by other ones for eating food. The mother will search food for their chicks. The dominant rooster has a benefit in competition for food.

Assume the number of roosters is denoted by RC, the number of hens is indicated by HC, the number of chicks is represented by CC, and the number of mother hens is denoted by $_{\text{MH}}$ respectively. All $C$ virtual chickens are denoted by their locations at $tst$ time steps in $ds$ dimensional space is given by $X_{a,b}^{tst}(a \in [1, \cdots, C], b \in [1, \cdots, ds)$. As the roosters having the best fitness values has more priority in acquiring the food so that those chickens can search for food in vast places, which is formulated in Eq. (30), in which standard deviation is indicated by $\sigma^2$ and is given in Eq. (31), and the random number is denoted by $Ran$:

$$X_{a,b}^{tst+1} = X_{a,b}^{tst} * (1 + Ran(0, \sigma^2)) \tag{30}$$

$$\sigma^2 = \begin{cases} 1, \; iffn_a \leq fn_c, \\ \exp\left(\frac{(fn_c - fn_a)}{|fn_a| + \varepsilon}\right), \; otherwise, \; c \in [1, C], c \neq a \end{cases} \tag{31}$$

From Eq. (31), the fitness value is indicated by $fn$, $c$ denotes rooster's index, and $\varepsilon$ used to avoid zero division

error. The submissive hens have less benefit in having food compared to the dominant hens. The corresponding equation is formulated in Eq. (32), where, a random number is indicated by *Ran* over [0, 1]. Moreover, the rooster index of $a^{th}$ group hen's groupmate is denoted by $rt1 \in [1, \cdots, C]$, and the rooster index of chicken selected at random is given by $rt2 \in [1, \cdots, C]$. The fitness of $fn_a > fn_{rt1}$, $fn_a > fn_{rt2}$ therefore $S2 < 1 < S1$. The equations for $S1$ and $S2$ are shown in Eqs. (33), and (34), respectively. The chicks move around their mother in order to find the food. This is given in Eq. (35):

$$X_{a,b}^{tst+1} = X_{a,b}^{tst} + S1 * Ran * \left(X_{rt1,b}^{tst} - X_{a,b}^{tst}\right) + S2 * Ran * \left(X_{rt2,b}^{tst} - X_{a,b}^{tst}\right) \tag{32}$$

$$S1 = \exp\left(\left(fn_a - fn_{rt1}\right)/\left(abs(fn_a) + \varepsilon\right)\right) \tag{33}$$

$$S2 = \exp\left(\left(fn_{rt2} - fn_a\right)\right) \tag{34}$$

$$X_{a,b}^{tst+1} = X_{a,b}^{tst} + FLM * (X_{mo,b}^{tst} - X_{a,b}^{tst}) \tag{35}$$

In Eq. (35), $X_{mo,b}^{tst}$ is the position of $a^{th}$ chicks mother, $FLM(FLM \in (0, 2))$ is a parameter. Moreover, *FLM* every chick is randomly chosen from 0 to 2, and the individual differences are considered. The pseudo-code of traditional CSO is shown in Algorithm 1.

### 4.5 Firefly algorithm

The inspiration of FF [36] is inspires the social behaviour of fireflies in tropical summer sky. By using bioluminescence with different flashing patterns, they communicate, search for prey, and locate mates. Moreover, few characteristics of FF are idealized for developing a FF algorithm. In order to make simple, only three rules are employed that are as follows:

(1) The whole FF's are unisex so one FF is attracted to others irrespective of the sex.
(2) The pleasant appearance is relative concerning contrast. Less contrast FF is moving in the direction of more contrast one. Moreover, attractiveness is proportional to the contrast that reduces by increasing the distance among the FF's. When there are no brighter FF's than a specific FF, it force to move about in the freedom at random.
(3) The contrast of FF is in some way connected to analytical form of cost function. In order to maximize the issue, contrast is being proportional to the cost function.

The starting positions of agents are defined in the space randomly, which is shown in Eq. (36), in which the minimum and the maximum allowable values for $p^{th}$ variable are given by $X_{p,\min}$, and $X_{p,\max}$. The initial value of $p^{th}$ variable for $q^{th}$ agent is given by $X_{p,q}^{(in)}$:

$$X_{p,q}^{(in)} = X_{p,\min} + rndm \cdot \left(X_{p,\max} - X_{p,\min}\right), p = 1, 2, \cdots, D \tag{36}$$

### 4.6 Hybrid FC-CSO algorithm

The conventional CSO has the ability to attain a high precision convergence rate but comparing with low dimensional cases the searching time is prolonged in the discrete optimization problems. Hence, to operate with a high convergence rate, the beneficial concept of FF is used here. In the proposed FC-CSO algorithm, the chick update is replaced by the FF update that works better for performing the proposed segmentation and classification effectively. Algorithm 1 shows the pseudo-code for the proposed FC-CSO.

**Algorithm 1**: Pseudocode of developed FC-CSO

Initialize the chickens population as $C$ and describe the related parameters
Validate the fitness value of $C$ chickens, $tst$ is initialized to 0
**while** $(tst < Max\_Gen)$
**if** $(tst\%TS = 0)$
Perform the fitness ranking
Categorize the cat swarms
**end if**
**for** $a = 1:C$
**if** $a = rooster$
Update its location by Eq. (30)
**end if**
**if** $a = hen$
Update its solution using Eq. (32)
**end if**
**if** $a = chick$
Update its solution by FF based on Eq. (36)
**end if**
Validate the new solution
If the new solution is best than the earlier one, update it
**end for**
**end while**

### 4.7 Hybrid classifier

The combination of CNN and RNN termed as CRNN is used for classifying the images into normal, benign, and malignant. An optimized hidden layer concept is to develop the novelty of the proposed model into a hybrid CRNN algorithm. The working flow of the optimized hybrid classifier is shown in Fig. 3.

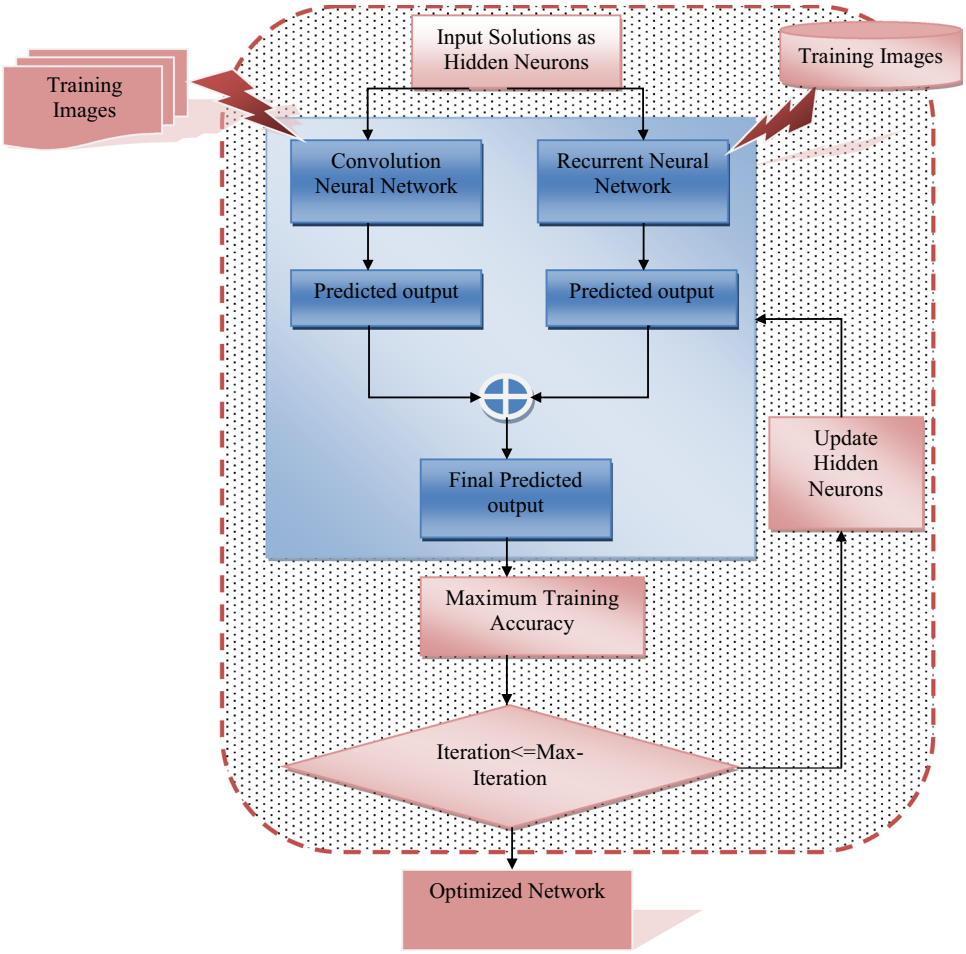


**Fig. 3** Working principle of the optimized hybrid classifier by FC-CSO algorithm

**CNN**: The basic structure of CNN comprises of several layers like convolutional layers, pooling layers, and one or more fully connected layers. It involves the combination of three layers. In CNNs the information flows from input to output direction.

***Convolutional layers***: This type of layer act an extractor, and therefore these layers help to learn the representation of input images. The neurons are structured into feature maps. Each neuron is associated with the neighborhood of the prior layer by the group of trainable weights, and it is named as a filter bank. The $d^{th}$ output feature map for $Co_d$ is computed using Eq. (8), in which $INP$ is the input image, the convolutional filter linked with $d^{th}$ feature map is denoted by $fm_d$, and the 2D convolutional operator is denoted by $*$, which is utilized for measuring the interior result of the filter in each region of the input image. Moreover, the non-linear creation function is denoted by $f(.)$. This allows for extracting the non-linear features:

$$Co_d = f(fm_d * INP) \tag{37}$$

Moreover, the two kinds of functions, such as hyperbolic tangent and sigmoid functions, are usually utilized.

***Pooling layers***: The feature map is utilized to decrease the spatial resolution which helps to attain the spatial invariance to input translations and distortions. It is a general tradition to employ common pooling aggregation layers for transmitting the normal input values of the minute region of an image to the succeeding level. The convolution by multiple input maps are merged in each resultant map, and this equation is represented in Eq. (38):

$$Co_d^{CL} = f(\sum_{k \in A_l} Co_d^{CL-1} * KN_{kl}^{CL} + AB_l^{CL}) \tag{38}$$

In Eq. (9), the convolutional layer is denoted by $CL$, and the downsampling layer is denoted by $CL-1$, the input features of downsampling layer is indicated by $Co^{CL-1}$, the kernel maps and additive bias of convolutional layer is indicated by $KN_{kl}^{CL}$ and $AB_l^{CL}$, respectively, the selection input maps are done by $A_l$, and the input and output are denoted as $k, l$.

***Fully connected layers***: The final phase of CNNs includes a generic multi-layer network. Several layers are stacked to extract the abstract feature depictions by moving in the network. This type of layer follows the feature depictions of pooling and convolutional layers for performing high-level reasoning functions.

***Training***: Both ANNs and CNNs make use of learning algorithms for altering the public parameters for attaining the necessary outcome of the network. The usual algorithm utilized for this objective is the backpropagation algorithm.

**RNN**: It [37] is a classification of ANN, where the links among the nodes create a directed graph with a progression of information. It can work with time-series data effectively, and therefore the resultant appears to be best when determining the present and the earlier data. LSTM is one type of RNN that consists of a memory cell unit. It consist of three gate units, namely input, forget, and output gates. Moreover, LSTM is used to resolve the gradient mass and explosion. On updating the memory cell unit's state by three gates, LSTM removes the unnecessary data and seize the required data in series efficiently. GRU is introduced as a special kind of LSTM, which is used to build the RNN model for enhancing performance. GRU combines the output and forgets gates into a unique update gate $UG_e$, where the linear interpolation assists in acquiring the current result. Assume, $g_e \leftarrow Fs_{com}$ is the $e^{th}$ input feature, and $h_{e-1}$ is the earlier hidden state. The updated gate is represented in Eq. (39), and the reset gate is shown in Eq. (40):

$$UG_e = Acf\left(WM^{gUG} g_e + WM^{hUG} h_{e-1}\right) \tag{39}$$

$$RG_e = Acf\left(WM^{gRG} g_e + WM^{hRG} h_{e-1}\right) \tag{40}$$

From Eqs. (39) and (40), the activation function is denoted by $Acf$, which is a logistic sigmoid function. The weight matrix is denoted by $WM^e = \left\{WM^{gUG}, WM^{hUG}, WM^{gRG}, WM^{hRG}\right\}$ that has to be tuned appropriately for minimizing the fault discrepancy among the predicted and actual output. The candidate state of the hidden unit is measured using Eq. (41), in which the element-wise multiplication is indicated by $\otimes$:

$$\tilde{h}_e = \tan\left(WM^{gh} g_e + WM^{hh}\left(h_{e-1} \otimes RG_e\right)\right) \tag{41}$$

The linear interpolation among $h_{e-1}$ and candidate state $\tilde{h}_e$ is named as $e^{th}$ hidden activation function $h_e$ of GRU, and the equation is given in Eq. (42):

$$\begin{aligned} h_e &= \left(1 - UG_e\right) \otimes \tilde{h}_e + UG_e \otimes h_{e-1} h_e \\ &= \left(1 - UG_e\right) \otimes \tilde{h}_e + UG_e \otimes h_{e-1} \end{aligned} \tag{42}$$

Here, the proposed FC-CSO algorithm optimizes the number of hidden layers of CNN and RNN. On performing the AND operation to both of the output generated from CNN and RNN, the final classified output is generated. Thus the three classes, such as normal, benign, and malignant categories, are obtained from the input mammogram images.

## 5 Results and discussions

### 5.1 Experimental set-up

The implemented breast cancer detection model was proposed using MATLAB 2018a, and also performance examination was performed. The Type I, and Type II measures are analyzed using performance analysis. Type I measures are positive measures like Accuracy, Sensitivity, Specificity, Precision, NPV, F1Score and MCC, and Type II measures are negative measures like FPR, FNR, and FDR. The state-of-the-art techniques were compared from the proposed model of the experimental results through validation with graphs and tabulations. The population size and iterations considered for the experiment were 10 and 25, respectively. The performance analysis of the suggested approach was compared over CSO [35], CSA [38], FF [36], and IAP-CSA [39]. The classification analysis was compared over NN [40], SVM [41], fuzzy [42], CNN [43], RNN [37], CRNN [43, 37].

### 5.2 Performance measures

In this paper, the following ten performance measures are evaluated:

(a) Accuracy: It is a "ratio of the observation of exactly predicted to the whole observations," as shown in Eq. (26).

(b) Sensitivity: It measures "the number of true positives, which are recognized exactly,"

$$Sen = \frac{Tr^P}{Tr^P + Fa^N} \tag{43}$$

(c) Specificity: It measures "the number of true negatives, which are determined precisely."

$$Spe = \frac{Tr^N}{Fa^P} \tag{44}$$

(d) Precision: It is "the ratio of positive observations that are predicted exactly to the total number of observations that are positively predicted".

$$\Pr e = \frac{Tr^P}{Tr^P + Fa^P} \tag{45}$$

(e) FPR: It is computed as "the ratio of the count of false-positive predictions to the entire count of negative predictions."

$$FPR = \frac{Fa^P}{Fa^P + Tr^N} \tag{46}$$

(f) FNR: It is "the proportion of positives which yield negative test outcomes with the test."

$$FNR = \frac{Fa^N}{Tr^N + Tr^P} \tag{47}$$

(g) NPV: It is the "probability that subjects with a negative screening test truly don't have the disease."

$$NPV = \frac{Fa^N}{Fa^N + Tr^N} \tag{48}$$

(h) FDR: It is "the number of false positives in all of the rejected hypotheses."

$$FDR = \frac{Fa^P}{Fa^P + Tr^P} \tag{49}$$

F1 score: It is the "harmonic mean between precision and recall. It is used as a statistical measure to rate performance".

$$F1score = \frac{Sen \bullet \Pr e}{\Pr e + Sen} \tag{50}$$

(j) MCC: It is a "correlation coefficient computed by four values."

$$MCC = \frac{Tr^P \times Tr^N - Fa^P \times Fa^N}{\sqrt{\left(Tr^P + Fa^P\right)\left(Tr^P + Fa^N\right)\left(Tr^N + Fa^P\right)\left(Tr^N + Fa^N\right)}} \tag{51}$$

### 5.3 Segmentation analysis

The experimental outcomes of breast cancer detection are shown in Fig. 4. The analysis of the proposed region growing algorithm and conventional tumor segmentation in mammogram images are tabulated in Table 2. The accuracy of the proposed FC-CSO-region growing algorithm is truly defined the true observations from the whole observations. It is 16.1% superior to conventional region growing, 1.3% superior to CSO, 0.3% superior to CSA, 0.1% superior to FF and IAP-CSA-based region growing. Moreover, the sensitivity of the developed FC-CSO-region growing is correctly determined the true observations. Therefore, the sensitivity of the suggested IAP-CSA-based region growing is 79.7% better than the region growing, 56.7%, 2.1%, 3% 1.9% superior to CSO, CSA, FF, and IAP-CSA-region growing. The precision of the developed FC-CSO-based-region growing is 7.6% improved than region growing, 55.7% improved than CSO-region growing, 59.6% improved than CSA- region growing, and 4.7% improved than FF and IAP-CSA-region growing. From the results, it has been confirmed that the suggested FC-CSO-region growing is superior in segmenting the tumor from the mammogram images.

### 5.4 Performance analysis

The conventional heuristic-based CRNN and the performance analysis is shown in Fig. 5. From Fig. 5a, the accuracy of the developed FC-CSO-CRNN is determined

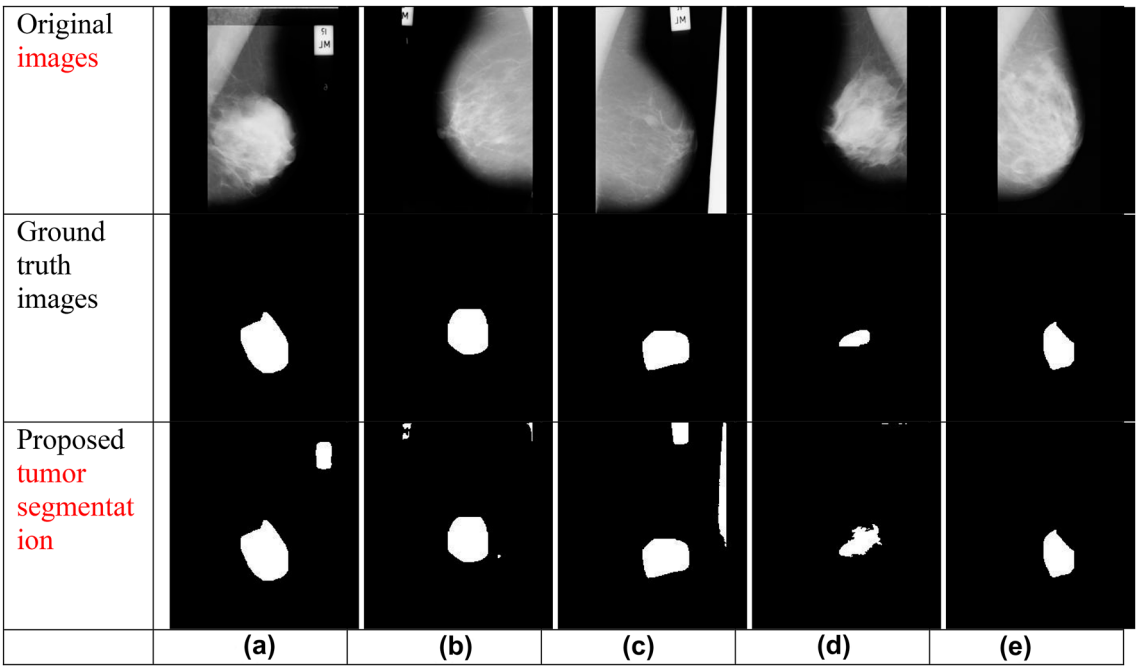


**Fig. 4** Experimental results of tumor segmentation for breast cancer detection for **a** image 1, **b** image 2, **c** image 3, **d** image 4, and **e** image 5

**Table 2** Segmentation analysis of the proposed FC-CSO-region growing algorithm and conventional tumor segmentation in mammogram images using region growing algorithm

| Metrics | Region growing [34] | CSO-region growing [35] | CSA-region growing [38] | FF-region growing [36] | IAP-CSA-region growing [39] | FC-CSO-region growing |
|---|---|---|---|---|---|---|
| Accuracy | 0.84781 | 0.97157 | 0.98098 | 0.98348 | 0.98387 | 0.98489 |
| Sensitivity | 0.12029 | 0.3796 | 0.60829 | 0.57756 | 0.58386 | 0.59521 |
| Specificity | 0.99805 | 0.98727 | 0.98653 | 0.99892 | 0.99892 | 0.99996 |
| Precision | 0.9272 | 0.44167 | 0.40253 | 0.95308 | 0.95308 | 0.99814 |
| FPR | 0.00195 | 0.012727 | 0.013465 | 0.001082 | 0.001081 | 4.30E−05 |
| FNR | 0.87971 | 0.6204 | 0.39171 | 0.42244 | 0.41614 | 0.40479 |
| NPV | 0.99805 | 0.98727 | 0.98653 | 0.99892 | 0.99892 | 0.99996 |
| FDR | 0.0728 | 0.55833 | 0.59747 | 0.046921 | 0.046921 | 0.001863 |
| F1-score | 0.21295 | 0.40829 | 0.48447 | 0.71926 | 0.72412 | 0.74573 |
| MCC | 0.30249 | 0.39499 | 0.48573 | 0.73505 | 0.73922 | 0.76478 |

correctly for all the learning percentages in a precise manner. The accuracy of the proposed FC-CSO-CRNN is 0.5%, 1%, 1.5%, 4% superior to CSO-CRNN, CRNN, FF-CRNN, and IAP-CSA-CRNN at learning percentage 45%. Moreover, the sensitivity of the modified FC-CSO-CRNN model has precisely defined the true positive values, which is shown in Fig. 5b. Thus, the sensitivity of the suggested FC-CSO-CRNN is 2.3% enhanced than CSA, 6.6% enhanced than FF, 8.4% enhanced than CRNN, 9% enhanced than CSO, and 10.9% enhanced than IAP-CSA-based CRNN at learning percentage 65%. Moreover, the precision of the developed FC-CSO-CRNN is positive observations from the whole positive observations are shown in Fig. 5d. It is 2.3%, 2.9%, 6%,10.1% superior to FF, CRNN, CSA, and IAP-CSA-CRNN at learning percentage 45%.

The overall performance of the proposed and existing methods based on CRNN is shown in Table 3. The accuracy of the proposed FC-CSO-CRNN is 3.3%, 2.3%, 0.4%, 0.9% better than CSO, CSA, FF, and IAP-CSA-based CRNN, respectively. Similarly, the precision of the developed FC-CSO-CRNN is defined correctly. The precision of the recommended FC-CSO-CRNN is 8.3%, 6.5%, 14%, 17.7% enhanced than CSO, CSA, FF, and IAP-CSA-CRNN. Finally, it is confirmed that the developed FC-CSO-CRNN is performing well in breast cancer detection comparing to other improved CRNN.

### 5.5 Comparative analysis over different classifiers

The comparative analysis of the proposed model over various classifiers is given in Table 4. The total positive and negative observations correctly determine the accuracy of the suggested FC-CSO-CRNN. Here, the accuracy of the modified FC-CSO-CRNN is 8.2% improved than NN, 34.2% improved than SVM, 40.2% improved than Fuzzy, 1.1% improved than CNN, 11.7% improved than RNN, and 34.4% improved than CRNN. Also, the sensitivity of the suggested FC-CSO-CRNN has exactly defined the true observations from the whole observations. The sensitivity of the developed FC-CSO-CRNN is 47.8%, 60.5%, 5.6%, 4.1%, 29.1%, 81.2% better than NN, SVM, fuzzy, CNN, RNN, and RNN. Similarly, the precision of the improved FC-CSO-CRNN is 7.4% enhanced than NN, 65.5% enhanced than SVM, 61.9% enhanced than fuzzy, 9.6% enhanced than CNN, 9.2% enhanced than RNN, and 45.6% enhanced than CRNN. The proposed model offers the best performance for the conventional models in detecting breast cancer in mammogram images.

## 6 Conclusion

A new method has been developed for detecting breast cancer in mammogram images using an optimized hybrid classifier. Here, the proposed model has undergone a few steps, such as image pre-processing, tumor segmentation, feature extraction, and detection. At first, the noise was removed from the mammogram image with the help of median filtering. In the segmentation phase, the tumor was segmented using an optimized region growing algorithm, which was based on a hybrid meta-heuristic algorithm by the combination of two algorithms such as CSO and FF, called FC-CSO. Later, feature extraction was performed using the GLCM and GRLM features. In the detection stage, the combination of two deep learning approaches like CNN and RNN was integrated. In addition, for RNN, GLRM features were given as input, whereas the segmented binary image was given to CNN for obtaining high accuracy. The accuracy of the proposed FC-CSO-CRNN was 3.3%, 2.3%, 0.4%, 0.9% better than CSO, CSA, FF, and IAP-CSA-based CRNN, respectively. Similarly, the precision of the developed FC-CSO-CRNN was 8.3%, 6.5%, 14%, 17.7% better than CSO, CSA, FF, and IAP-CSA-CRNN. The classification analysis provided the results as the accuracy of the suggested

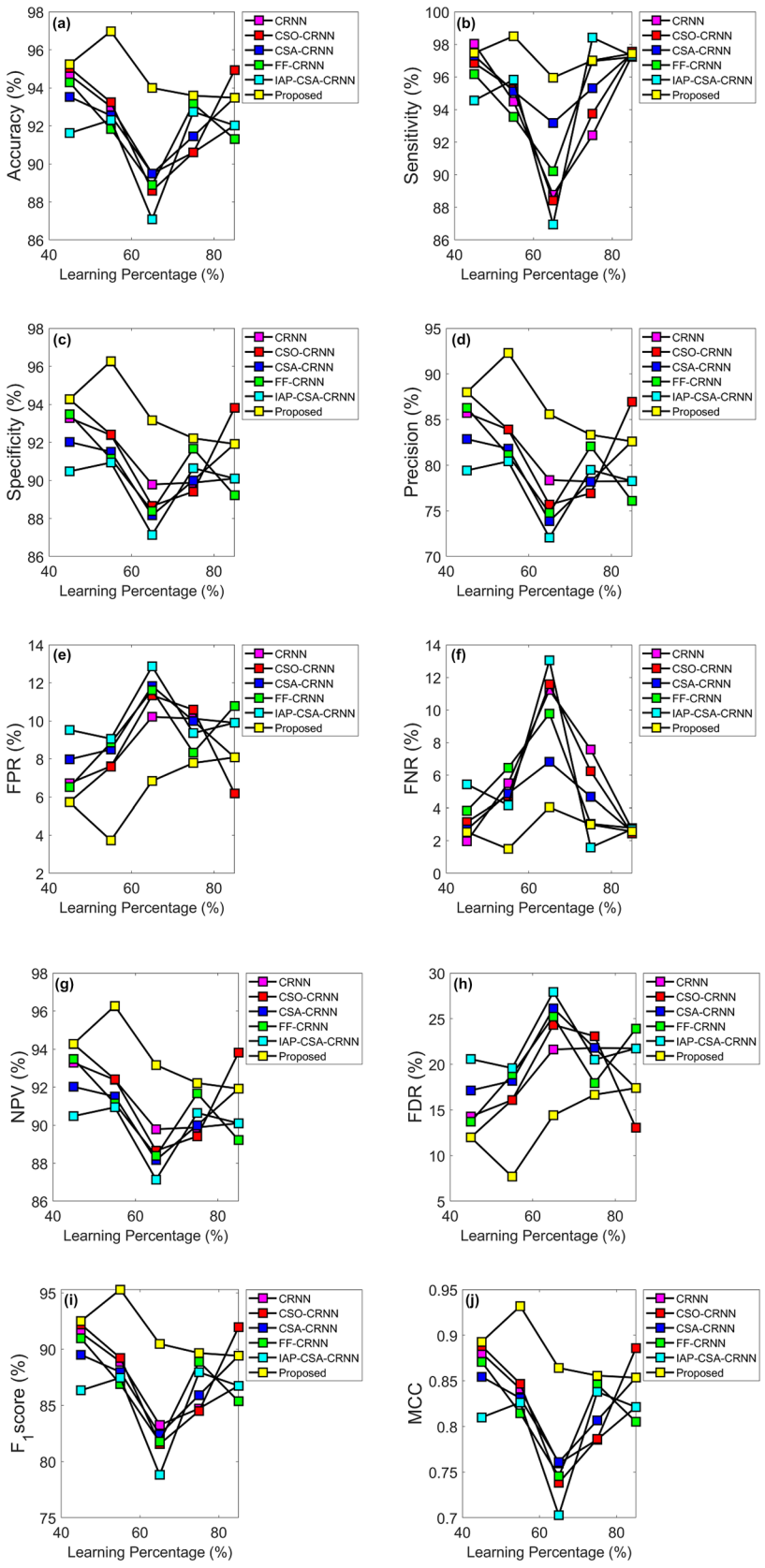


**Fig. 5** Performance analysis of proposed and conventional heuristic-based CRNN oriented breast cancer detection models with performance metrics **a** accuracy, **b** sensitivity, **c** specificity, **d** precision, **e** FPR, **f** FNR, **g** NPV, **h** FDR, **i** F1-score, and **j** MCC

**Table 3** Overall performance analysis of the proposed and conventional CRNN-based breast cancer detection models

| Performance metrics | CSO-CRNN [35] | CSA-CRNN [38] | FF-CNN+RNN [36] | IAP-CSA-CRNN [39] | FC-CSO-CRNN |
|---|---|---|---|---|---|
| Accuracy | 0.90598 | 0.91453 | 0.93162 | 0.92735 | 0.9359 |
| Sensitivity | 0.9375 | 0.95313 | 0.9697 | 0.98413 | 0.97015 |
| Specificity | 0.89412 | 0.9 | 0.91667 | 0.90643 | 0.92216 |
| Precision | 0.76923 | 0.78205 | 0.82051 | 0.79487 | 0.83333 |
| FPR | 0.10588 | 0.1 | 0.083333 | 0.093567 | 0.077844 |
| FNR | 0.0625 | 0.046875 | 0.030303 | 0.015873 | 0.029851 |
| NPV | 0.89412 | 0.9 | 0.91667 | 0.90643 | 0.92216 |
| FDR | 0.23077 | 0.21795 | 0.17949 | 0.20513 | 0.16667 |
| F1-score | 0.84507 | 0.85915 | 0.88889 | 0.87943 | 0.89655 |
| MCC | 0.78637 | 0.80671 | 0.84611 | 0.83796 | 0.85566 |

**Table 4** Analysis on proposed breast cancer detection models over conventional classifiers

| Performance metrics | NN [40] | SVM [41] | Fuzzy [42] | CNN [43] | RNN [37] | CRNN [43, 37] | FC-CSO-CRNN |
|---|---|---|---|---|---|---|---|
| Accuracy | 0.83681 | 0.675 | 0.64583 | 0.91667 | 0.81053 | 0.67368 | 0.90598 |
| Sensitivity | 0.625 | 0.36458 | 0.875 | 0.8875 | 0.71579 | 0.51 | 0.92424 |
| Specificity | 0.94271 | 0.7526 | 0.53125 | 0.93125 | 0.85789 | 0.76216 | 0.89881 |
| Precision | 0.84507 | 0.26923 | 0.48276 | 0.86585 | 0.71579 | 0.53684 | 0.78205 |
| FPR | 0.057292 | 0.2474 | 0.46875 | 0.06875 | 0.14211 | 0.23784 | 0.10119 |
| FNR | 0.375 | 0.63542 | 0.125 | 0.1125 | 0.28421 | 0.49 | 0.075758 |
| NPV | 0.94271 | 0.7526 | 0.53125 | 0.93125 | 0.85789 | 0.76216 | 0.89881 |
| FDR | 0.15493 | 0.73077 | 0.51724 | 0.13415 | 0.28421 | 0.46316 | 0.21795 |
| F1-score | 0.71856 | 0.30973 | 0.62222 | 0.87654 | 0.71579 | 0.52308 | 0.84722 |
| MCC | 0.62094 | 0.10548 | 0.39161 | 0.81381 | 0.57368 | 0.27553 | 0.78568 |

FC-CSO-CRNN was 8.2% improved than NN, 34.2% improved than SVM, 40.2% improved than Fuzzy, 1.1% improved than CNN, 11.7% improved than RNN, and 34.4% improved than CRNN. Finally, it is concluded that the proposed FC-CSO-CRNN is superior in detecting breast cancer from mammogram images. The proposed method detects the breast cancer using the mammogram images, which reduces the need of chemotherapy or breast removal and saves lives by detecting the breast cancer during early stage. In this work, we used the median filter at the pre-processing stage for noise removal. Even though the median filter has several benefits, such as straightforward, offers a sensible noise removal performance, it has a drawback of removing blurs image details and thin lines even at low noise densities. Hence, in future, we will use other filters, like Wiener filter for noise removal.

## References


1. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, Jemal A (2017) Colorectal cancer statistics, 2017. CA Cancer J Clin 67(3):177–193
2. Sapate S, Talbar S, Mahajan A, Sable N, Desai S, Thakur M (2019) Breast cancer diagnosis using abnormalities on ipsilateral views of digital mammograms. Biocybern Biomed Eng 40:290–305
3. Pawar MM, Talbar SN (2016) Genetic fuzzy system (GFS) based wavelet co-occurrence feature selection in mammogram classification for breast cancer diagnosis. Perspect Sci 8:247–250
4. Raghavendra U, Acharya UR, Fujita H, Gudigar A, Tan JH, Chokkadi S (2016) Application of gabor wavelet and locality sensitive discriminant analysis for automated identification of breast cancer using digitized mammogram images. Appl Soft Comput 46:151–161
5. Ji Z, Lou C, Yang S, Xing D (2012) Three-dimensional thermoacoustic imaging for early breast cancer detection. Med Phys 39(11):6738–6744
6. Mohanty AK, Senapati MR, Lenka SK (2016) Retraction note to an improved data mining technique for classification and detection of breast cancer from mammograms. Neural Comput Appl 27(1):249–249
7. Timmers JM, van Doorne-Nagtegaal HJ, Zonderland HM, van Tinteren H, Visser O, Verbeek AL, den Heeten GJ, Broeders MJ (2012) The breast imaging reporting and data system (BI-RADS) in the Dutch breast cancer screening programme: its role as an assessment and stratification tool. Eur Radiol 22(8):1717–1723
8. Al-antari MA, Al-masni MA, Park S-U, Park JH, Metwally MK, Kadah YM, Han S-M, Kim T-S (2018) An automatic computer-aided diagnosis system for breast cancer in digital mammograms via deep belief network. J Med Biol Eng 38(3):443–456

9. Ogiela MR, Krzyworzeka N (2016) Heuristic approach for computer-aided lesion detection in mammograms. Soft Comput 20(10):4193–4202
10. Bhateja V, Misra M, Urooj S, Lay-Ekuakille A (2013) A robust polynomial filtering framework for mammographic image enhancement from biomedical sensors. IEEE Sens J 13(11):4147–4156
11. Bhateja V, Patel H, Krishn A, Sahu A, Lay-Ekuakille A (2015) Multimodal medical image sensor fusion framework using cascade of wavelet and contourlet transform domains. IEEE Sens J 15(12):6783–6790
12. Bai PR, Liu QY, Li L, Teng SH, Li J, Cao MY (2013) A novel region-based level set method initialized with mean shift clustering for automated medical image segmentation. Comput Biol Med 43(11):1827–1832
13. Li X, Radulovic M, Kanjer K, Plataniotis KN (2019) Discriminative pattern mining for bBreast cancer histopathology image classification via fully convolutional autoencoder. IEEE Access 7:36433–36445
14. Wei D, Weinstein S, Hsieh M, Pantalone L, Kontos D (2019) Three-dimensional whole breast segmentation in sagittal and axial breast MRI with dense depth field modeling and localized self-adaptation for chest-wall line detection. IEEE Trans Biomed Eng 66(6):1567–1579
15. Chiang T, Huang Y, Chen R, Huang C, Chang R (2019) Tumor detection in automated breast ultrasound using 3-D CNN and prioritized candidate aggregation. IEEE Trans Med Imaging 38(1):240–249
16. Wang X-F, Min H, Zou L, Zhang Y-G (2015) A novel level set method for image segmentation by incorporating local statistical analysis and global similarity measurement. Pattern Recognit 48(1):189–204
17. Shin H-C, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics, and transfer learning. IEEE Trans Med Imaging 35(5):1285–1298
18. Wang Z, Li M, Wang H, Jiang H, Yao Y, Zhang H, Xin J (2019) Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features. IEEE Access 7:105146–105158
19. Geweid GGN, Abdallah MA (2019) A novel approach for breast cancer investigation and recognition using M-level set-based optimization functions. IEEE Access 7:136343–136357
20. Singh SP, Urooj S, Lay-Ekuakille A (2016) Breast cancer detection using PCPCET and ADEWNN: a geometric invariant approach to medical X-ray image sensors. IEEE Sens J 16(12):4847–4855
21. Li H, Zhuang S, Li D-a, Zhao J, Ma Y (2019) Benign and malignant classification of mammogram images based on deep learning. Biomed Signal Process Control 51:347–354
22. da Cruz TN, da Cruz TM, dos Santos WP (2018) Detection and classification of mammary lesions using artificial neural networks and morphological wavelets. IEEE Lat Am Trans 16(3):926–932
23. Kaur P, Singh G, Kaur P (2019) Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification. Inf Med Unlocked 16:100151
24. Hai J, Tan H, Chen J, Wu M, Qiao K, Xu J, Zeng L, Gao F, Shi D, Yan B (2019) Multi-level features combined end-to-end learning for automated pathological grading of breast cancer on digital mammograms. Comput Med Imaging Graph 71:58–66
25. Duraisamy S, Emperumal S (2017) Computer-aided mammogram diagnosis system using deep learning convolutional fully complex-valued relaxation neural network classifier. IET Comput Vis 11(8):656–662
26. Razmjooy N, Loschi HJ, Estrela VV (2019) A study on metaheuristic-based neural networks for image segmentation purposes. Data Sci. https://doi.org/10.1201/9780429263798-2
27. Guo G, Razmjooy N (2019) A new interval differential equation for edge detection and determining breast cancer regions in mammography images. Syst Sci Control Eng 7(1):346–356
28. Razmjooy N, Mousavi BS, Soleymani F, Khotbesara MH (2013) A computer-aided diagnosis system for malignant melanomas. Neural Comput Appl 23(7–8):2059–2071
29. Razmjooy N, Ashourian M, Karimifard M, Estrela VV, Loschi HJ, do Nascimento D, França RP, Vishnevski M (2020) Computer-aided diagnosis of skin cancer: a review. Curr Med Imaging Rev. https://doi.org/10.2174/1573405616666200129095242
30. Razmjooy N, Sheykhahmad FR, Ghadimi N (2018) A hybrid neural network–world cup optimization algorithm for melanoma detection. Open Med 13(1):9–16
31. Zhu Y, Huang C (2012) An improved median filtering algorithm for image noise reduction. Phys Proc 25:609–616
32. Malegori C, Franzetti L, Guidetti R, Casiraghi E, Rossi R (2016) GLCM, an image analysis technique for early detection of biofilm. J Food Eng 185:48–55
33. Radhakrishnan M, Kuttiannan T (2012) Comparative analysis of feature extraction methods for the classification of prostate cancer from TRUS medical images. IJCSI Int J Comput Sci Issues 9(1):171–179.
34. Feng Q, Gao B, Lu P, Woo WL, Yang Y, Fan Y, Qiu X, Gu L (2018) Automatic seeded region growing for thermography debonding detection of CFRP. NDT E Int 99:36–49
35. Meng X, Liu Y, Gao X, Zhang H (2014) A new bio-inspired algorithm: chicken swarm optimization In: International Conference in Swarm intelligence, pp 86–94
36. Gandomi AH, Yang X-S, Talatahari S, Alavi AH (2013) Firefly algorithm with chaos. Commun Nonlinear Sci Numer Simul 18(1):89–98
37. Li F, Liu M (2019) A hybrid convolutional and recurrent neural network for hippocampus analysis in Alzheimer's disease. J Neurosci Methods 323:108–118
38. Askarzadeh A (2016) A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm. Comput Struct 169:1–12
39. Patil RS, Biradar N (2019) Improved region growing segmentation for breast cancer detection: progression of optimized fuzzy classifier. In Communication
40. Fernández-Navarro F, Carbonero-Ruz M, Alonso DB, Torres-Jiménez M (2017) Global sensitivity estimates for neural network classifiers. IEEE Trans Neural Netw Learn Syst 28(11):2592–2604
41. Yu S, Tan KK, Sng BL, Li S, Sia ATH (2015) Lumbar ultrasound image feature extraction and classification with support vector machine. Ultrasound Med Biol 41(10):2677–2689
42. Sousa MJ, Moutinho A, Almeida M (2019) Classification of potential fire outbreaks: a fuzzy modeling approach based on thermal images. Expert Syst Appl 129:216–232
43. Namatēvs I (2017) Deep convolutional neural networks: structure feature extraction, and training. Inform Technol Manag Sci 20:40–47