Enhanced Firefly Optimization Based Classifier to Diagnose Multimodality Breast Cancer Images

https://doi.org/10.3991/ijoe.v19i06.37891

Anupama Y.K.[1(✉)], Amutha S.[1], Ramesh Babu D.R.[1], S. Sathish Kumar[2]
[1]Dayananda Sagar College of Engineering, Affiliated to VTU University, Bengaluru, Karnataka, India
[2]RNS Institute of Technology, Affiliated to VTU University, Bengaluru, Karnataka, India
anupama-cs@dayanandasagar.edu

Abstract—Breast cancer is one of the most affecting carcinoma for women from long time. Early detection is necessary to increase the lifespan of patients. In this study deep learning and machine learning approaches are applied to histopathological, mammogram and ultrasound breast cancer modalities. In-order to increase the efficacy of diagnosis of these modalities. Study has been carried out in majorly four phases. First phase involved collection of the datasets of all the three modalities mentioned earlier. Second phase consists of extracting relevant features using ResNet-18. Third phase involves feeding the extracted information to enhanced firefly or to the existing optimization techniques. Fourth phase consists of considering selected features as input to the classifiers. Then enhanced firefly based classifier compared with the existing ant colony and genetic algorithm based classifier. Enhanced firefly based classifier displays better results compared to the state of art approaches.

Keywords—ResNet-18, firefly optimization, ant colony, genetic, modalities

1 Introduction

Breast cancer is one of the frequently affecting cancers among women in all over the world. Types of breast cancer are invasive and non-invasive. In the invasive breast cancer from the affected area of the breast it will spread to the distant parts of the body system. In the non-invasive it can be of affected breast mass, but it will not spread to the other parts of the body system. Sometimes dense non-vulnerable breast mass may exist. The main factors contribute to this carcinoma are menopause and life style. Due to these reasons, early diagnosis and prognosis is necessary to prolong the life of patients. For detecting carcinoma various screening procedures like Mammogram, ultrasound, tomography, etc… can be used.

Machine learning technique (MLT) is a strategy to extract the valuable data from the given data. MLT has procedures to remove the irrelevant data, optimizing of selected data, image classification and displaying results. MLT's can be categorized into known label techniques and unknown label techniques. In the known label techniques data provided has known input and expected output. Examples for known label techniques are simple linear regression, Logistic regression, etc… Whereas, in the unknown labels techniques the data provided has only known input. Examples for unknown label techniques are clustering, SVM, etc… MLT's can be applied on the histopathological, mammogram and ultrasound images. These images undergo through the MLT processes. End result obtained is diagnosed patient details. Details are the predictions obtained by MLT's. These predictions significantly help the radiologists in medical prominent decisions.

2 Related work

Related work has been carried out to explore the different research articles with respect to applying machine learning/deep learning strategies on mammogram, ultrasound and histopathological breast cancer image modalities. Along with this optimization techniques and other research articles are also explored.

2.1 Mammogram research article

Y. Shen, N. Wu, J. Phang et al. [1] investigated a technique for local processing and global processing using neural network structure. Local processing and global processing strategies are carried out using globally aware multiple instance classifier (GMIC) on mammogram images. GMIC technique is a neural network fusion structure performs better compared to the existing techniques such as ResNet and R-CNN. Technique is applied on only mammogram images. Future enhancement suggested extending work on tomosynthesis, MRI, ultrasound, etc.

Wessam M. Salama and Moustafa H Aly [2] proposed an automated framework for classifying of breast image. Various methods of classifiers like InceptionV3(IV3), DenseNet121(D121), ResNet50(RN50), VGG16 and MobileNetV2(MNV2) are used to segment the MIAS, DDSM and CBIS-DDSM into cancerous and non-cancerous breast mass. Two views Mediolateral Oblique and Cranio Caudal are considered in the dataset. These views helped to get more features and lead to improving the efficacy of diagnosis techniques. IV3 and adopted U-net achieved better results compared to the D121, RN50 and MNV2 techniques.

Xiangchun Yu, Wei Pang et al. [3] explored a technique to find the abnormalities in mammogram images has two different steps. In the first step identified patches of interest. In the second step applied a technique to classify them as normal or abnormal tissues. Proposed technique performed better than the existing techniques.

Siham A. Mohammed et al. [4] proposed a technique to improve the efficacy of diagnosis of breast carcinoma dataset. Class imbalance problem solved using oversampling technique. Then Bayes, J48 techniques applied on WBC dataset. Proposed technique increased accuracy of breast carcinoma diagnosis results.

Adam Yala et al. [5] investigated a method to improve diagnosis prediction of mammograms. Based on mammogram density and list of threats to women health, developed a hybrid deep model for prediction. Model performed in a better way compared to the existing techniques. 73% accuracy obtained using the hybrid technique.

2.2 Ultrasound research article

Essam H. Houssein et al. [6] explored applying deep learning techniques on breast cancer ultrasound images. Study shows that there is a greater significance of deep learning techniques as they help clinicians to improve their diagnosis. From the survey challenges are listed as requirement of training data, adopting of transfer learning technique to tackle the over fitting issues and robust techniques. Also survey covered applicability of different techniques to different modalities such as mammogram, ultrasound and tomography breast cancer images.

Yuan Xua et al. [7] proposed a technique to automatically classify images into four types of calcifications. Layer by layer network structure of neural such as CNN applied on ultrasound data. 80% accuracy obtained using CNN. Proposed extension of this work as to apply CNN on MRI and tomography based images.

Ge-Ge Wu et al. [8] investigated a technique for identifying abnormalities in ultrasound modality. In this approach examined cancer at the early stage using AI techniques. Technique automatic examining the images lessen the burden on radiologists.

2.3 Histopathological research article

C Kaushal et al. [9] explored a technique to improve the identification of significant patterns of data sets of H&E histopathological, 40x magnified and UDIAT-Centre Diagnostic cancer images of breast. Study shows flashing feature of firefly used as one of the feature for segmentation of an image. Firefly is used in improving identifying significant patterns. Technique outperformed the existing technique. Extension of future work suggested as incorporating other nature based optimization technique to improve further efficacy of diagnosis.

2.4 Optimization research article

Hu Peng Wenhua Zhu et al. [10] proposed fusion technique to increase the accuracy of cancer image classification of breast. Composite firefly (CFF) technique is been used. CFF involves opting for best firefly and any other two fireflies flashing features. Than just opting for best flashing firefly CFF displayed more efficient results. Significant features are selected in the fusion technique using firefly approach was better compared with the state of art techniques such as RaFa and Lifa.

2.5 Other research articles

OM Al-hazaimeh OM Al-hazaimeh et al. [11] proposed a new method that combines image processing and deep learning to diagnose diabetic retinopathy (DR) in retinal fundus images. The method involves four stages: image preprocessing, lesion segmentation, feature extraction, and classification. The authors used a publicly available dataset of retinal fundus images to train and evaluate the proposed model. The results show that the proposed method achieves a high accuracy of 94.5% in classifying DR severity levels. Moreover, the model has comparable performance to human experts in diagnosing DR. The authors suggest that the proposed method can be a useful tool for screening and diagnosing DR in clinical settings. Future enhancements could include incorporating additional clinical features and using larger datasets for training to improve the model's generalization capabilities.

Tan PH et al. [12] presents a revised classification system for breast tumors based on their molecular characteristics and clinical behavior. The methodology involves a comprehensive review of the literature, international expert consensus meetings, and analysis of large datasets. The new classification system aims to improve diagnostic accuracy and prognostic value and guide treatment decisions. The performance of the model is evaluated through studies showing improved prediction of patient outcomes and treatment response. The revised classification includes four main groups: luminal, HER2-positive, basal-like, and normal-like, with further subtypes based on additional molecular features. The authors recommend incorporating this new classification into routine clinical practice. The potential future enhancement for the WHO classification of breast tumors could be the incorporation of multi-omics data, such as genomics, transcriptomics and proteomics, to complement the existing morphological and immunohistochemical analysis. This could provide a more comprehensive understanding of the underlying biology of different breast tumor subtypes and potentially lead to the identification of new therapeutic targets.

Obaida M. Al-Hazaimeh et al. [13] proposes a geometrical-based approach for human image detection that is robust to occlusion and clutter. The approach involves detecting the human head and torso using a multi-scale HOG descriptor and geometric constraints and then refining the detection using an AdaBoost classifier. The proposed approach outperforms several state-of-the-art methods on two benchmark datasets, achieving high detection rates while maintaining low false positive rates. The approach is also robust to occlusion and clutter, making it suitable for real-world scenarios. Overall, the paper presents a promising approach for robust human image detection. Future enhancements include incorporating deep learning techniques and exploring the effectiveness of the proposed approach on different types of images.

3 Resources and existing techniques

In this section resources such as data sets used for experiment are detailed. ResNet-18 as a feature extractor explained in detail. Existing optimized methodology based classifier are discussed followed by ResNet-18.

3.1 Dataset

Three image datasets of breast cancer were considered for experiment. Histopathological (Breakhis) [14] consists of 9109 images. Out of these 2480 are non-cancerous and 5429 are cancerous images. Image Samples were collected by encountering biopsy of 82 patients with different telescopic factors. Mammogram CBIS-DDSM (Curated Breast Imaging Subset of DDSM) is a simplified and standardized form of the Digital Database for Screening Mammography (DSM) [15]. Data set consists 6775 of normal, non-cancerous, and cancerous images. Ultrasound (ultra) images were collected by Hammasat University and Queen Sirikit Center [16]. Out of these 296 cancerous and non-cancerous images were available for experiment.

3.2 ResNet-18 feature extractor

ResNet-18 [17] is used as a feature extractor. This model is trained on ImageNet database. ImageNet has of millions of images. ResNet-18 is one of the pre-trained models. 512 features are extracted from the last fully connected layer and size of each convolution block is as mentioned in the Figure 1.

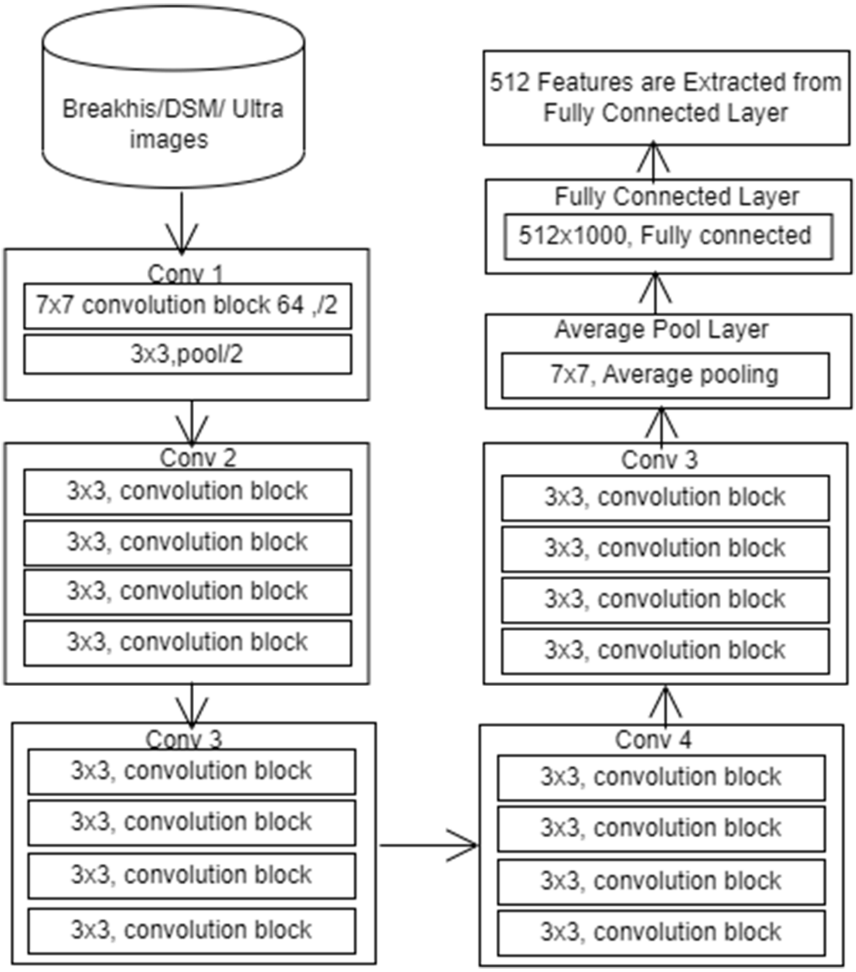

Fig. 1. ResNet-18 feature extractor

3.3 Methodology with existing optimization techniques

Optimization strategies are applied on the input parameters in order to tune them, to improve the performance of the classifier. In this study existing ant colony optimization (ACO) or genetic optimization (GAO) technique based classifier summarization is as shown in Figure 2.

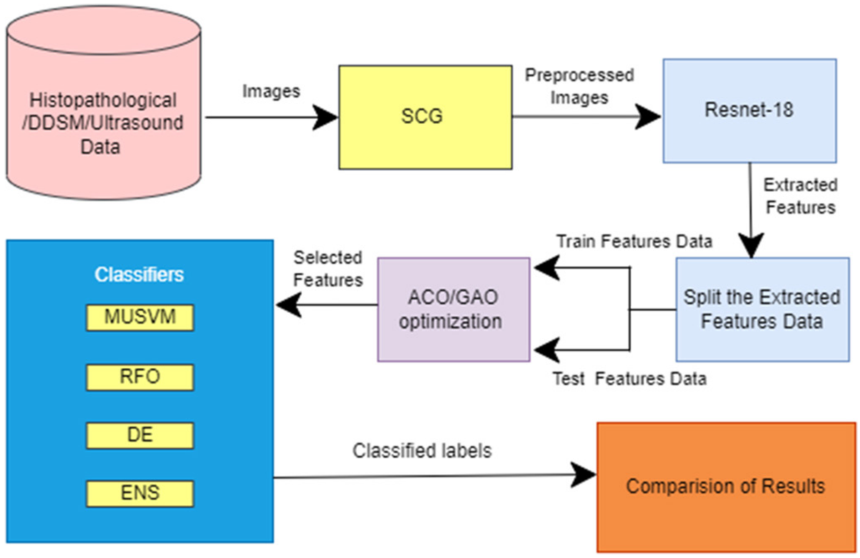

Fig. 2. ResNet18 and ACO/GAO based classifier

Steps followed in procedure are:

Step1: Collection of Histopathological, Mammogram and Ultrasound breast cancer data.
Step2: Applying data preprocessing denoising nonlinear edge preserving filters Sobel, Canny and Gabor filter (SCG) on the collected data.
Step3: Obtained preprocessed images are taken as input to ResNet-18 to extract the relevant features.
Step5: Extracted features are split as Train data and Test Data.
Step6: Split data taken as input to ACO or GAO approach to select significant features.
Step7: Significant features are then fed through the Multiclass Support Vector Machine (MUSVM), Random Forest (RFO), Decision Tree (DE) and Ensemble (Ens) to classify significant features as desired class labels as benign and malignant. In Ens classifier majority polling concept applied on classifiers MUSVM, DE, RFO, K Nearest Neighbor (KNE) and Naïve Bayes (NBA).

4 Proposed methodology

In this section Enhanced Firefly (EFF) optimization based classifier detailed. EFF derived from Traditional Firefly (TFF) optimization technique.

TFF [9] based on the strategy of moving one firefly towards another firefly or attracting prey based on the intensity of light produced by them.

Following terms are used in TFF:

- γ- Absorption co-efficient depend on the foggy and dark environmental factors
- α- Stability randomized factor
- β- Attractive coefficient depends on the light intensity produced by firefly
- dm is the Euclidean distance metric
- € normal distribution between 0 and 1

Whereas in EFF [10] considered probability of visibility of intensity of light of fireflies in brightness. Light intensity is denoted by 'IL'. Depending on the surface of high or low intensity values available in the given problem 'IL' varies from 0 to 1. In the enhanced firefly algorithm environmental factors are represented as 'n1' (dark) and 'n2' (foggy). The association between 'n1' and 'n2' can be defined using environmental factor 'E' of Eq. (1).

$$E = \frac{1}{1 + e^{-(n2/(n1(\ln(\mathrm{i}_{indexgen}))))}} \tag{1}$$

Variable indexgen varies from 1 to $\mathrm{i}_{\mathrm{highgen}}$. Overall summary of EFF is as in Figure 3.

Characterize Target function P(X)

Generate Community X_i with a uniform distribution over workspace

Initialize $\gamma,\alpha,\beta^{new},I_L$,n1 and n2

Calculate robustness values of each firefly $P(X_i)$,i=1,2,3,…………$n_{community}$

While ite<highgen

Calculate Es =1-($i_{indexgen}/i_{highgen}$) where ($1 \leq i_{indexgen} \leq i_{highgen}$)

if Es≥rd

Calculate $\beta^{new}=\beta(I_L)+E$ Where rd≤Es

else

Calculate $\beta^{new}=\beta(I_L)+E$

end if

for i=1 to $n_{community}$

for **j=1 to** $n_{community}$

if $P(X_j)>P(X_i)$

Moving firefly toward j in dm-dimension

end if

Update $\beta=\beta^{new}e^{-\gamma(dm)}$

Calculate solution $X_i=X_i+\beta e^{-\gamma(dm)ij}(X_i-X_j)+\alpha€$

end for j

end for i

Calculate Score of the fireflies and find the current best in the community

Ite=ite+1

end while

Note:--

β^{new}—New attractive coefficient

rd is a random number obtained from uniform distribution varies between 0 and 1

Fig. 3. EFF

4.1 Recommended methodology with EFF optimization technique

Investigated an architecture to classify the given modality images. Architecture is based on the EFF optimization technique is as shown in Figure 4.

Steps followed in investigated architecture algorithm are:

Step1: Collected Histopathological, Mammogram and Ultrasound breast images data.
Step2: Applied denoising nonlinear edge preserving filter Sobel canny and Gabor filter (SCG) on the image data.
Step3: Obtained preprocessed images are taken as input to ResNet-18(R18) to extract the relevant features
Step5: Extracted features are split as Train data and Test Data
Step6: Split data taken as input to EFF optimization approach and extracted significant features.
Step7: Significant features are then fed through the MUSVM, DE, RFO and ENS to classify significant features as desired class labels. In Ens classifier majority polling technique applied on classifiers MUSVM, DE, RFO, KNE and NBA.

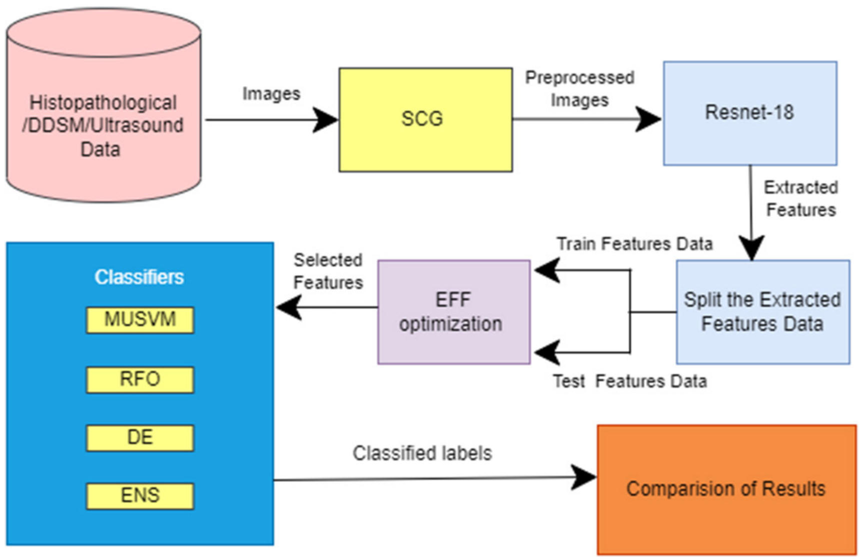

Fig. 4. ResNet18 and EFF based classifier

5 Assessment parameters and experiment results

Assessing parameters Confusion Matrix (CMA), Accuracy (Acc), Sensitivity (Sens), Specificity (Spec), F1 score (F1S) and Matthews correlation coefficient (MCC) are discussed in this section. Acc, Sens, Spec, F1S and MCC equations are as in Eq. (2), Eq. (3), Eq. (4), Eq. (5) and Eq. (6) respectively. Followed by this Experiment results are considered.

CMA has four terms. TRP is about the class which is actually positive (P) and labeled as positive. TRN is about the class which is actually negative (N) and labeled as negative. FAN is about the ones are actually positive but labeled as negative. FAP is about the ones are actually negative and labeled as positive.

Actual vs. Prediction		**Prediction**	
		P	**N**
Actual	**P**	TRP	FAN
	N	FAP	TRN

$$Acc = \frac{TRP + TRN}{TRP + TRN + FAN + FAP} \tag{2}$$

$$Sens = \frac{TRP}{TRP + FAN} \tag{3}$$

$$Spec = \frac{TRN}{TRN + FAP} \tag{4}$$

$$F1S = \frac{2 * P * Sens}{P + Sens} \tag{5}$$

Where $P = \dfrac{TRP}{TRP + FAP}$

$$MCC = \frac{(TRP * TRN) - (FAN * FAP)}{\sqrt{(TRP + FAN)(TRP + FAN)(TRN + FAP)(TRN + FAN)}} \tag{6}$$

5.1 Experiment results

EFF based classifier tested with Acc, Sens, Spec, F1S and MCC parameters for Breakhis Data. Proposed approach EFF compared with available ACO and GAO based classifiers. 3500 Breakhis images results are tabulated for EFF, ACO and GAO optimization techniques based classifiers MUSVM, RFO, DE and Ens as in Tables 1, 2 and 3. Results of these techniques displayed as in Figure 5.

Table 1. EFF based classifier performance for breakhis data

	Acc	**Sens**	**Spec**	**F1S**	**MCC**
MUSVM	50.00	25.00	25.00	33.33	0.00
RFO	96.52	96.75	96.75	96.52	93.27
DE	97.93	97.93	97.93	97.93	95.87
Ens	99.85	99.70	100.00	99.85	99.70

Table 2. ACO based classifier performance for breakhis data

	Acc	**Sens**	**Spec**	**F1S**	**MCC**
MUSVM	94.30	94.31	94.31	94.30	88.62
RFO	91.23	91.52	91.52	91.21	82.75
DE	92.24	92.25	92.25	92.24	84.48
Ens	94.51	93.94	95.11	94.54	89.03

Table 3. GAO based classifier performance for breakhis data

	Acc	**Sens**	**Spec**	**F1S**	**MCC**
MUSVM	50.00	25.00	25.00	33.33	0.00
RFO	97.18	97.32	97.32	97.18	94.50
DE	97.48	97.48	97.48	97.48	94.96
Ens	99.55	99.10	100.00	99.55	99.10

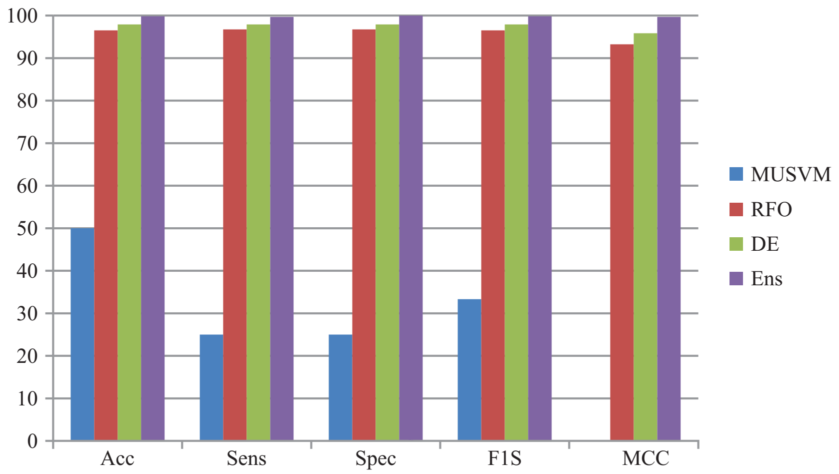

Fig. 5. ResNet18 and EFF based classifier results for breakhis data

EFF based classifier tested with Acc, Sens, Spec, F1S and MCC parameters for DSM data. Proposed approach EFF compared with available ACO and GAO based classifiers. 1000 DSM images results are tabulated for EFF, ACO and GAO optimization techniques based classifiers MUSVM, RFO, DE and Ens as in Tables 4, 5 and 6. Results of these techniques displayed as in Figure 6.

Table 4. EFF based classifier performance for DSM data

	Acc	Sens	Spec	F1S	MCC
MUSVM	98.53	97.11	99.26	97.81	96.71
RFO	94.20	89.12	97.04	91.54	87.27
DE	87.46	77.77	93.34	82.29	72.99
Ens	99.26	98.53	100.00	99.26	98.52

Table 5. ACO based classifier performance for DSM data

	Acc	Sens	Spec	F1S	MCC
MUSVM	66.14	49.50	79.81	56.39	30.75
RFO	66.99	50.48	80.58	57.28	32.49
DE	62.66	45.67	77.24	52.67	24.08
Ens	83.47	82.54	84.35	81.90	66.71

Table 6. GAO based classifier performance for DSM data

	Acc	Sens	Spec	F1S	MCC
MUSVM	97.45	96.95	99.22	97.69	96.53
RFO	92.48	86.10	96.12	89.14	83.58
DE	87.79	78.32	93.53	82.73	73.69
Ens	98.01	98.17	99.88	99.01	98.04

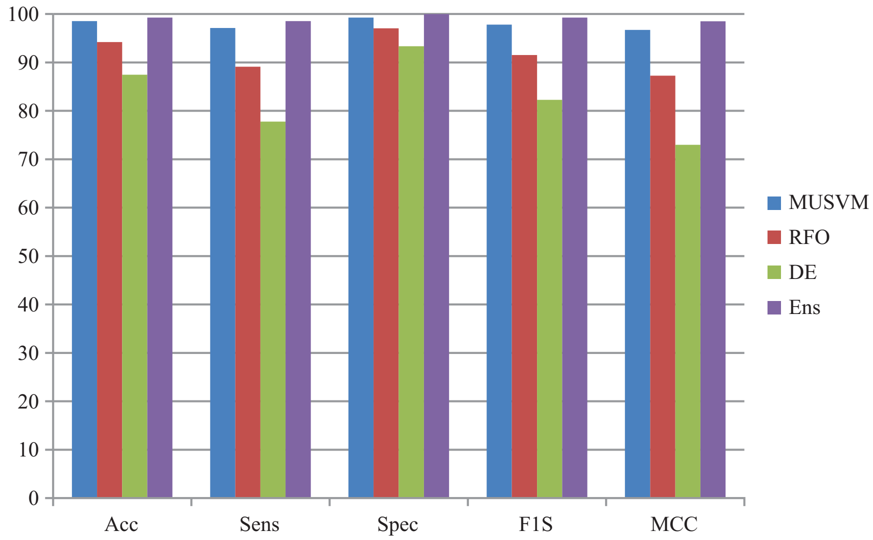

Fig. 6. ResNet18 and EFF based classifier results for DSM data

EFF based classifier tested with Acc, Sens, Spec, F1S and MCC parameters for ultra data. Proposed approach EFF compared with available ACO and GAO based classifiers. 100 Ultra images results are tabulated for EFF, ACO and GAO optimization techniques based classifiers MUSVM, RFO, DE and Ens as in Tables 7, 8 and 9. Results of these techniques displayed as in Figure 7.

Table 7. EFF based classifier performance for ultra data

	Acc	Sens	Spec	F1S	MCC
MUSVM	55.00	66.37	66.37	43.13	20.07
RFO	95.00	95.00	95.00	95.00	90.00
DE	91.67	92.26	92.26	91.64	83.92
Ens	99.17	100.00	98.46	99.78	98.73

Table 8. ACO based classifier performance for ultra data

	Acc	Sens	Spec	F1S	MCC
MUSVM	81.94	84.24	84.24	81.39	65.99
RFO	79.17	79.46	79.46	79.12	58.63
DE	77.78	78.41	78.41	77.64	56.18
Ens	83.33	87.41	80.75	81.95	67.40

Table 9. GAO based classifier performance for ultra data

	Acc	Sens	Spec	F1S	MCC
MUSVM	99.31	99.36	99.36	99.30	98.66
RFO	94.44	95.02	95.02	94.41	89.46
DE	88.19	88.45	88.45	88.18	76.64
Ens	99.06	98.72	97.00	98.33	98.44

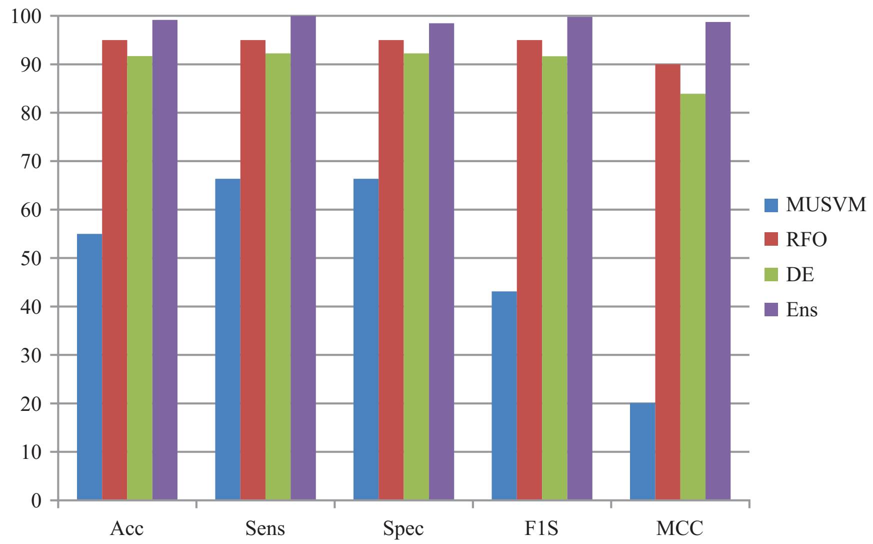

Fig. 7. ResNet18 and EFF based classifier results for ultra data

5.2 Discussion

Experiment carried out on three modalities. GAO based Ens classifier has accuracy of 99.55, 98.01 and 99.06 for Breakhis, DSM and Ultra breast cancer image data respectively. ACO based Ens classifier has accuracy of 94.51, 83.47 and 83.33 for Breakhis, DSM and Ultra breast cancer images. EFF based Ens classifier has accuracy of 99.85, 99.26 and 99.17 for Breakhis, DSM and Ultra breast cancer images. As per experiment results displayed in Figures 4, 5 and 6 EFF based Ens classifier has outperformed ACO and GAO based Ens classifier.

6 Conclusion

EFF based proposed methodology performed better compared with ACO and GAO based methodologies for Breakhis, DSM and Ultra image data. As EFF based methodology displays accuracy of 99.85, 99.26 and 99.17. Early diagnosis of cancer required to improve the number of survival of patients. Investigated EFF based architecture classifier in this study shows that it can be used as a fundamental diagnosis by clinical radiologists.

As an extension of this work EFF based classifier can be applied on tomography, CT scan and MRI modalities.

7 References

[1] Shena, Yiqiu, et al., (2021), An interpretable classifier for high-resolution breast cancer screening images utilizing weakly supervised localization: Medical Image Analysis 68:101908. https://doi.org/10.1016/j.media.2020.101908

[2] Salama, A., Wessam, M., and Aly (2021), Deep learning in mammography images segmentation and classification: Automated CNN approach: Alexandria Engineering Journal 60:4701–4709. https://doi.org/10.1016/j.aej.2021.03.048

[3] Yu, Xiangchun, et al., (2020), Mammographic image classifcation with deep fusion learning: Scientific Reports 10:14361. https://doi.org/10.1038/s41598-020-71431-x

[4] Mohammed, Siham A., et al., (2020), Analysis of breast cancer detection using different machine learning techniques: DMBD 2020, CCIS 1234:108–117. https://doi.org/10.1007/978-981-15-7205-0_1

[5] Yala, Adam, et al., (2019), A deep learning mammography-based model for improved breast cancer risk prediction: Radiology 292(1):60–66. https://doi.org/10.1148/radiol.2019182716

[6] Houssein, Essam, H., et al., (2020), Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review: Expert Systems with Applications 167:114161. https://doi.org/10.1016/j.eswa.2020.114161

[7] Xua, Yuan, et al., (2018), Medical breast ultrasound image segmentation by machine learning: Ultrasonics 91. https://doi.org/10.1016/j.ultras.2018.07.006

[8] Wu, Ge-Ge, et al., (2019), Artificial intelligence in breast ultrasound: World J Radiol 11(2): 19–26. https://doi.org/10.4329/wjr.v11.i2.19

[9] Kaushal, C., Kaushal, Kirti and Singla, A., (2020), Firefly optimization based segmentation technique to analyze medical images of breast cancer: International Journal of Computer Mathematics 98(6):1–19. https://doi.org/10.1080/00207160.2020.1817411

[10] Wenhua Zhu, Hu Peng, Deng Kun Changshou, Zhijian Wu, and Yu (2020), Composite firefly algorithm for breast cancer recognition: Concurrency Computat Pract Exper, 2020:e6032. https://doi.org/10.1002/cpe.603

[11] Obaida M. Al-hazaimeh, et al., (2022), Combining artificial intelligence and image processing for diagnosing diabetic retinopathy in retinal fundus images: International Journal of Online Engineering 18(3). https://doi.org/10.3991/ijoe.v18i13.33985

[12] Tan, PH., et al., (2020), The 2019 world health organization classification of tumours of the breast: WHO Classification of Tumours Editorial Board, Histopathology 77(2):181–185. https://doi.org/10.1111/his.14091

[13] Obaida, M. and Al-Hazaimeh, et al., (2019), Geometrical-based approach for robust human image detection: Multimed Tools Appl 78:7029–7053. https://doi.org/10.1007/s11042-018-6401-y

[14] Spanhol, F., (2016), A dataset for breast cancer histopathological image classification: IEEE transactions on biomedical engineering (TBME) 63(7):1455–1462. https://doi.org/10.1109/TBME.2015.2496264

[15] Sawyer Lee, Rebecca, et al., (2017), A curated mammography data set for use in computer-aided detection and diagnosis research: Scientific Data volume 4, Article number: 170177. https://doi.org/10.1038/sdata.2017.177

[16] Rodtook, A., et al., (2018), "Automatic initialization of active contours and level set method in ultrasound images of breast abnormalities: Pattern Recognition 79:172–182. https://doi.org/10.1016/j.patcog.2018.01.032

[17] ImageNet. http://www.image-net.org

8 Authors

Mrs. Anupama Y.K., Assistant Professor of Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bengaluru. She obtained her M.Tech and B.E. degree from VTU University. She has published in 9 journals and conferences. Her research interests include data mining, digital image processing and computer vision. (e-mail: anupama-cs@dayanandasagar.edu)

Dr. Amutha S., Professor of Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bengaluru. She obtained her Ph.D. degree from VTU University. She has more than 24 years of teaching experience. She has published in 29 journals and conferences. Her research interests include Big Data, data mining, and cloud computing. (e-mail: amutha-cs@dayanandasagar.edu)

Dr. Ramesh Babu D.R., HOD & Professor of Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bengaluru. He obtained his Ph.D. degree from Mysore University. He has more than 20 years of teaching experience. He has published in 28 journals and 20 conferences. His research interests are data mining, digital image processing and computer vision. (e-mail: bobrammysore@gmail.com)

Dr. S. Sathish Kumar, Professor of Department of Information Science and Engineering, RNSIT college of Engineering, Bengaluru. He obtained his B.E degree from PSNA College of Engineering, M.E from Regional Engineering College, Trichy, Tamil Nadu and Ph.D. from Dr. MGR Educational and Research Institute University. He has more than 35 years of teaching experience. He has published in 28 journals and conferences. His research interests are Big Data and Cloud computing. (e-mail: sathish_tri@yahoo.com)

Article submitted 2023-01-05. Resubmitted 2023-03-07. Final acceptance 2023-03-13. Final version published as submitted by the authors.