BBRC
Bioscience Biotechnology
Research Communications

Classification of Breast Cancer in Histopathology Image using Modified Ant Lion Optimizer and Capsule Network Architecture

Shwetha G.K [1], K R Udaya Kumar Reddy[2]
Department of Computer Science & Engineering, NMAM Institute of Technology, Nitte
Visvesvarya Technological University, Belagavi, Karnataka State, India

ABSTRACT

The exact recognition of breast cancer disease utilizing histology pictures is a difficult assignment, because of the variety of generous tissue and heterogeneity of cell development. In this exploration, a proper component choice and classification methods are proposed for programmed bosom malignancy discovery and characterization. At first, the cores and non-cores cells are portioned from the histological pictures by utilizing Fuzzy C Means (FCM) grouping algorithm. At that point, the component vectors from the sectioned cores and non-cores cells are separated by Speeded up Robust Features (SURF) and shading second highlights. Moreover, Modified Ant Lion Optimizer (MALO) calculation is used to choose the dynamic or ideal element vectors from the removed highlights. In MALO calculation, competition choice system is utilized to choose the people from the irregular populace to maximize the assembly rate that assists in accomplishing better characterization. At long last, the Capsule Network Architecture (CNA) is used to group the breast cancer disease as benign or malignant. The BreaKHis and Stanford Tissue Microarray Dataset (TMAD) are utilized to research the suggested model presentation. The division and order execution of the suggested model is assessed by methods for exactness, review, f-score, accuracy, jaccard and dice coefficient. In the trial segment, the suggested model enhanced least of 0.17% and limit of 8.04% of exactness in BreaKHis, and TMAD identified with the current models.

KEY WORDS: BREAST CANCER DETECTION, CAPSULE NETWORK ARCHITECTURE, HISTOPATHOLOGICAL IMAGING, IMAGE NORMALIZATION, MODIFIED ANT LION OPTIMIZER, SPEEDED UP ROBUST FEATURES.

INTRODUCTION

In recent decades, the most common cancer is the breast cancer among females in India, where out of 100 individuals, around 30% of the individuals are affected from breast cancer (Parameshachari et al. 2020; [2] Prabu et al. 2019). In the developing countries, the incidence of breast cancer is increased highly, due to the change in life expectancy, adoption of western lifestyles, and increase in urbanization (Kumar, et al. 2020; Yan et al. 2020; Budak et al. 2019). Hence, the early diagnosis of breast cancer significantly diminishes the mortality rate worldwide. Where the most commonly utilized imaging techniques for accurate breast cancer detection and classification are mammogram, computerized tomography, diffusion tensor imaging, magnetic resonance imaging, histology, positron emission tomography, etc. (Sudharshan et al. 2019; Dordea et al. 2013). Among these imaging techniques, histology images plays a crucial role in breast cancer detection, because it delivers a more comprehensive view of normal and abnormal tissues and also it effectively preservers the underlying tissues (Wahab et al. 2017; Li et al. 2019).

A few existing approaches used in histopathological breast cancer detection and classification are Support

289

Vector Machine (Spanhol 2019), cascaded method with multi-SVM (Wan et al. 2015), Deep Convolutional Neural Network (DCNN) (Gecer et al. 2018), CNN (Araújo et al. 2017; Dordea et al. 2013; Komura & Ishikawa 2018), etc. Due to the complexity of histological photos and the drastic workload in terms of makes the segmentation and classification tasks time consuming, where the results are subjected to pathologist subjectivity. In order to focus this concern, a new model for component choice and classification methods are proposed for programmed bosom malignancy discovery and characterization.

In this research article, a new semi-supervised model is proposed for enhancing the detection of breast cancer and classification using histopathological images.The histopathological images initially were obtained from two online databases such as BreaKHis and TMAD. Then, the image normalization method was used for improving the visibility of both nuclei and non-nuclei cells in the collected histopathological images. The normalization method not only diminishes the noise but also brings the range of pixel intensity values to the normal distribution. In addition, the nuclei and non-nuclei cells were segmented from the pre-processed pictures using FCM approach. In this approach, the data points were exclusively belongs to one cluster center, so it gives better results for overlapped datasets compared to other clustering approaches.

Then, feature extraction was accomplished by using SURF and color moments (skewness, standard deviation and mean) to extract the features from the segmented nuclei and non-nuclei cells. The undertaken feature extraction techniques were extremely robust to noise and invariant to geometric and photometric transformations that helps in extracting the relevant features. Then, the optimal feature vectors were selected by utilizing MALO technique. In MALO, tournament selection scheme was used to choose the people from the random population that significantly improves the pace of convergence for better classification. The selected optimal feature vectors were classified by using CNA classifier that was robust in unstructured data conditions. In the resulting segment, the proposed model efficacy was investigated by means of accuracy, recall, f-score, precision, jaccard and dice coefficient.

This research article is organized as below. A couple of recent research publications on breast cancer detection is surveyed in the Section 2. Section 3 details about the undertaken methodologies for breast cancer detection with mathematical expression. The experimental analysis of the proposed model is indicated in the Section 4. Section 5 details about the conclusion of the research work.

Literature Review: (Zheng et al. 2017) implemented a new deep learning network (BreastNet) for breast cancer recognition on the basis ofConvolutional Neural Network (CNN). Initially, augmentation approaches were utilized to change the features of the collected images such as shift, brightness, flip and rotation. Then, select the important key regions from the histopathological images using hyper-column technique. The developed BreastNet consists of four blocks such as convolutional, residual, dense and pooling blocks. In this research study, the histopathology breast images were acquired from BreakHis dataset. In the experimental section, the developed network performance was compared with VGG-19, AlexNet and VGG-16 models on BreakHis database.

Experimental result demonstrates that the developed network achieved better performance in light of accuracy. Hence, the developed BreastNet model contains only high level layers, where it was not suitable for low resolution histological images. S. Reis, et al, (Reis et al. 2017) implemented a new model for breast cancer recognition on the basis of texture level feature extraction. At first, local binary pattern and multi-scale basic image features were utilized for extracting the feature values from the collected histopathological images. Then, random decision tree classification technique was applied to classify the breast pictures as benign and malignant. This research shows the ability of texture based image analysis for classifying the breast cancer on H&E stained slides. From the experimental investigation, the developed system showed significant performance in breast cancer recognition compared to prior methods in light of recall, precision, f-score and area under curve.

Since, the random decision tree classifier was not applicable for multiclass classification and it is only suitable for binary classification that was considered as a major concern in this research study. (Saha et al. 2018) developed a new supervised model for detecting mitosis from breast histopathological images. In this research, the deep learning model was developed on the basis of hand-crafted features, which were issued from the prior medical challenges; AMIDA-13, MITOS @ ICPR 2012, and the project expertise (MICO ANR TecSan). In this study, the developed deep learning network comprises of 2 fully connected layers, 5 convolution layers, 4 Rectified Linear Units (ReLU) and 4 max pooling layers. After every convolution layer, ReLU was utilized as an activation function. After 1st fully connected layer, drop-out layer was included in order to avoid over-fitting. Usually, the handcrafted features comprises of intensity, textural, and morphological features. The ReLU activation function was active only when the units were zero and positive, or else it leads to bias shift and dead neurons, which was a major concern in the developed model.

(Saha et al. 2018) presented a new feature extraction system on the basis of CNN for histopathological breast cancer detection. Initially, the nuclei were detected from the histological images and then CNN with 3 hierarchy structures were used to train the images. In addition, the image level features (spatial and pattern distribution) of the nuclei were extracted by the trained network. In the experimental segment, the developed features were investigated on a histological image dataset of

breast lesion. The experimental consequences shows that the developed features significantly denotes the histopathological images and the developed system attains an effective classification performance related to the prior methods. By using only texture level features, the semantic space between the extracted features were high, which may leads to misclassification.

(Khosravi et al. 2018) used CNN for identifying the breast cancer related markers, cancer tissues, sub-types and their staining scores. Additionally, the developed CNN discriminates five biomarkers of breast cancer, two subtypes of lung cancer, and four biomarkers of bladder cancer. In the clinical conditions, the deep learning based approaches delivers precise status assessments. Generally, the accuracy of CNN depends on the complexity and size of the architecture and noise occurred in the database. In this research, the histopathological breast images were acquired from TMAD. Experimental outcome demonstrates that the developed network achieved better performance in breast cancer detection by means of accuracy, area under curve, recall, precision, Cohen's kappa and jaccard coefficient. As previously mentioned, CNN comprises of only high level layers, so it was not adaptable for low resolution images. In order to overcome these concerns, in this research paper for automatic breast cancer detection , a new frame-work is suggested using histopathological images.

Proposed Model: In the medical field, histopathological breast cancer detection and classification gained more attention among the researchers, because breast cancer is the leading cause of mortality among women worldwide related to other cancer types. In this article, the proposed model comprises of six phases for breast cancer detection and classification such as image collection: BreaKHis and TMAD, image preprocessing: normalization, segmentation: FCM, extraction of features: SURF and color moments, optimization of features: MALO and classification: CNA. The flow diagram of the model proposed is stated in figure 1.

1Image Collection and Pre-Processing: In this study, the input histopathological images are collected from TMAD and BreaKHis datasets. TMAD consists of 205,161 histological images that archives 350 probes on 1490 tissue microarray slides. Among these 205,161 images, 31,300 histological images are available online. For breast cancer, TMAD database comprises of 67 benign images and 278 malignant histology images. In addition, BreaKHis dataset consists of 7909 image samples for malignant and benign classes. Hence, the benign class contains 2440 image samples and the malignant class consists of 5429 image samples.

In BreaKHis database, the histology images are collected from 82 subjects using different magnifying factors like 400×,200×,100×and 40×. After image collection, normalization technique is utilized to find the variations and deformations occurred in the images that helps in enhancing the collected image quality. The formula to estimate normalization in the histological images is mathematically represented in equation (1).

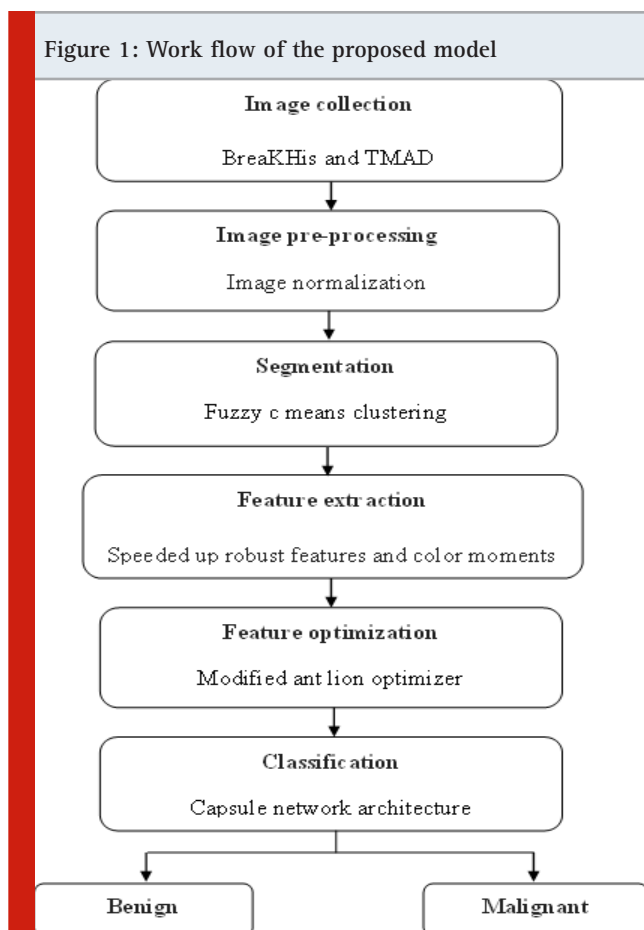Figure 1: Work flow of the proposed model

Image collection

BreaKHis and TMAD

Image pre-processing

Image normalization

Segmentation

Fuzzy c means clustering

Feature extraction

Speeded up robust features and color moments

Feature optimization

Modified ant lion optimizer

Classification

Capsule network architecture

Benign   Malignant

Figure 2: Sample histological image, a) malignant class, and b) benign class
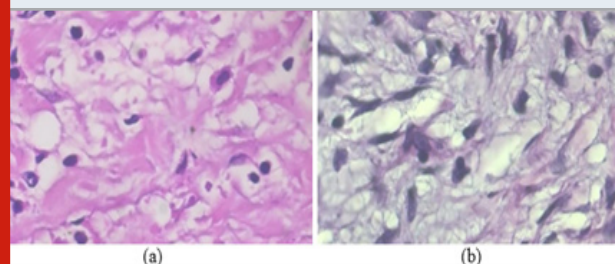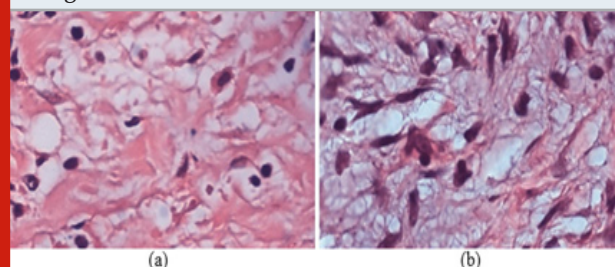
(a)   (b)

Figure 3: Normalized image, a) malignant class, and b) benign class

(a)   (b)

$$IN = (I - Min) + \frac{newMax - newMin}{Max - Min} + newMin \qquad (1)$$

Where, normalized image is indicated as IN, collected histological images are denoted as I, the normalized image pixel intensity range is specified as newMax-newMin, and intensity range for pixels of collected histological images areMin=0,and Max=255. Sample collected and normalized breast histological images of malignant and benign are indicated in the figures 2 and 3.

2 Image Segmentation: After normalizing the histological images, segmentation is carried-out for segmenting the nuclei and non-nuclei cells. Initially, select the clusters and then assign the coefficients randomly to each data points. Further, calculate the centroid for every cluster and its coefficients of being in the clusters.Let IN be a normalized image that comprises of a set of x_i grey scale images at pixel i (i=1,2,...,N) and X={$x_1,x_2,x_3...x_N$}$\subset R^k$ respectively in the k-dimensional area with the cluster centers v={$v_1,v_2,v_3...v_c$}, where cis stated as positive integer (2<c<<N). By assigning a different membership value to all pixels, the clusters created in the image space are combined in the FCM algorithm. Therefore, the objective function of FCM is mathematically written in equation (2).

$$FCM = \sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^{m} \|x_i - v_j\|^2 \qquad (2)$$

Where, m is indicated as weight exponent,u_{ij} is denoted as value of membership for each pixel i in jth cluster(j=1,2,...c), N is represented as data points, m>1 and $\|x_i - v_j\|^2$ be the gray scale Euclidean distance between i andv_j that is mathematically stated in equation (3).

$$\sum_{j=1}^{c} u_{ij} = 1, u_{ij} \in [0,1], 0 \leq \sum_{i=1}^{N} u_{ij} \leq N \qquad (3)$$

However, the centers of clusters are iteratively updated by utilizing the membership function as stated in the equations (4) and (5). In this scenario, the accuracy a is measured by membership function from one iteration k to the next iteration k+1 that is determined using equation (6).

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} (\|x_i - v_j\|^2 / \|x_i - v_k\|^2)^{1/(m-1)}} \qquad (4)$$

Figure 4: Segmented image, a) malignant class, and b) benign class

(a)　　　　(b)
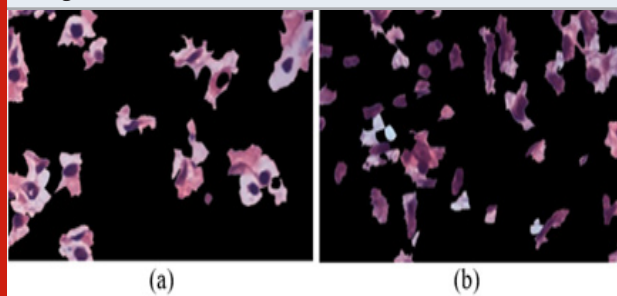
$$v_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} x_i}{\sum_{i=1}^{N} u_{ij}^{m}} \qquad (5)$$

$$a = \Delta_i^{N} \Delta_i^{c} |u_{ij}^{k+1} - u_{ij}^{k}| \qquad (6)$$

Where, Δ is indicated as largest vector value, u_{ij}^{k+1} and u_{ij}^{k} are stated as degree of membership of iterations k+1 and k. After segmentation, feature extraction is employed to extract the feature vectors from the partitioned nuclei and non-nuclei cells. Segmented image is graphically denoted in figure 4.

3 Feature Extraction and Optimization: After segmentation, feature vectors are calculated from partitioned nuclei and non-nuclei cells by using SURF (Iesmantas et al. 2018) and color moments (Bay et al. 2008). By combining the low and high descriptors for level characteristics, the semantic space between the feature subsets are decreased that helps in attaining batter classification performance. After extracting the feature vectors, optimization is carried-out by utilizing MALO algorithm. Hence, ALO is a multi-objective optimizer that delivers multiple solution to optimize the issues related to "curse of dimensionality". The ALO algorithm shows the action of ant lions and also it is a algorithm based on population that provides solution for approximation to the concern with a random collection of remedies. The working process of MALO algorithm is given below.

Step 1: At first, ant is assigned with random walk.

Step 2: For each iteration, fitness of the ant is analysed by utilizing the objective function.

Step 3: Ant walks in the search space utilizing random walks.

Step 4: In 1st iteration, ant-lion location is considered as an ant position, and the ant lion position is changed on the basis of ant movement.

Step 5: An ant lion is allocated for each ant. Where, the ant lion position is updated if the ant rate is reduced.

Step 6: Usually, ant lions are elites which have impact on the ant movements in all directions.

Step 7: The ant lions are replaced with elite if it provides better influence on the movement of an ant.

Step 8: Repeat the steps 2 to 7, until the algorithm attains a satisfactory result.

Step 9: The elite ant lions fitness value and position provides better estimation.

The main aim of MALO algorithm is to find and investigate the ant's position. Initially, the ant walks in the search space utilizing random walks, which is

292

assigned using equation (7).

$$x^t_i = \frac{(x^t_i - k_i) \times (n^t_i - m^t_i)}{(l_i - k_i)} + m^t_i \qquad (7)$$

Where, t is indicated as random walk, m^t_i is denoted as minimum of ith variable at t^{th} iteration, n^t_i is denoted as maximum of ith variable, k^t_i is indicated as minimum random walk and l^t_i is represented as maximum random walk. After random walk, catch the ant by using equation (8).

$$AL^t_j = A^t_i, \quad If\ f(A^t_i) < f(AL^t_j) \qquad (8)$$

Where, AL is indicated as ant-lions and A is denoted as ants, j is stated as selected ant-lions, and i is the given position.

The final step in MALO is elitist, where the fitness of ant-lion is selected and stored. The random walk in the ant-lions gravity is towards the selected ant lion and the elite ant-lion using tournament selection methodology. By using equation (9), select the corresponding ant-lion.

$$A^t_i = \frac{R^t_A + R^t_E}{2} \qquad (9)$$

Where, R is stated as ant-lion random walk and E is denoted as elite random walk at i^th iteration. For solving the multi-objective issue in crowd-sourcing using MALO algorithm, it has to modify based on the equation (10).

$$AL^t_j = A^t_i, If\ f(A^t_i) < f(AL^t_j) \qquad (10)$$

Where, $\underline{t}$ is indicated as iteration, AL^t_j is stated as ant-lion in i^{th} position at t^{th} iteration.

Table 1. Parameter setting of the developed architecture

	Type of layer	Capsules/ maps and neurons	Dimension of the capsule
0	Input	3M×512N×512N	1×1
1	Convolutional	64M×255N×255N	4×4/2
2	Convolutional	128M×126N×126N	4×4/2
3	Convolutional	256M×61N×61N	6×6/2
4	Convolutional	256M×28N×28N	6×6/2
5	Convolutional	256M×11N×11N	8×8/2
6	Capsule layer	3872 C	8
7	CancerCaps layer	4C	16

4 Classification: After selecting the optimal features, classification is carried out by using CNA. Usually, CNN has several conceptual concerns; (i) it does not considered spatial relationship between the simpler objects, and (ii) the process of max pooling rejects details pertaining to the position of a few entities that the network tries to identify. To address these problems, a concept of capsule is presented in CNN classifier, where capsule is a group of neurons whose output is interpreted as the property of the similar object. Every capsule includes two components; activation probability and pose matrix. The output vector length of a capsule is interpreted as the probability, whose entities are denoted as the capsule in the current input. The parameter setting of the developed architecture is indicated in table 1.

In primary capsule layer, each capsule is connected with CancerCaps layer that enables better learning compared to max-pooling routing (Iesmantas &Alzbutas 2018). In regular neural networks, squashing function is a multi-dimensional activation function that is mathematically denoted in equation (11).

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2}\frac{}{\|s_j\|} \qquad (11)$$

Where, v_j is indicated as vector output of capsule and s_j is stated as total input. The architecture of CNA is given figure 5.

Experimental Analysis: In this research study, Python programming environment was used for experimental simulation with 2 TB memory, windows 10 operating system, 64 GB RAM, i7 3.0 GHz processor and 8 GB GPU. Additionally, the proposed model performance was related with a few benchmark models like BreastNet (Togaçar et al. 2020) and CNN (Khosravi et al. 2018) in order to evaluate the advantage of proposed model over the benchmark models. In this research, the proposed model performance was investigated on two datasets like BreaKHis and TMAD. Meanwhile, the proposed model performance was evaluated by the performance metrics like the coefficient of precision, recall, accuracy, jaccard and dice. Hence, the mathematical representation of the undertaken performance metrics is denoted in the equations (12-17).

$$Accuracy = \frac{TP+TN}{FN+TP+FP+TN} \times 100 \qquad (12)$$

$$Recall = \frac{TP}{FN+TP} \times 100 \qquad (13)$$

$$F-score = \frac{2TP}{2TP+FN+TP} \times 1 \qquad (14)$$

$$Precision = \frac{TP}{FP+TP} \times 100 \qquad (15)$$

$$Jaccard\ coefficient = \frac{TP}{FP+TP+FN} \times 100 \qquad (16)$$

$$Dice\ coefficient = \frac{2TP}{2TP+FP+FN} \times 100 \qquad (17)$$

Figure 5: Architecture of CAN

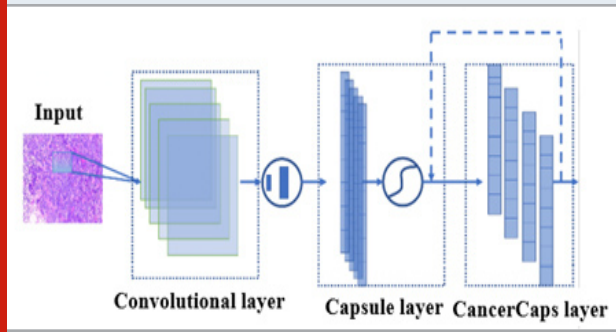Input — Convolutional layer — Capsule layer — CancerCaps layer

Table 2. Performance investigation of the proposed model in light of jaccard and dice coefficient on BreaKHis database

Classifiers	Optimizers	Jaccard coefficient (%)	Dice coefficient (%)
RNN	ALO	82.39	77.51
DNN		90.81	87.07
CNA		92.03	90.98
RNN	MALO	88.05	81.83
DNN		93.28	90.83
CNA		97.78	96.93

In this dataset, the CNA classifier attained 97.78% of jaccard coefficient and 96.93% of dice coefficient, which is higher related to other techniques; RNN and DNN. Figure 6 represents the graphical valuation of proposed model in light of jaccard and dice coefficient on BreaKHis database.

In table 3, the classification efficiency of the model proposed is investigated by means of recall, precision, f-score and accuracy. By investigating table 3, the CNA achieved 98.97% of accuracy in breast cancer classification that showed maximum of 4.77% and

Where, false negative is indicated as FN, false positive is stated as FP, true negative is denoted as TN, and true positive is represented as TP.

1 Quantitative Analysis on BreaKHis Dataset: In this segment, BreaKHis dataset is undertaken to analyze the proposed model performance with dissimilar classification techniques like Recurrent Neural Network (RNN) and Deep Neural Network (DNN) and optimization techniques like ALO and MALO by means of the coefficient of precision, recall, accuracy, jaccard and dice. In the table 2, the proposed model performance is investigated by means of jaccard and dice coefficient. Here, the performance analysis is done for 100 histopathological images (50 samples for malignant class and 50 samples for benign class). Though, the CNA classifier achieved superior performance in breast detection and classification related other classification techniques with MALO algorithm.

Figure 6: Graphical evaluation of the proposed model in light of jaccard and dice coefficient on BreaKHis Database.
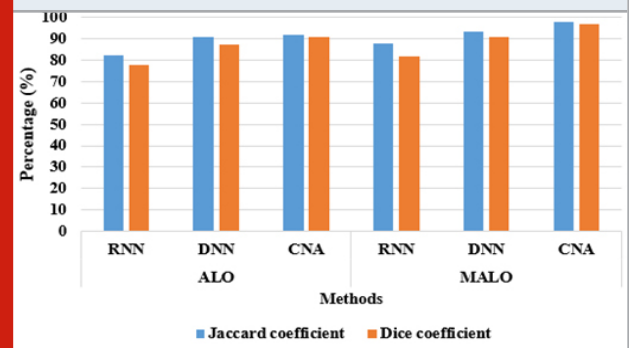
minimum of 4.05% improvement in accuracy related to the comparative techniques like RNN and DNN. In addition, recall, precision and f-score of the proposed model is significant compared to other classification techniques, especially using MALO algorithm. The graphical valuation of the proposed model in light of recall, precision, f-score and accuracy on BreaKHis database is represented in figure 7.

Table 3. Performance investigation of the model proposed in light of recall, precision, f-score and accuracy on BreaKHis database.

Classifiers	Optimizers	Precision (%)	Recall (%)	F-score (%)	Accuracy (%)
RNN	ALO	89	84.04	89.033	90.82
DNN		92.43	89.89	87.03	91.18
CNA		94.03	90.09	91.04	93.01
RNN	MALO	90.02	90.90	91.92	94.20
DNN		94.33	93.83	96.04	94.92
CNA		97	96.89	98.63	98.97

2 Quantitative Analysis on TMAD: In this section, TMAD is used to analyze the proposed model performance with dissimilar classification and optimization techniques in light of accuracy, recall, f-score, precision, jaccard and dice coefficient. In table 4, the proposed model performance is analyzed by means of jaccard and dice coefficient on TMAD. In this scenario, the performance investigation is accomplished for 50 histological images

(25 samples for malignant class and 25 samples for benign class). By analyzing table 4, the CNA classifier attained good performance in breast detection related

other classification and optimization techniques on TMAD. Figure 8 represents the graphical valuation of the proposed model in light of jaccard and dice coefficient on TMAD.

Figure 7: Graphical evaluation of the model proposed in light recall, precision, f-score and accuracy on BreaKHis database.
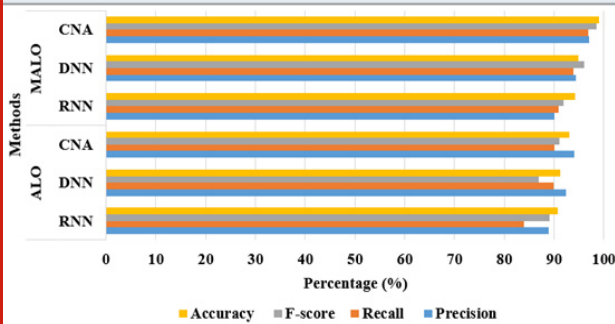

Figure 8: Graphical evaluation of the model proposed in light of jaccard and dice coefficient on TMAD
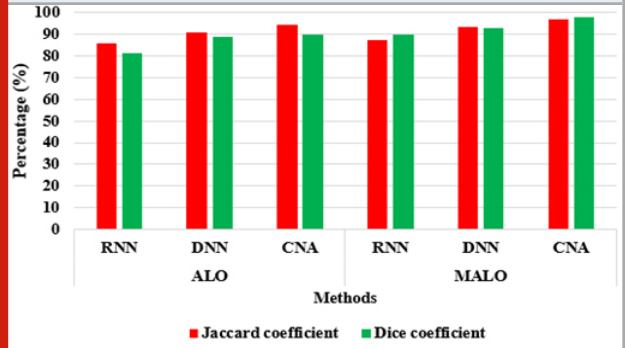
Table 4. Performance investigation of the model proposed in light of jaccard and dice coefficient on TMAD

Classifiers	Optimizers	Jaccard coefficient (%)	Dice coefficient (%)
RNN	ALO	85.55	81.10
DNN		90.82	88.89
CNA		94.31	90.04
RNN	MALO	87.09	89.82
DNN		93.38	92.97
CNA		96.92	97.94

Figure 9: Graphical evaluation of the model proposed in light recall, precision, f-score and accuracy on TMAD
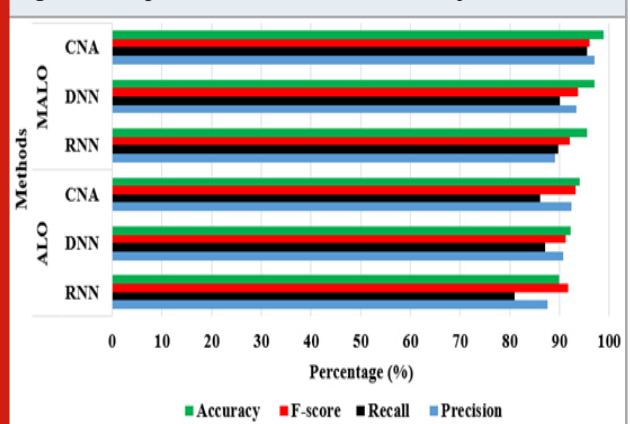
Table 5. Performance investigation of the model proposed in light of recall, precision, f-score and accuracy on TMAD.

Classifiers	Optimizers	Precision (%)	Recall (%)	F-score (%)	Accuracy (%)
RNN	ALO	87.62	80.90	91.73	90
DNN		90.72	87.04	91.20	92.30
CNA		92.37	86.09	93.302	94
RNN	MALO	89.09	89.83	92.08	95.60
DNN		93.45	90.05	93.72	97
CNA		97.08	95.52	96	98.84

In table 5, the proposed model performance is investigated on TMAD database in light of recall, f-score, precision and accuracy. From the experimental study, the classification accuracy of CNA is 98.84%, and the existing techniques (RNN and DNN) achieves 95.60% and 97%. In addition, the recall, precision, and f-score of CNA is 95.52%, 97.08%, and 96%. In contrast, the comparative classification techniques (RNN and DNN) achieves minimum recall, precision, and f-score

compared to CNA classifier. Graphical valuation of the proposed model in light of recall, precision, f-score and accuracy on TMAD is represented in figure 9.

3 Comparative Investigation: Table 6 states the comparative study of the proposed and existing models. M. Togaçar, et al, [15] developed a new deep learning network (BreastNet) for breast cancer recognition on the basis of CNN. In this study, the histology breast

images were acquired from BreakHis dataset. In the experimental phase, the developed network performance was related with VGG-19, AlexNet and VGG-16 models on BreakHis dataset. Experimental result shows that the developed network attained 98.80% of accuracy in classification. In addition, P. Khosravi, et al, [19] used CNN for identifying the breast cancer related markers, cancer tissues, sub-types and their staining scores. In this article, the histological breast images were acquired from TMAD. Experimental outcome shows that the developed model attained 90.80% of accuracy in breast cancer detection. Related to these articles, the proposed model showed good result in breast cancer detection. In this work, feature optimization plays a vital role in breast cancer detection, which significantly reduces the "curse of dimensionality" issue. The major advantage of using MALO algorithm is to diminish the input entities and also to calculate the more useful feature vectors from the segmented image regions that helps in attaining better performance in breast cancer classification.

Table 6. Comparative investigation

Methodology	Dataset	Accuracy (%)
BreastNet (Tosaçar et al. 2020)	BreakHis	98.80
CNN (Khosravi et al. 2018)	TMAD	90.80
Proposed model	BreakHis	98.97
	TMAD	98.84

CONCLUSION

In this article, a new feature optimization technique is proposed with CNA classifier for enhancing the breast detection and classification performance. Initially, FCM methodology is undertaken for segmenting the nuclei and non-nuclei cells from the images. Then, the optimal features are selected by applying MALO algorithm after feature extraction. The corresponding selected feature values are classified as benign or malignant classes by using CNA classifier. In the resulting phase, the proposed model performance is validated in light of accuracy, recall, f-score, precision, jaccard and dice coefficient. In breast cancer recognition, the model proposed enhanced minimum of 0.17% and maximum of 8.04% of accuracy in BreaKHis, and TMAD related to the current versions of models (BreastNet and CNN). In future work, a novel optimization based segmentation technique can be included in the proposed model for further improving the breast cancer classification performance in the histopathological images.

REFERENCES

Parameshachari, B.D., Panduranga, H.T. and liberata Ullo, S., 2020, September. Analysis and Computation of Encryption Technique to Enhance Security of Medical Images. In IOP Conference Series: Materials Science and Engineering (Vol. 925, No. 1, p. 012028). IOP Publishing.

Prabu, S., Balamurugan, V. and Vengatesan, K., 2019. Design of cognitive image filters for suppression of noise level in medical images. Measurement, 141, pp.296-301.

Kumar, S.K. Singh, S. Saxena, K. Lakshmanan, A.K. Sangaiah, H. Chauhan, S. Shrivastava, and R.K. Singh, "Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer", Information Sciences, vol.508, pp.405-421, 2020.

R. Yan, F. Ren, Z. Wang, L. Wang, T. Zhang, Y. Liu, X. Rao, C. Zheng, and F. Zhang, "Breast cancer histopathological image classification using a hybrid deep neural network", Methods, vol.173, pp.52-60, 2020.

Ü. Budak, Z. Cömert, Z.N. Rashid, A. Sengür, and M. Çıbuk, "Computer-aided diagnosis system combining FCN and Bi-LSTM model for efficient breast cancer detection from histopathological images", Applied Soft Computing, vol.85, pp.105765, 2019.

P.J. Sudharshan, C. Petitjean, F. Spanhol, L.E. Oliveira, L. Heutte, and P. Honeine, "Multiple instance learning for histopathological breast cancer image classification", Expert Systems with Applications, vol.117, pp.103-111, 2019.

M. Dordea, H. Colvin, P. Cox, A.P. Nicolas, V. Kanakala, and O. Iwuchukwu, "Clinical and histopathological factors affecting failed sentinel node localization in axillary staging for breast cancer", the surgeon, vol.11, no.2, pp.63-66, 2013.

N. Wahab, A. Khan, and Y.S. Lee, "Two-phase deep convolutional neural network for reducing class skewness in histopathological images based breast cancer detection", Computers in biology and medicine, vol.85, pp.86-97, 2017.

X. Li, M. Radulovic, K. Kanjer, and K.N. Plataniotis, "Discriminative pattern mining for breast cancer histopathology image classification via fully convolutional autoencoder", IEEE Access, vol.7, pp.36433-36445, 2019.

F.A. Spanhol, L.S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification", IEEE Transactions on Biomedical Engineering, vol. 63, no.7, pp.1455-1462, 2015.

T. Wan, J. Cao, J. Chen, and Z. Qin, "Automated grading of breast cancer histopathology using cascaded ensemble with combination of multi-level image features", Neurocomputing, vol. 229, pp.34-44, 2017.

B. Gecer, S. Aksoy, E. Mercan, L.G. Shapiro, D.L. Weaver, and J.G. Elmore, "Detection and classification of cancer in whole slide breast histopathology images using deep

convolutional networks", Pattern recognition, vol.84, pp.345-356, 2018.

T. Araújo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, C. Eloy, A. Polónia, and A. Campilho, "Classification of breast cancer histology images using convolutional neural networks", PloS one, vol.12, no.6, 2017.

D. Komura, and S. Ishikawa, "Machine learning methods for histopathological image analysis", Computational and structural biotechnology journal, vol.16, pp.34-42, 2018.

M. Togaçar, K.B. Özkurt, B. Ergen, and Z. Cömert, "BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer", Physica A: Statistical Mechanics and its Applications, vol.545, pp.123592, 2020.

S. Reis, P. Gazinska, J.H. Hipwell, T. Mertzanidou, K. Naidoo, N. Williams, S. Pinder, and D.J. Hawkes, "Automated classification of breast cancer stroma maturity from histological images", IEEE Transactions on Biomedical Engineering, vol.64, no.10, pp.2344-2352, 2017.

M. Saha, C. Chakraborty, and D. Racoceanu, "Efficient deep learning model for mitosis detection using breast histopathology images", Computerized Medical Imaging and Graphics, vol.64, pp.29-40, 2018.

Y. Zheng, Z. Jiang, F. Xie, H. Zhang, Y. Ma, H. Shi, and Y. Zhao, "Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classification", Pattern Recognition, vol.71, pp.14-25, 2017.

P. Khosravi, E. Kazemi, M. Imielinski, O. Elemento, and I. Hajirasouliha, "Deep convolutional neural networks enable discrimination of heterogeneous digital pathology images", EBioMedicine, vol.27, pp.317-328, 2018.

H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)", Computer vision and image understanding, vol.110, no.3, pp.346-359, 2008.

N. Keen, "Color moments", School Of Informatics, University Of Edinburgh, pp.3-6, 2005.

T. Iesmantas, and R. Alzbutas, "Convolutional capsule network for classification of breast cancer histology images" In International Conference Image Analysis and Recognition, Springer, Cham, pp. 853-860, 2018.