PRZEGLĄD ELEKTROTECHNICZNY

**Ukazuje się od 1919 roku** **1'2023**

Organ Stowarzyszenia Elektryków Polskich • Wydawnictwo SIGMA-NOT Sp. z o.o.

**P.Esther Jebarani[1], N.Umadevi[2]**

Department of Computer Science & IT, Sri Jayendra Saraswathy Maha Vidyalaya College of Arts and Science, Coimbatore,India(1)(2).
ORCID: 1. https://orcid.org/0000-0003-3685-2670; 2. https://orcid.org/0000-0001-9021-2637

doi:10.15199/48.2023.01.01

# Grey Wolf Optimization Based Breast Cancer Detection using 1D Convolution LSTM Classifier

***Abstract***. *Women are particularly vulnerable to breast cancer. Breast cancer diagnosis has benefited greatly from the utilization of ultrasound imaging. Breast UltraSound (BUS) image segmentation remains a difficult challenge due to low image quality. Furthermore, BUS image segmentation, as well as classification, is an important stage in the analysis process. Initially, the image associated with breast cancer is gathered from MIAS database. The gathered image undergoes pre-processing operation using the adaptive median filtering technique. Subsequently, the segmentation is performed in the pre-processed images using the hybrid method consisting of GMM and K-Means. These segmented images undergo the feature extraction steps further where the features are extracted by utilizing the Gray Level Co-occurrence Matrix (GLCM). Grey Wolf Optimization (GWO) selects the optimal features for further classification using a novel 1D Convolution LSTM. Here, the pooling layer of 1D CNN is replaced by the LSTM. The objective function behind the optimal feature selection and classification is the accuracy maximization. Finally, the novel One Dimensional Convolution Long Short Term Memory (1 DCLSTM) classifies the outcome into normal, benign, and malignant, respectively. The proposed method is compared with the other state of art methods related to this research.*

***Streszczenie***. *Kobiety są szczególnie narażone na raka piersi. Diagnostyka raka piersi bardzo skorzystała na wykorzystaniu obrazowania ultrasonograficznego. Segmentacja obrazu UltraSound (BUS) piersi pozostaje trudnym wyzwaniem ze względu na niską jakość obrazu. Ponadto segmentacja obrazu BUS, a także klasyfikacja, jest ważnym etapem procesu analizy. Początkowo obraz związany z rakiem piersi pozyskiwany jest z bazy MIAS. Zgromadzony obraz jest poddawany wstępnemu przetwarzaniu przy użyciu techniki adaptacyjnego filtrowania medianowego. Następnie na wstępnie przetworzonych obrazach przeprowadzana jest segmentacja metodą hybrydową składającą się z GMM i K-Means. Te podzielone na segmenty obrazy przechodzą kolejne etapy ekstrakcji cech, w których cechy są wyodrębniane przy użyciu macierzy współwystępowania poziomu szarości (GLCM). Optymalizacja Gray Wolf (GWO) wybiera optymalne funkcje do dalszej klasyfikacji przy użyciu nowatorskiego rozwiązania 1D Convolution LSTM. W tym przypadku warstwa łączenia 1D CNN zostaje zastąpiona przez LSTM. Funkcją celu stojącą za optymalnym doborem i klasyfikacją cech jest maksymalizacja dokładności. Wreszcie, powieść jednowymiarowa pamięć krótkoterminowa z konwolucją jednowymiarową (1 DCLSTM) klasyfikuje wynik odpowiednio na normalny, łagodny i złośliwy. Proponowana metoda jest porównywana z innymi nowoczesnymi metodami związanymi z tymi badaniami.* (***Wykrywanie raka piersi w oparciu o optymalizację Gray Wolf przy użyciu klasyfikatora 1D Convolution LSTM***)

**Keywords:** Breast Cancer Detection; Classification; Grey Wolf Optimization; Optimal Feature Selection; One Dimensional Convolution Long Short Term Memory; Accuracy maximization
**Słowa kluczowe**: rak piersi, algorytm szarego wilka, diagnostyka

## Introduction

The UDH, ADH, and DCIS are a collection of cytological and structural heterogeneous that often are restricted to the lobular system [11]. These are relevant because they are linked to an enhanced chance of developing cancer in the future, albeit in varying degrees [12]. It has greater risk of invasive cancer, correspondingly, when compared to the reference group [13]. Recent immunological phenotypic variations caused by clonal gene mutations pose a danger of invasion and metastasis of varying magnitudes [14]. Consequently, there exists no indication that UDH is a precursor lesion at this time. Every year, roughly 2,50,000 new instances of intraductal lesions are expected to be identified in United States. Percutaneous needle biopsies are the normal version of therapy for diagnosing image-detected and palpable breast abnormalities. Patients having UDH should have regular checkups as well as therapeutic treatments [15]. As a result, the care of individuals diagnosed with ADH/DCIS and UDH may differ drastically based on the findings of the percutaneous biopsy.

Cancer describes a condition in which cells continue to proliferate uncontrollably. The therapy may be very effective when the cancer is detected early [16]. As per TNM staging approach, a comprehensive procedure of breast cancer diagnosis on pathological imaging includes grading (cell analysis), cancer classification/detection, and staging (the degree of the illness) [17]. It is, nevertheless, time demanding, and even the most experienced pathologists make mistakes. Furthermore, trained pathologists are in short supply, particularly in rural regions. With the advanced soft computing paradigm, one of the computing technique called CNN algorithms for image identification, image semantic segmentation, and other challenges have been presented, with promising results [18]. Machine learning algorithm is used to train the pathological images, to perform auto-classification as well as detection of the early breast cancer in women [19]. Classification process is challenging for pathologists, it requires scanned and labelled medical images, and the size may be Giga pixel range [20]. To validate the classifier with the huge datasets by tuning hyper parameter of the model.

The image's low-level visual contents such as edges, brightness, contrast, texture, contour, and so on are fine-tuned by pre-processing technique [21]. When the borders are in continuous or not sufficiently apparent, segmentation based on such low-level image attributes frequently fails [22]. This is because majority of the traditional segmentation algorithms (such as Snake and level group) trying to discover edges in order to spot and extract an item without knowing the complete process [23]. Even though

few learning-based algorithms may categorize or cluster pixels to detect distinct areas, they rely on low-level image attributes and are unable to comprehend the many objects [24]. However, the large amount of breast image data is required, then it is trained to improve the learning rate for determining tumor area and its different kinds [25].
The contribution of the article is:

- To develop the novel breast cancer segmentation and classification for early diagnosis in women.
- To classify the trained features using a novel 1 DCLSTM, where the pooling layer of 1D CNN is replaced by the LSTM.
- To optimize feature selection phase and the classification phase using GWO, where the features, epoch count of 1D CNN, and hidden neuron count of LSTM are tuned with the intention of accuracy maximization.

The organization of this article is as follows: Section I depicts the introduction of breast cancer. Section II explains the literature survey. Section III describes the proposed model and data description for the introduced breast cancer detection and classification. Section IV expresses the pre-processing technique for the proposed model. Section V deals with the feature extraction and classification for the proposed breast cancer detection. Section VI explains the results and discussion. Section VII draws the conclusion of this research work.

## Literature Survey

### A. Related Works

Chen *et al.,* 2021 [1] on a patch level, researchers looked at the difficulty of recognizing breast cancer metastases in pathological images. When compared to the standard supervised technique, experimental findings suggested that the method was beneficial in improving patch-level classification accuracy. Dundar *et al.,* 2011 [2] have unveiled a prototype approach for identifying breast microscopic tissues autonomously. This method analysed with scanned tissue slides for specific cytological criteria as well as repeatability when utilized in tandem. Zhou *et al.,* 2018 [3] have described about SWE data to identify aggressive and benign breast cancers using a segmentation-free radiomics technique. The approach was designed to combine the benefits of both SWE and CNN in automated feature extraction and appropriate classification. The suggested technique, in contrast to standard techniques was aimed to extract characteristics from a dataset without the need for segmentation or human intervention. This preserved peri-tumor data, which was lost in segmentation-oriented approaches. The suggested technique's improved performance compared to traditional approaches, as well as its automated nature, showed that it contained a lot of promise for clinical detection.

Saha and Chakraborty, 2018 [4] with low user participation, have developed an approach for recognizing breast cancer images. The manual measurement of HER2 has been a long-standing challenge for pathologists and hence suggested HER2 DNN (Her2Net). Several key components exist for the suggested HER2Net framework's convolutional as well as de-convolutional sections. For classification as well as error estimation, a softmax layer and a fully connected layer were also utilized. Ultimately, the categorization findings were used to compute HER2 scores. The deployment score were the primary contributions of the introduced HER2Net system. The suggested Her2Net's has better accuracy and broad application in the framework of HER2.

Ashraf *et al.,* 2013 [5] a methodological approach for multichannel MRFs has been proposed. To find characteristics that meet conditional independence requirements that employed conditional mutual information. This approach is to include kinetic feature maps improved MRI to MRF as observation channels for tumour segmentation. It is compared with the results to those of the well-known normalized cuts segmentation technique, as well as other widely utilized algorithms for breast tumour segmentation, such as FCM and the recent technique of running FCM-VES. The AUC of these preceding approaches were 0.92, 0.88, and 0.60, accordingly. It is looked at the effect of superior segmentation in tumour characterization as well as feature extraction. This is investigated with the impact of better segmentation on forecasting the likelihood of tumor recurrence using a confirmed tumour gene expression test.

Huang *et al.,* 2020 [6] have suggested a new way to separate breast tumours using semantic classification as well as patch merging. To crop a ROI on the original image, the suggested approach first picked two diagonal points. Utilizing SLIC, the cropped image is separated into multiple super-pixels. A bag-of-words method might also be developed by extracting some features from the super-pixels. A BPNN and KNN might be used to achieve the classification process. The repository from the BUS dataset with 320 sample sets is used to train and test with the target data. Segmentation technique is used to determine the hand-labeled tumor contours with better TP and FP compared to the existing approaches.

Zhou *et al.,* 2021 [7] have suggested a unique multi-task training model for ABUS image tumour segmentation as well as classification. The suggested architecture is made up of two sub-networks: a classification light-weight multi-scale network and a segmentation encoder-decoder network. There are imprecise borders of tumors in ABUS images in the previous iterations, few more iterations are taken to minimize the artifacts.

Wang *et al.,* 2016 [8] BCH images have been subjected to an automated quantitative image analysis approach. To produce ROIs, WDMR were coupled, resulting in accurate localization. To divide overlapping cells with greater precision and resilience, a DSSM approach is used. The features were derived for cell nuclei categorization. SVM having CAGA data is used to find the best feature set. Normal as well as malignant cell images might be classified with an accuracy of 96.19% (0.31%), a sensitivity of 99.05% (0.27%), and a specificity of 93.33% (0.81%).

Mouelhi *et al.,* 2018 [9] have proposed a novel method for segmentation and classification for tumor nuclei. It is used to recover the entire stained nuclei areas and divided overlapping nuclei in the suggested segmentation technique. To obtain quick tumour nuclei detection, stromal cells were eliminated from the segmented image utilizing adaptive criteria. Furthermore, a combination of four standard colour separation methods are used a unsupervised learning for classification of tumour nuclei for further Allred cancer grading.

Belhaj *et al.,*2020 [10] have suggested a one-stage end-to-end UNet framework for detecting breast tumour. Utilizing the publically accessible datasets in breast and DDSM, the suggested method is tested with respect to its effectiveness in segmenting and categorizing breast tumour. These datasets has the mass segmentation as well as classification process showed an IOU score of 90.50% and an AUC of 99.88%.

## Proposed model and datasets for breast cancer classification

### A. Proposed Model

Initially, the datasets are gathered from MIAS repository that consist of normal, benign, and malignant images. The

collected data undergoes the pre-processing using adaptive median filtering technique. Further, the segmentation is performed using the hybrid method that consists of GMM and K-means approach. The segmented images undergo the feature extraction for extracting the features. Here, the feature extraction is done by GLCM. Since, the extracted feature seems to be lengthy, it is necessary to select the optimal features for reducing the complexity. This is accomplished by GWO such that it chooses the optimal features. These optimally selected features undergo the final classification that is done using the enhanced CNN technique. Here, the pooling layer of 1D CNN is replaced using LSTM, thus the enhanced CNN here is referred as 1DCLSTM. This 1DCLSTM classifies the final output as whether it is normal, benign, and malignant. Figure 1 depicts the proposed model for image segmentation and classification.

### B. Datasets

The data associated with the classification part of breast cancer are gathered from MIAS database that consist of normal, benign, and malignant images. Some of the sample images associated with each category is described in Figure 2.

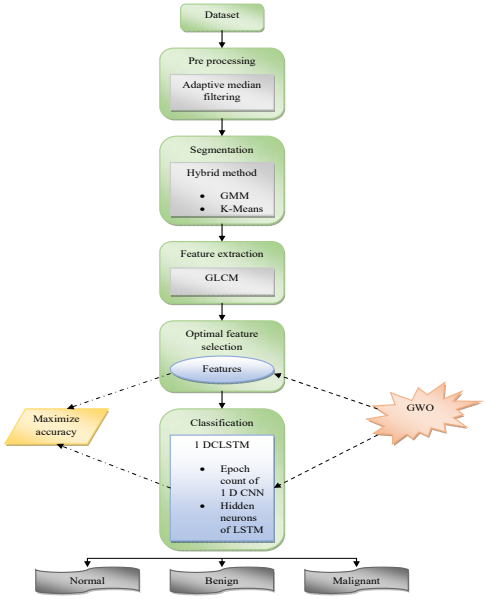


Fig.1. Proposed model for image segmentation and classification

## Pre-processing for the introduced breast cancer

### A. Pre-processing

Pre-processing is the most important stage in breast cancer analysis because the acquired breast cancer image is of poor quality. It is critical for correcting and adjusting the breast cancer image in preparation for subsequent analysis and processing [49]. To reduce impulsive noise, the adaptive median filter [26] employs noise detection and filtering methods. The size associated with the window is considered to be adaptive, is nothing but if the stated criteria are not met, the window size is enlarged. Let $J_{\min}$ be the window's minimum pixel value, $J_{jk}$ be the corrupted image's pixel, $J_{\max}$ be the window's maximum pixel, $X_{\max}$ be the maximum window size $X$ be the current window, that may be attained, and $J_{med}$ be the window's median. The algorithm for this filtering approach then accomplishes in two stages, as follows:

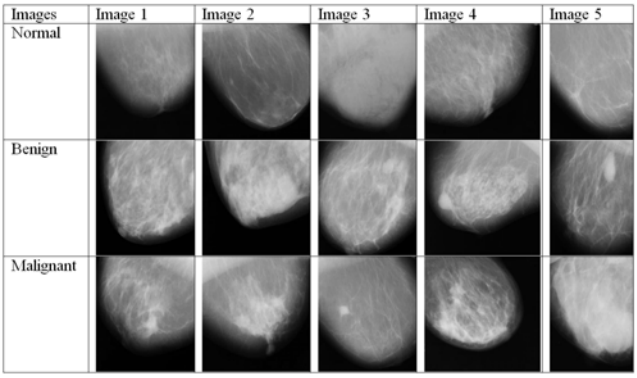


Fig.2. Sample imges used for the breast cancer classification

**Level A:**

- If $J_{\min} < J_{med} < J_{\max}$ then it is not impulse, it proceeds to Level B to determine if it is an impulse.
- Otherwise, the window size is raised and Level A is continued till the median is no longer an impulse, at which point it switches to Level B; else the maximum size is achieved, at which point the median becomes the filtered pixel.

**Level B:**

- If $J_{\min} < J_{jk} < J_{\max}$ the present pixel is not an impulse, and so the pixel remains unaffected.
- If the pixel is equal to $J_{\max}$ or $J_{\min}$, the pixel is classified the Level A median.

These are commonly employed to filter images that have been denoised with a noise level of more than 20%.

### B. Segmentation

Breast cancer diagnosis is dependent on segmentation methods. Image analysis relies heavily on segmentation, which encompasses detection, feature extraction, classification, and treatment. For treatment planning, segmentation assists physicians in quantifying the amount of tissue in the breast. Here, the segmentation is accomplished by hybrid technique that consists of GMM and the K-means method.

**GMM [27]:** As a traditional nonparametric as well as unsupervised adaptive threshold selection approach, the Otsu method uses the image's histogram.

This approach takes into account not only the grey information but also the pixels, reduced noise and the impact of the segmentation stage to some extent. Simultaneously, this system is adaptable and can autonomously pick a threshold, reducing the negative impacts of human intervention. There are probable to be local static regional as well as local rapid variation in the similar image scene. Set up distinct areas the count of pixels Gaussian distribution as per the actual situation and adequately enhance the count of Gaussian distribution. As a result, a technique is suggested for adaptively adjusting the count of Gaussian distribution pixels, which can be based on the local scene complexity [28].

Entropy represents the average quantity of data at information theory.

$$(1) \quad \begin{cases} Y_{ch} = \sum_{j=1}^{C} x_{j,u} \overline{\omega}_{j,u} \\ C = \arg\min\left(\sum_{c} x_j > U\right) \end{cases}$$

The global fuzzy entropy is lower than the local fuzzy entropy. As a result, it should use formula 2.

(2) $\left|y_{j,u}-\mu_{j,u}\right|<2.5\gamma_{j,u}$

However, in practice, owing to factors like noise interference, the histogram does not always show a clear peak value, and the Otsu technique cannot achieve excellent separation outcomes because the threshold is determined primarily by the gray-level histogram and the parameter must be upgraded as below.

(3) $x_{j,u}=(1-\alpha)x_{j,u-1}+\alpha N_{j,u}$

**K-means [29]:** The most often utilized partitional approach is the K-Means Algorithm (KMA) [30] [31], as illustrated in Algorithm I. Its most used distance measurement is Euclidean distance, which is defined as:

(4) $\sqrt{\sum_{j=1}^{e}\left(y_j-z_j\right)^2}$

In a d-dimensional Euclidean space, $y_j$ and $z_j$ represent two points. The cluster centroid $d_l$ can be changed as:

(5) $d_l=\dfrac{\sum_{y_{j\in d_l}} y_j}{d_l}$

| **Algorithm 1:** K-Means Algorithm |
|---|
| Make a set of $l$ random integers and $l$ cluster centroids |
| For every instance |
| Instance assignment to cluster having nearest centroid |
| Centroid update |
| Instance reassignment to clusters |
| Till no update |

**Feature extraction and classification process for the introduced breast cancer model**

**A. Feature Extraction**

Feature extraction describes dimensionality reduction of the extracted feature vector. One among the major familiar texture analysis approaches is the GLCM [32]. It is also known as a second-order statistic. GLCM can calculate statistical characteristics from the image's intensity [33]. Various features are considered such as correlation, homogeneity, energy, entropy and contrast. $Q(j,k)$ defines the normalised value of GLCM in the entire equations. The count of the pair of grey levels $j$ and $k$, is represented by every item $(j,k)$. The count of grey levels is denoted by $O_h$. Features that quantify coarseness, smoothness, and texture-oriented information with strong discriminating power may be determined using the Co-occurrence matrix. Energy, correlation, homogeneity, contrast, and entropy are a few examples of such measurements [34], which are provided by:

(6) $Energy=\sum_{j,k=0}^{O_h-1} q(j,k)^2$

(7) $Homogeneity=\sum_{j,k=0}^{O_h-1}\dfrac{1}{1+(j,-k)^2}Q(j,k)$

(8) $Correlation=\dfrac{\sum_{j,k=0}^{O_h-1}(j-\mu_j)(k-\mu_k)Q(j,k)}{\sigma_j\sigma_k}$

(9) $Contrast=\sum_{j,k=0}^{O_h-1}(j,k)^2 Q(j,k)$

(10) $Entropy=\sum_{j,k=0}^{O_h-1}Q(j,k)\left[\ln Q(j,k)\right]$

The means are $\mu_j$ and $\mu_k$, while the standard deviations are $\sigma_j$ and $\sigma_k$ in this equation.

(11) $\mu_j=\sum_{j,k=0}^{O_h-1} j\left[Q(j,k)\right]$

(12) $\mu_k=\sum_{j,k=0}^{O_h-1} k\left[Q(j,k)\right]$

(13) $\sigma_j=\sqrt{\sum_{j,k=0}^{O_h-1}Q(j,k)(j-\mu_j)^2}$

(14) $\sigma_k=\sqrt{\sum_{j,k=0}^{O_h-1}Q(j,k)(k-\mu_k)^2}$

All of these characteristics can help to identify two types of images with a good classification rate [35].

**B. Optimal Feature Selection**

To improve classification outcomes, recent sophisticated feature selection approaches employ the power of optimization algorithms to choose a subset of important features. Some optimization methods include a number of regulating parameters that must be tuned for optimal performance. It will reduce insignificant factors and increase classification accuracy as well as performance. Since, the extracted features are too huge; GWO is used to reduce the dimensionality of extracted feature vector. Here, out of the total 44 extracted features, 10 features are optimally selected.

**C. Objective Model**

The foremost intention of the introduced technique is to optimize the features of optimal feature selection phase, epoch count of 1D CNN, and hidden neurons of LSTM with the consideration of accuracy maximization. It is modelled as in Eq. (15).

(15) $Obj=\underset{\left\{Fe^{opt},EC^{CNN},HN^{LSTM}\right\}}{\arg\max}\left(Accuracy\right)$

Here, the objective or the fitness function is shown by $Obj$, optimally selected features are shown by $Fe^{opt}$, epoch count of 1D CNN is shown by $EC^{CNN}$, and hidden neurons of LSTM are shown by $HN^{LSTM}$ respectively. Accuracy is defined as, “correct prediction count to the total prediction count”. It is shown as below.

(16) $Accuracy=\dfrac{P^T+N^T}{P^T+N^T+P^F+N^F}$

In the above equation, the true positive is shown by $P^T$, true negative is shown by $N^T$, false positive is shown by $P^F$, and false negative is shown by $N^F$ respectively.

**D. Classification**

Breast cancer is classified mostly on the basis of the histological aspect of tissue in the tumor, however this is not always the case. Here, the classification is done using 1D CLSTM, where the pooling layer of 1D CNN is replaced by LSTM. The description of each of these classifiers is shown below.

**1D CNN [36]:** These were developed in [37] [38] to handle 1D data and offers a number of advantages, including computational gains. Convolution is conducted on data vectors in 1D-CNNs, using an input signal vector $y$ of length $O$ with a filter vector $\omega$ of length $M$. This function

may be written as (1), and it produces a 1D output layer $d$ with the length $(O-M+1)$ without zero-padding.

(17) $$d(k)=g\left(\sum_{j=0}^{M-1}\omega(j)\times(k-j)+c\right),k=0,1,\cdots,O-1$$

Here, $c$ denotes the bias term and $g(\cdot)$ defines a non-linear function, which in this case is the ReLU [39]. In a kernel window mechanism, it reaches a maximum value $v$ by having size $n\times 1$ where stride is $t$ and applies it to an input vector $d$ , yielding an output vector $e$ specified as:

(18) $$e=\max\left(v(n\times 1,t)d\right)$$

The FC layer links all of its neuron nodes using the feature map output from the previous layer as input. BN describes a technique for normalising and scaling the FC layer in order to modify its distribution and speed up training [40]. During network training, an activation function produces the expected output and it is calculated by cross-entropy loss $m$ as:

(19) $$m=\sum_{l=1}^{D}u(l)\log\left(q(l)\right)$$

Here, $D$ shows the count of classes, u(l) defines the outcome of comparing the forecasted label to the ground truth label, and q(l) describes the predicted likelihood that the input corresponds to a particular class $l$ .

But, it needs vast training data, and also does not encode the orientation as well as the position of object. Hence, to overcome the limitations, the pooling layer of 1D CNN is replaced by the LSTM thus the novel classifier method is called as 1 DCLSTM.

**LSTM [41]:** LSTM architecture includes peephole connections in the identical cell [42] [43].

It calculates a mapping from an input pattern $y=(y_1,\cdots,y_U)$ to an output series $z=(z_1,\cdots,z_U)$ by recursively computing network unit activations from $u=1$ to $U$ utilising the below equations:

(20) $$j_u=\sigma\left(X_{jy}y_u+X_{jn}n_{u-1}+X_{jd}d_{u-1}+c_j\right)$$

(21) $$g_u=\sigma\left(X_{gy}y_u+X_{gn}n_{u-1}+X_{gd}d_{u-1}+c_g\right)$$

(22) $$d_u=g_u\,e\,a_{n-1}+j_u\,e\,h\left(X_{dy}y_u+X_{dn}n_{u-1}+c_d\right)$$

(23) $$p_u=\sigma\left(X_{py}y_u+X_{pn}n_{u-1}+X_{pd}d_u+c_p\right)$$

(24) $$n_u=p_u\,e\,i(d_u)$$

(25) $$z_u=\phi\left(X_{zn}n_u+c_z\right)$$

Here, the $X$ terms signify weight matrices, $X_{jd}$ , $X_{gd}$ , $X_{pd}$ describes the diagonal weight matrices, the $c$ terms symbolise bias vectors ( $c_j$ shows the input gate bias vector), $\sigma$ defines the logistic sigmoid operation where $j$ , $p$ , $g$ and $d$ describes the input gate, output gate, forget gate and cell activation vector. Among all of which have the identical size as the cell output activation vector $n$ , where $h$ and $i$ describes the cell input and output activation functions. Here, $e$ depicts the element-wise product associated with the vectors. The framework of 1 DCLSTM utilized for the breast cancer classification is shown in Figure 3.

### A. GWO

The GWO is used for enhancing the optimal feature selection and the classification phases of the labelled model. It optimizes the features in the optimal feature selection phase, epoch count of 1 D CNN as well as hidden neurons of LSTM with the consideration of accuracy maximization. GWO [44] is designed by four kinds of grey wolves namely alpha, delta, omega and beta.

The hunting processes are executed as: seeking for prey, attacking the prey and encircling the prey.

Grey wolves engage in group hunting, which shows an intriguing social characteristic apart from their social hierarchy. The following are the key stages of grey wolf hunting:

- The prey is monitored and approached.
- Pursue prey, encircle it, and annoy it.
- Attacking the prey is a good idea.

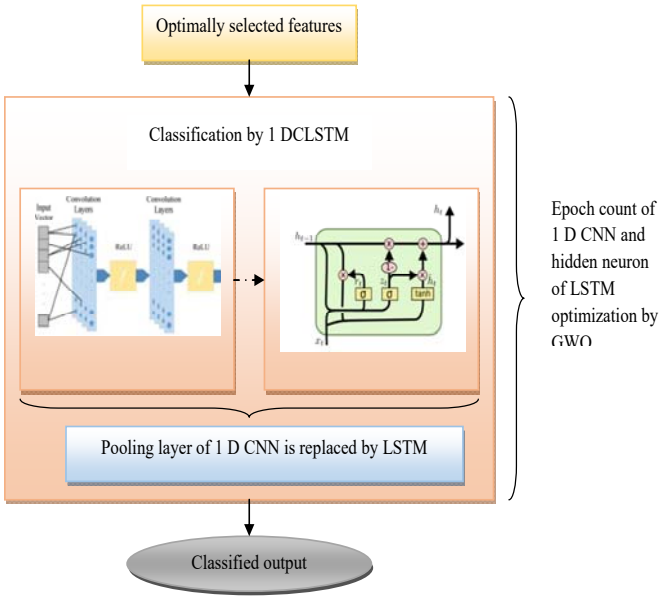


Fig.3. 1 DCLSTM for the introduced breast cancer classification

**Social hierarchy:** When creating GWO, the fitness solution for the alpha $(\alpha)$ is utilized. As a result, the second and third optimal solutions are known as beta $(\beta)$ and delta $(\delta)$ . All of the remaining are presumed to be omega $(\omega)$ . It uses $\alpha$ , $\beta$ , and $\delta$ to lead the searching (optimization).

**Encircling prey:** Grey wolves encircle victims, as previously stated. The below equations describe the behaviour:

(26) $$\vec{E}=\left|\vec{D}\cdot\vec{Y}_q(u)-\vec{Y}(u)\right|$$

(27) $$\vec{Y}(u+1)=\vec{Y}_q(u)-\vec{B}\cdot\vec{E}$$

Here, $u$ denotes the current iteration, $\vec{B}$ and $\vec{D}$ denote coefficient vectors, $\vec{Y}_q$ denotes the prey's position vector, and $\vec{Y}$ shows a grey wolf's position vector. The following is how the vectors $\vec{B}$ and $\vec{D}$ are determined:

(28) $$\vec{B}=2\vec{b}\cdot\vec{s}_1-\vec{b}$$

(29) $\vec{D} = 2 \cdot \vec{s}_2$

Over the course of repetitions, the components of $b$ are linearly lowered from 2 to 0, while $s_1$, $s_2$ shows random vectors in the range [0, 1].

**Hunting:** In this approach, the below formulae are presented.

(30) $\vec{E}_\alpha = \left|\vec{D}_1 \cdot \vec{Y}_\alpha - \vec{Y}\right|, \vec{E}_\beta = \left|\vec{D}_2 \cdot \vec{Y}_\beta - \vec{Y}\right|, \vec{E}_\delta = \left|\vec{D}_3 \cdot \vec{Y}_\delta - \vec{Y}\right|$

(31) $\vec{Y}_1 = \vec{Y}_\alpha - \vec{B}_1 \cdot \left(\vec{E}_\alpha\right), \vec{Y}_2 = \vec{Y}_\beta - \vec{B}_2 \cdot \left(\vec{E}_\beta\right), \vec{Y}_3 = \vec{Y}_\delta - \vec{B}_3 \cdot \left(\vec{E}_\delta\right)$

(32) $\vec{Y}\left(u+1\right) = \dfrac{\vec{Y}_1 + \vec{Y}_2 + \vec{Y}_3}{3}$

**Attacking prey (exploitation):** Grey wolves finish the hunt by assaulting the victim, as previously stated. We reduce the value of $\vec{b}$ to statistically represent nearing the prey.

**Search for prey (exploration):** Grey wolves process the searching primarily using the alpha, delta and beta positions. The pseudo code of GWO is given in Algorithm 2.

**Algorithm 2:** GWO [44]

Start

Grey wolf population initialization $Y_j \left(j = 1, 2, \cdots, o\right)$

Initialization of $b$, $B$, and $D$

Fitness computation of every search agent

Best search agent= $Y_\alpha$

Second best search agent= $Y_\beta$

Third best search agent= $Y_\delta$

while $u < u_{\max}$

For every search agent

$$\vec{Y}\left(u+1\right) = \frac{\vec{Y}_1 + \vec{Y}_2 + \vec{Y}_3}{3}$$

End for

Update $b$, $B$, and $D$

Fitness computation of complete search agents

Update $Y_\alpha$, $Y_\beta$, and $Y_\delta$

$u = u + 1$

End while

return $Y_\alpha$

Stop

## Results and discussion

### A. Experimental Setup

The introduced model was implemented in MATLAB and the outcomes were analyzed. The proposed approach was differentiated with conventional methods such as PSO [45], WOA [46], GA [47], and ABC [48] to describe the introduced model's superiority. The total number of samples considered is 900 and the processing timing is 3.028s.

### B. Performance Metrics

The different performance measures used here are shown below.

**Precision:** It quantifies the count of appropriate positive predictions done.

(33) $\Pr ecision = \dfrac{P^T}{P^T + P^F}$

**Recall (Sensitivity):** It quantifies the count of accurate predictions done to the entire positive predictions that have been done.

(34) $\mathrm{Re}call\left(Sensitivity\right) = \dfrac{P^T}{P^T + N^F}$

**F1 Score:** It is described as the harmonic mean among the recall and precision.

(35) $F1\,Score = \dfrac{2 \times \left(\mathrm{Re}call * \Pr ecision\right)}{\left(\mathrm{Re}call + \Pr ecision\right)}$

**Specificity:** It describes a method's capability in predicting true negatives of every available category.

(36) $Specificity = \dfrac{N^T}{\left(N^T + P^F\right)}$

**MCC:** It computes the dissimilarity among the predicted and the actual values.

(37) $MCC = \dfrac{P^T \times N^T - P^F \times N^F}{\sqrt{\left(P^T + P^F\right)\left(P^T + N^F\right)\left(N^T + P^F\right)\left(N^T + N^F\right)}}$

### C. Epoch Analysis

It can indicate a connection among two data sequences or uncover periodicities within a time series. The epoch analysis with respect to the introduced breast cancer segmentation and classification is shown here. It can be seen that as epoch increases, the error decreases, thereby describing the betterment of the proposed method. Table 1 describes the epoch analysis of the proposed model.

Table 1. Epoch Analysis

| Epoch | Error |
|---|---|
| 500 | $6.6667 \times 10^{-3}$ |
| 1000 | $5.9217 \times 10^{-3}$ |
| 1500 | $5.9514 \times 10^{-3}$ |
| 2000 | $4.9918 \times 10^{-3}$ |
| 2500 | $3.0012 \times 10^{-3}$ |
| 3000 | $3.1398 \times 10^{-3}$ |

### D. Classification Analysis

The classification analysis of the different breast cancer classification methods is revealed in Table 2. Each technique is measured using the terms such as classifiers, dataset, sensitivity, accuracy and specificity.

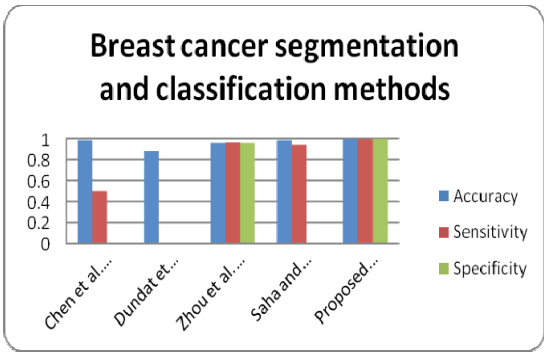


Fig.4. Classification analysis of various methods for the developed breast cancer segmentation and classification

### A. Confusion Matrix Analysis

It is described as a table that is frequently utilized for defining the classifier performance on a collection of test data and its true values are known. The confusion matrix for the developed breast cancer segmentation and classification is shown here for the actual and predicted values in terms of normal, benign, and malignant.

Table 2. Classification Analysis of Different Techniques

| Authors | Feature extraction methods | Year | Classifier | Data | Sensitivity | Accuracy | Specificity |
|---|---|---|---|---|---|---|---|
| Chen et al. [1] | Binary classifier | 2021 | Few-shot classification method | Unlabeled dataset | 0.500 | 0.984 | - |
| Dundat et al. [2] | Statistical features | 2011 | MIL Learning | Study database | - | 0.879 | - |
| Zhou et al. [3] | CNN | 2018 | CNN | SWE data | 0.962 | 0.958 | 0.957 |
| Saha and Chakraboty [4] | CNN | 2018 | TLSTM | HER2 image database | 0.940 | 0.983 | - |
| Proposed Method | GLCM + GWO | 2022 | 1 DCLSTM | MIAS dataset | 0.991 | 0.994 | 0.996 |

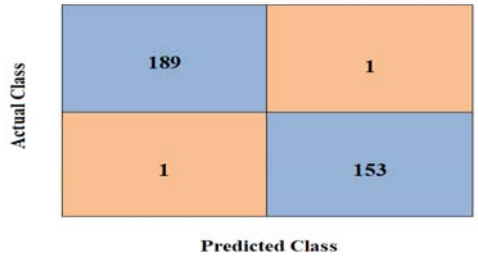


Fig.5.Confusion Matrix

Table 3. Confusion Matrix Analysis

| | | Predicted value | | |
|---|---|---|---|---|
| | | Normal | Benign | Malignant |
| Actual value | Normal | 300 | 0 | 0 |
| | Benign | 3 | 297 | 0 |
| | Malignant | 0 | 3 | 297 |

## B. Processing Time Analysis

The processing time is defined as the time it consumes to finish a defined process. Here, the output is enhanced by minimizing the processing time. The processing time in terms of feature extraction and classification with consideration of different samples for the developed breast cancer segmentation and classification is shown.

Table 4. Processing Time Analysis

| Operation | Number of samples | Processing time |
|---|---|---|
| **Feature extraction** | 40 | 180s |
| **Classification** | 40 | 240s |
| **Feature extraction** | 19 | 50s |
| **Classification** | 19 | 84s |

## C. ROC Analysis

The ROC curves are employed for assessing the diagnostic accuracy associated with a test for selecting the optimal cut-off related to a test and for comparing the diagnostic accuracy of various tests. Figure 6 shows the ROC analysis of various methods for the breast cancer segmentation and classification.

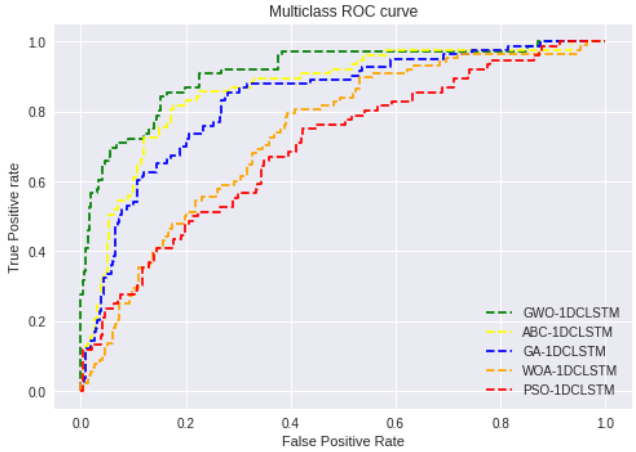


Fig.6.ROC Analysis

## D. Proposed Method Analysis

The proposed method analysis for the breast cancer segmentation and classification in terms of various measures such as recall, precision, MCC, F1 Score, specificity, and accuracy in terms of normal and abnormal classes is displayed in Tables 5 and 6. It clearly describes the measure being achieved in various classes.

Table 5.Proposed method analysis

| Class | Precision | Recall | F1 Score | MCC | Accuracy | Specificity |
|---|---|---|---|---|---|---|
| **Normal** | $9.9337 \times 10^{-1}$ | $9.9333 \times 10^{-1}$ | $9.9333 \times 10^{-1}$ | $9.9002 \times 10^{-1}$ | $9.9433 \times 10^{-1}$ | $9.9667 \times 10^{-1}$ |
| **Abnormal** | $9.9112 \times 10^{-1}$ | $9.9278 \times 10^{-1}$ | $9.8415 \times 10^{-1}$ | $9.8902 \times 10^{-1}$ | $9.9117 \times 10^{-1}$ | $9.8786 \times 10^{-1}$ |

Table 6. Comparative Analysis of Proposed method with existing works

| Table 6. Comparative Analysis of Proposed method with existing works**Measures** | **Chen et al. [1]** | | **Dundat et al. [2]** | | **Zhou et al. [3]** | | **Saha and Chakraborty [4]** | | **Proposed method** | |
|---|---|---|---|---|---|---|---|---|---|---|
| Classes | **Normal** | **Abnormal** | **Normal** | **Abnormal** | **Normal** | **Abnormal** | **Normal** | **Abnormal** | **Normal** | **Abnormal** |
| **Recall** | 0.500 | 0.499 | - | - | 0.962 | 0.961 | 0.940 | 0.932 | 0.993 | 0.993 |
| **Precision** | - | - | - | - | - | - | 0.966 | 0.965 | 0.993 | 0.992 |
| **Accuracy** | 0.984 | 0.978 | 0.879 | 0.869 | 0.958 | 0.951 | 0.983 | 0.981 | 0.994 | 0.991 |
| **MCC** | - | - | - | - | - | - | - | - | 0.990 | 0.989 |
| **F1 Score** | - | - | - | - | - | - | - | - | 0.993 | 0.991 |

### E. Comparative Analysis

The comparative analysis of the introduced model with the existing methods in terms of various measures is displayed. It can be clearly demonstrated that the introduced approach reveals better outcomes than the existing methods, thus revealing its superiority.

### F. Conclusion

This research aimed to utilize intelligent technologies in order to create a new breast cancer segmentation and classification system. The image of a breast cancer patient was initially acquired from MIAS database. The adaptive median filtering technique was used to pre-process the obtained image. Following that, a hybrid approach combining GMM and K-Means was used to segment the pre-processed images.

The feature extraction stage was performed on these segmented images, and the features were recovered using the GLCM. Because the number of features retrieved was large, the GWO was used to select the best ones. In order to classify the breast cancer at the earlier stage, the final classification phase employed the innovative 1 DCLSTM method. The LSTM replaced the pooling layer of a 1 D CNN in this case. The accuracy maximization function was the intention of the best feature selection and classification. Finally, the novel 1 DCLSTM segregated the results into three categories namely normal, malignant and benign. The simulation results demonstrated the functionality of the developed method. It can be concluded that the proposed technique obtains the accuracy of 99.43%, specificity of 99.67% and sensitivity of 99.13% of which it outperforms the other existing techniques.

## REFERENCES


[1] Chen J., Jiao J., He S., Han G., Qin J., Few-Shot Breast Cancer Metastases Classification via Unsupervised Cell Ranking, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, (2021), No.5, 1914-1923.

[2] Murat Dundar, Sunil Badve, Gokhan Bilgin, Computerized classification of intraductal breast lesions using histopathological images, *IEEE Transactions on Biomedical Engineering*, 58 (2011), 1977-1984.

[3] Yongjin Zhou, Jingxu Xu, Qiegen Liu, Cheng Li, Zaiyi Liu, Meiyun Wang, Hairong Zheng, Shanshan Wang, A Radiomics Approach With CNN for Shear-Wave Elastography Breast Tumor Classification, *IEEE Transactions on Biomedical Engineering*, 65 (2018), No. 9, 1935-1942.

[4] Saha M., Chakraborty C., Her2Net: A Deep Framework for Semantic Segmentation and Classification of Cell Membranes and Nuclei in Breast Cancer Evaluation, *IEEE Transactions on Image Processing*, 27(2018), No. 5, 2189-2200.

[5] Ashraf A B., Gavenonis S C., Daye D., Mies C., Rosen M A., Kontos D., A Multichannel Markov Random Field Framework for Tumor Segmentation With an Application to Classification of Gene Expression-Based Breast Cancer Recurrence Risk, *IEEE Transactions on Medical Imaging*, 32(2013), No. 4, 637-648.

[6] Qinghua Huang, Yonghao Huang, Yaozhong Luo, Feiniu Yuan, Xuelong Li, Segmentation of breast ultrasound image with semantic classification of superpixels, *Medical Image Analysis*, 61(2020) 1-31.

[7] Yue Zhou, Houjin Chen, Yanfeng Li, Qin Liu, Xuanang Xu, Shu Wang, Dinggang Shen, Pew-Thian Yap, Multi-task learning for segmentation and classification of tumors in 3D automated breast ultrasound images, *Medical Image Analysis*, 70(2021), 1-12.

[8] Pin Wang, Xianling Hu, Yongming Li, Qianqian Liu, Xinjian Zhu, Automatic cell nuclei segmentation and classification of breast cancer histopathology images, *Signal Processing*, 122(2016), 1-13.

[9] Aymen Mouelhi, Hana Rmili, Jaouher Ben Ali, Mounir Sayadi, Raoudha Doghri, Karima Mrad, Fast unsupervised nuclear segmentation and classification scheme for automatic all red cancer scoring in immunohistochemical breast tissue images, *Computer Methods and Programs in Biomedicine*, 165(2018), 37-51.

[10] Khaoula Belhaj Soulami, Naima Kaabouch, Mohamed Nabil Saidi, Ahmed Tamtaoui, Breast cancer: One-stage automated detection, segmentation, and classification of digital mammograms using UNet model based-semantic segmentation, *Biomedical Signal Processing and Control*, 66(2020), 1-12.

[11] Rozai J., Borderline epithelial lesions of the breast, *The American Journal of Surgical Pathology*, 15(1991), No. 3, pp. 209–221, 1991.

[12] W. A. Berg et al., Detection of breast cancer with addition of annual screening ultrasound or a single screening mri to mammography in women with elevated breast cancer risk, The Journal of the American Medical Association, 307(2012), No. 13, 1394–1404.

[13] Fitzgibbons P. L., Page D. L., Weaver D., Thor A. D., Allred D. C., Clark G. M., Ruby S. G., Malley F. O, Simpson J. F., Connolly J. L., Hayes D. F., Edge S. B., Lichter A., Schnitt S. J, Prognostic factors in breast cancer, Archives of Pathology & Laboratory Medicine Pathol, 124(2000), No. 7, 966–978.

[14] Woo Kyung Moon, Chiun Sheng Huang, Wei Chih Shen, Etsuo Takada, Analysis of elastographic and b-mode features at sonoelastography for breast tumor classification, Ultrasound in Medicine & Biology, 11(2009), No.35, 1794-802.

[15] D. M. Regner et al., “Breast lesions: Evaluation with us strain imagingcclinical experience of multiple observers,” Radiology, vol. 238, no. 2, p. 425, 2006.

[16] Jin Chung , Won Kyung Lee, Eun-Suk Cha, Jee Eun Lee, Jeoung Hyun Kim, Young Hoon Ryu, Shear-wave elastography for the differential diagnosis of breast papillary lesions. Plos One, 11(2016), No.11, 1-11.

[17] Seokmin Han, Ho-Kyung Kang, Ja-Yeon Jeong, Moon-Ho Park, A deep learning framework for supporting the classification of breast lesions in ultrasound images, Physics in Medicine & Biology, 19(2017) No.62. .

[18] Su Hyun Lee, Woo Kyung Moon, Nariya Cho, Jung Min Chang, Hyeong-Gon Moon, Wonshik Han, Dong-Young Noh, Jung Chan Lee, Hee Chan Kim, Kyoung-Bun Lee, In-Ae Park, Shear-wave elastographic features of breast cancers: comparison with mechanical elasticity and histopathologic characteristics.” Investigative Radiology, 3(2014), No.49, 147-55.

[19] Woo Kyung Moon, Shao Chien Chang, Chiun Sheng Huang, Ruey Feng Chang, Breast tumor classification using fuzzy clustering for breast elastography, Ultrasound in Medicine & Biology, 37 (2011), No. 5, 700–708.

[20] Jung Hyun Yoon, Myung Hyun Kim, Eun-Kyung Kim, Hee Jung Moon, Jin Young Kwak, Min Jung Kim, Interobserver variability of ultrasound elastography: how it affects the diagnosis of breast lesions, Ajr American Journal of Roentgenology, 3(2011), No.196, 730-736.

[21] Qi Zhang, Yang Xiao, Shuai Chen, Congzhi Wang, Hairong Zheng, Quantification of elastic heterogeneity using contourletbased texture analysis in shear-wave elastography for breast tumor classification.” Ultrasound in Medicine & Biology, 2(1015), No.41, 588–600.

[22] Woo KyungMoon, Yao-SianHuan, Yan-WeiLee, Shao-ChienChang, Chung-MingLo, Min-ChunYang, Min SunBae, Su HyunLee, Jung MinChang, Chiun Sheng Huang, Yi-TingLin, Ruey Feng Chang, Computer-aided tumor diagnosis using shear wave breast elastography,” Ultrasonics, (2017), No.78, 125-133.

[23] Qi Zhang, Yang Xia, Jingfeng Suo, Jun Shi, Jinhua Yu, Yi Guo, Yuanyuan Wang, Hairong Zheng, Sonoelastomics for breast tumor classification: A radiomics approach with clustering-based feature selection on sonoelastography,” Ultrasound in Medicine & Biology, 5(2017), No.43, no. 5, p. 1058-1069.

[24] Qi Zhang, Yang Xiao, Wei Dag, Jingfeng Suo, Congzhi Wang, Jun Shi, Hairong Zheng, Deep learning based classification of breast tumors with shear-wave elastography, Ultrasonics, (2016), No. 72, 150–157.

[25] Krouskop T.A., Wheeler T.M.,Kallel F., Garra B.S., Hall T., Elastic moduli of breast and prostate tissues under compression, Ultrasonic Imaging, 4(1998), No. 20, 260 -274.

[26] Y. Zhao, D. Li, Z. Li, Performance enhancement and analysis of an adaptive median filter, International Conference on Communications and Networking, (2007), 651-653.
[27] Xinxin Xie, Wenzhun Huang, Harry Haoxiang Wang, and Zhe Liu, Image De-noising Algorithm based on Gaussian Mixture Model and Adaptive Threshold Modeling, Proceedings of the International Conference on Inventive Computing and Informatics (ICICI 2017) IEEE Xplore Compliant, (2017).
[28] Benilda Eleonor V. Commendador, Darwin Joseph L. Dela Cruz, Nathaniel C. Mercado, Ria A. Sagum, Diana C. Santiago, Sharlaine Grace C. Tagnines, Dialogue Transcription using Gaussian Mixture Model in Speaker Diarization, International Proceedings of Economics Development and Research, 1(2013), No.63, 1-5.
[29] Shreya Banerjee, Ankit Choudhary, Somnath Pal, Empirical Evaluation of K-Means, Bisecting KMeans, Fuzzy C-Means and Genetic K-Means Clustering Algorithms, 2015 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), December 2015.
[30] Jain A., Data clustering: 50 years beyond k-means, Pattern Recognition letters, 8(2010), No. 31, 651-666.
[31] J. B. MacQueen, Some methods for classification and analysis of multivariate observations, In Proceedings of 5th Berkley Symposium on Mathematical Statistics and Probability, Vol. I: Statistics, (1967), 281–297.
[32] Than Than Htay, Su Su Maung and Khine Thin Zar, Analysis on Results Comparison of Feature Extraction Methods for Breast Cancer Classification, International Journal of Advances in Scientific Research and Engineering (ijasre), 6(2020), No. 3, 95-102.
[33] Vijay Kumar, Priyanka Gupta, Importance of statistical measures in digital image processing, International Journal of Emerging Technology and Advanced Engineering, 8(2012), No.2, 56-62.
[34] Haralick R.M., Shanmugam K., Dinstein I., Textural features for image classification, IEEE Transactions on Systems, Man and Cybernetics, (1973), No.3, 610-621.
[35] Cheng H.D., Shan J., Ju W., Guo Y., Zhang L., Automated breast cancer detection and classification using ultrasound images: A survey, Pattern recognition, 1(2010), No.43, 299-317.
[36] Mitiche, Imene; Nesbitt, Alan; Conner, Stephen; Boreham, Philip; Morison, Gordon, 1D-CNN based real-time fault detection system for power asset diagnostics, IET Generation, Transmission & Distribution, 24(2020), No.14, 5766-5773.
[37] Abdeljaber, O., Avci, O., Kiranyaz, M.S., Boashash, B., Sodano, H., Inman, D.J., 1d-cnns for structural damage detection: Verification on a structural health monitoring benchmark data, Neurocomputing, (2018), No.275, 1308-1317.
[38] Woon, W.L., Aung, Z., El.Hag, A, Intelligent monitoring of transformer insulation using convolutional neural networks, 6th Data Analytics for Renewable Energy Integration. Technologies, Systems and Society (DARE), Dublin, Ireland, (2018), 127-136.
[39] Nair, V., E. Hinton, G., Rectified linear units improve restricted boltzmann machines, 27th International Conference on Machine Learning (ICML), Haifa, Israel, (2010), 807–814.
[40] Ioffe, S., Szegedy, C., Batch normalization: accelerating deep network training by reducing internal covariate shift, 32nd International Conference on Machine Learning (ICML), Lille, France, (2015), 448–456.
[41] Has¸im Sak, Andrew Senior, Franc¸oise Beaufays, Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling, (2014), 1-5.
[42] Gers F.A., Schmidhuber J., Cummins F., Learning to forget: Continual prediction with LSTM, Neural Computation, 10(2000), No.12, 2451-2471.
[43] Gers F.A., Schraudolph N.N., Schmidhuber J., Learning precise timing with LSTM recurrent networks, Journal of Machine Learning Research, (2003), No.3, 115-143.
[44] Seyedali Mirjalili, Seyed Mohammad Mirjalili, Andrew Lewis, Grey Wolf Optimizer, Advances in Engineering Software, (2014), No.69, 46-61.
[45] Pedersen M.E.H., Chipperfield A.J., Simplifying Particle Swarm Optimization, Applied Soft Computing, 2(2010), No.2, 618-628.
[46] Seyedali Mirjalili, Andrew Lewis, The Whale Optimization Algorithm, Advances in Engineering Software, (2016), No.95, 51-67.
[47] John Mc Call, Genetic algorithms for modelling and optimisation, Journal of Computational and Applied Mathematics, 1(2005), No.184, 205-222.
[48] Karaboga D., Basturk B., On the performance of artificial bee colony (ABC) algorithm, Applied Soft Computing, 1(2008), No.8, 687-697.
[49] Jayaseelan, Samuel Manoharan, Gopal, Sakthivel Thirumalai, Muthu Sangeetha, Selvaraju, Sivamani, Patel, Saad, A Hybrid Fuzzy based Cross Neighbor Filtering (HF-CNF) for Image Enhancement of fine and coarse powder Scanned Electron Microscopy (SEM) images, 6(2022), No.42, 6159-6169.