



Hybrid Feature Selection Using the Firefly Algorithm for Automatic Detection of Benign/Malignant Breast Cancer in Ultrasound Images

Dafni Rose Jesuharan^{1*}, Thason Thaj Mary Delsy², Vijayakumar Kandasamy¹, Pradeep Mohan Kumar Kanagasabapathy³

¹ Department of Computer Science and Engineering, St. Joseph's Institute of Technology, Chennai 600119, Tamil Nadu, India

² Sathyabama Institute of Science and Technology, Chennai 600119, Tamil Nadu, India

³ Department of Computing Technologies, SRM Institute of Science and Technology, Chennai 603203, Tamil Nadu, India

Corresponding Author Email: hodcsestaffaffairs@stjosephstechnology.ac.in

Copyright: ©2023 IIETA. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.400628>

ABSTRACT

Received: 21 March 2023

Revised: 1 August 2023

Accepted: 6 September 2023

Available online: 30 December 2023

Keywords:

cancer, ultrasound-imaging, ResUNet, ResNet18, firefly algorithm, classification

The incidence rate of breast cancer (BC) is progressively increasing worldwide, and early diagnosis can help reduce the mortality rate. Ultrasound imaging, a cost-effective imaging technique, is widely used for initial screening of patients suspected of having breast cancer. Categorizing breast ultrasound images into benign and malignant classes is crucial for planning appropriate treatment strategies to combat BC. This research proposes a Convolutional Neural Network (CNN) framework to classify breast ultrasound images. This framework comprises the following stages: (i) image collection and resizing, (ii) CNN segmentation to extract the cancerous region, (iii) deep feature mining, (iv) extraction of handcrafted features, (v) selection of optimal features based on the Firefly algorithm (FA) and serial concatenation of features to create the feature vector, and (vi) performance evaluation and validation. The proposed classification task is executed using (i) deep-feature-based classification and (ii) integrated deep and handcrafted (hybrid) features. Experimental outcomes confirm that the ResNet18-based deep features achieve a classification accuracy of 91% with the SoftMax classifier, while the proposed hybrid features provide a classification accuracy of 99.50% with the K-Nearest Neighbor (KNN). These results underscore the significance of the proposed scheme.

1. INTRODUCTION

In the modern era, numerous facilities exist to enhance human living conditions. Despite significant medical advancements, cancer incidence continues to rise due to various unavoidable factors. Early detection and treatment can significantly improve disease prognosis [1-3]. Typically, cancer is caused by abnormal cell or tissue growth and standard treatment methodologies include medication, radiotherapy, chemotherapy, and surgery. Early-stage cancer (benign) can often be treated with medication and chemotherapy, while advanced cancer (malignant) often requires more complex treatment procedures, including surgery [4, 5].

In 2020, cancer caused over 10 million deaths worldwide, accounting for nearly one-sixth of all deaths. According to the World Health Organization (WHO) report for 2020, cancer has emerged as one of the leading causes of death. Breast cancer (2.26 million), lung cancer (2.21 million), colorectal cancer (1.93 million), prostate cancer (1.41 million), and skin cancer (1.2 million) are the most commonly diagnosed cancers [6]. Despite the implementation of necessary treatments, cancer remains a leading cause of death, with the WHO report confirming 685,000 deaths from breast cancer alone. Recently, the WHO established a women's health chatbot to disseminate

information about breast cancer. It is available in multiple languages, including English, Greek, Hungarian, Russian, and Ukrainian, with plans to add more languages in the future [7].

To detect breast cancer (BC) at early stages, several screening procedures have been developed due to its rapid incidence and high fatality rate. Standard BC screening procedures include physical examination to identify irregularities in the breast, biomedical imaging-supported analysis, and needle biopsy for confirming the severity of BC. Once the exact stage of cancer is determined, appropriate treatment procedures such as medication, chemotherapy, and surgery are implemented to treat the patient.

Biomedical imaging-supported examinations are one of the most widely used screening methodologies for BC. Techniques such as Magnetic Resonance Imaging [8, 9], Thermal Imaging [10, 11], and Ultrasound Imaging (UI) [12] are extensively used to detect the disease and its severity. UI, a recently developed non-invasive procedure, is a powerful tool for early BC detection. Prior research has confirmed that UI helps detect early and acute forms of cancer with better accuracy than other methods. The image recording process of UI is straightforward and requires just a simple ultrasound transmitter-receiver setup to capture the breast section and detect abnormalities.

Convolutional Neural Network (CNN) supported schemes

have shown promising results in the literature. This research proposes a CNN approach integrating segmentation and classification methods. The stages of this scheme include data collection and initial processing, ResUNet-supported breast tumor segmentation, handcrafted feature extraction, deep feature extraction, Firefly Algorithm (FA) based feature optimization and integration of deep and handcrafted features, and finally, classification and validation.

To validate the performance of the developed scheme, this work implements the classification task using individual deep features and integrated deep and handcrafted features. The proposed work considers pretrained CNN models such as AlexNet, VGG16, VGG19, ResNet18, ResNet50, and ResNet101. Initially, the classification task is executed using the SoftMax classifier, and the CNN model providing the best classification result is considered for further analysis. This work also considers handcrafted features like Local Binary Patterns (LBP) and Discrete Wavelet Transform (DWT) features collected from the resized UI, as well as the Gray Level Co-occurrence Matrix (GLCM) mined from the segmented tumor. The optimized features of the CNN and the handcrafted features are then serially integrated to create a new feature vector for better benign and malignant classification.

This work considers a binary classification with a five-fold cross validation and the achieved results with individual and integrated features are separately analyzed. The result of this study confirms that the individual deep feature of ResNet18 achieves a 91.50% accuracy with deep features and the SoftMax classifier. The integrated feature provides a 99.50% accuracy with the K-Nearest Neighbor (KNN) classifier. This result confirms that the ResUNet and ResNet18-based classification provides superior results.

The contributions of this research include:

1. Detection of benign and malignant breast cancer using integrated CNN segmentation and classification.
2. Implementation of ResUNet-based segmentation and ResNet18-based classification.
3. Firefly algorithm-based feature selection and hybrid feature-supported classification executed on UI database.

This research is organized as follows: Section 2 presents related works, Section 3 outlines the methodology, and Sections 4 and 5 provide results and conclusions, respectively.

2. RELATED WORKS

The incidence rate of breast cancer (BC) in women is escalating globally due to a multitude of factors, including age, heredity, and personal habits such as obesity, delayed pregnancy, alcohol consumption, and having children at a later age [13-15]. Clinical-level screening of BC typically employs a chosen biomedical imaging modality, with ultrasound imaging (UI) being a commonly adopted approach in hospitals. Previous research on UI-based examinations validates the efficacy of this imaging modality in classifying BC into benign and malignant categories, which aids in strategizing and executing the appropriate treatment plans. A selection of UI-supported BC detection methods available in the literature is presented in Table 1. These prior studies affirm that UI-based diagnosis is instrumental in achieving superior results in BC classification.

Table 1. An overview of medical imaging schemes for the detection of breast cancer

Reference	Examination Procedure
Rajinikanth et al. [16]	Non-invasive Thermal Imaging (TI) methodology is presented for automatic detection of BC using machine learning.
Dey et al. [17]	The use of ML to detect early/acute BC is presented with the application of the TI, and this work has demonstrated detection accuracy more than 90%.
Fernandes et al. [18]	It has been demonstrated that by using the TI, an assessment of abnormal breast regions using joint thresholding and segmentation can be provided.
Thanaraj et al. [19]	In this paper, Shannon's entropy is used to automatically segment BC lesions in UI by combining level-set segmentation with Shannon's entropy.
Ilesanmi et al. [20]	Using an enhanced deep learning scheme, we present an automatic method for automatically extracting the BC lesion from UI.
Meraj et al. [21]	The purpose of this work is to perform U-Net segmentation and deep-feature based classification of the user interface (UI) in order to detect the BC and achieve a classification accuracy of over 98%.
Irfan et al. [22]	This work implemented parallel feature fusion approach to detect the BC in UI and achieved an accuracy of 98.97%
Jabeen et al. [23]	This work implemented probability-based optimal deep features fusion to classify the UI to recognize the BC and its category. This work provided a overall classification accuracy of 99.1%.
Vijayakumar et al. [24]	Implementation of ML based BC detection using the UI is presented and the Mayfly-Algorithm selected optimal features are adopted. This work provided a classification accuracy of >91%.
Zhou et al. [25]	This work presented computerized segmentation and classification of BC tumors in 3D UI using multi-task learning methodology.
Lei et al. [26]	Mask scoring R-CNN supported segmentation of the BC tumor in 3D UI is presented.
Sehgal et al. [27]	This work presents the review of the methods available to analyze the breast UI.

The studies discussed in Table 1 indicate that UI-based BC detection requires a carefully designed image examination framework. Research by Irfan et al. [22] illustrates that integrating CNN segmentation and classification yields better results during image-supported disease detection. Therefore, this proposed research also adopts a similar approach to segment the BC region from the UI, and subsequently classifies the test images into benign or malignant based on the tumor dimensions.

3. METHODOLOGY

The success of an automatic disease detection depends on the methodology designed and implemented to examine the chosen medical images. This section presents the procedures executed in the UI based BC detection task.

3.1 Implemented scheme

As shown in Figure 1, the proposed scheme uses CNN segmentation (ResUNet) and classification to detect BC classes using user interfaces. As a first step, the resized test images and ground-truths (GTs) are considered to train the ResUNet. The trained ResUNet scheme is then considered to extract the tumor section from the chosen UI database and these sections are then considered to mine the tumor features using GLCM.

The deep features from test images are mined using the chosen pretrained model and the handcrafted features are then mined using LBP and DWT. The handcrafted features, like GLCM, LBP and DWT are integrated together to form a real handcrafted feature vector. In order to reduce the deep and handcrafted features vector, this work considers the FA algorithm and the reduces features are serially integrated together to generate a new feature vector. The deep handcrafted (hybrid) future vector is then considered to classify the chosen UI images into benign and malignant class.

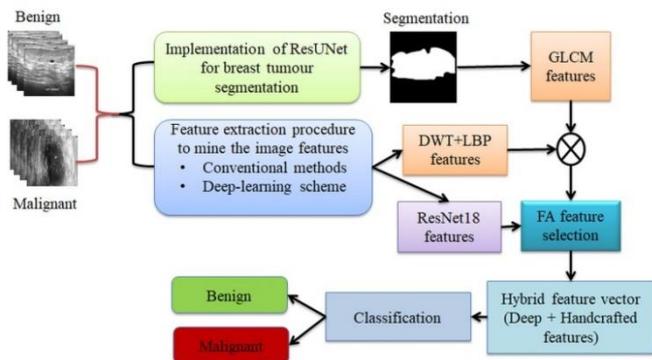


Figure 1. CNN supported scheme to detect the BC from the ultrasound imagery

3.2 Image database

The clinical-grade medical image choice is essential to bridge the gap between theoretical and practical research work. In this work, the clinical grade breast UI provided by Al-Dhabyani et al. [28] is used for the assessment. Before the assessment, the necessary picture augmentation is implemented to increase the number of test images to 1000 (500 benign and 500 malignant classes). Initially, each picture and GT are resized to $512 \times 512 \times 1$ pixels to train and test the ResUNet, and the necessary tumor section from every image is extracted. Later, the classification task is executed with 80% images ($400+400=800$ images) to train the proposed system and 20% of images ($100+100=200$ images) to validate the performance. The necessary details of the images are depicted in Table 2, and the example images are shown in Figure 2.

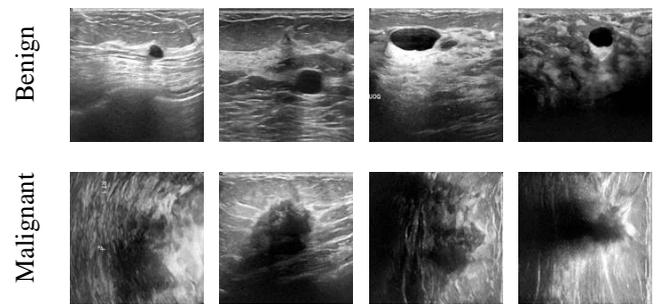
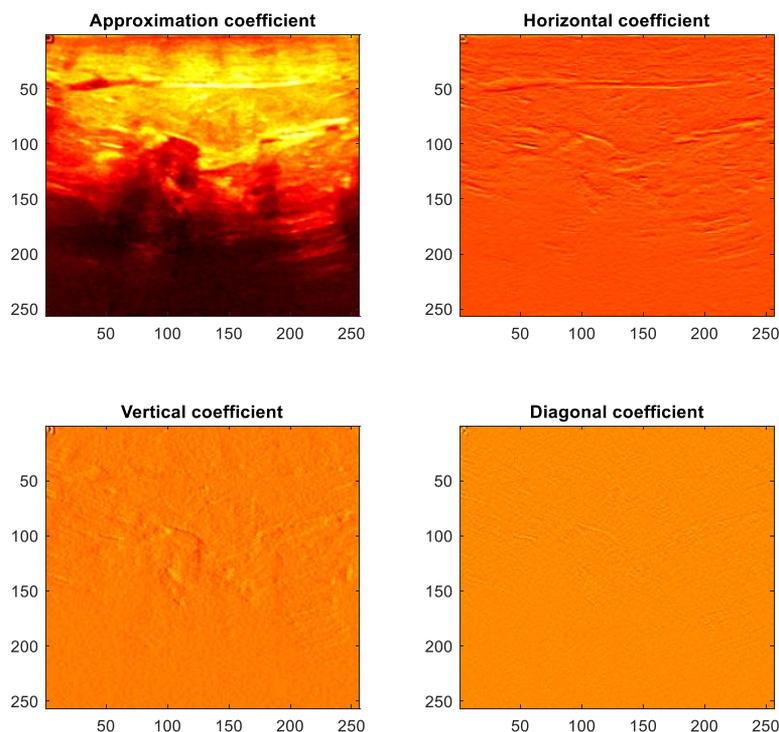


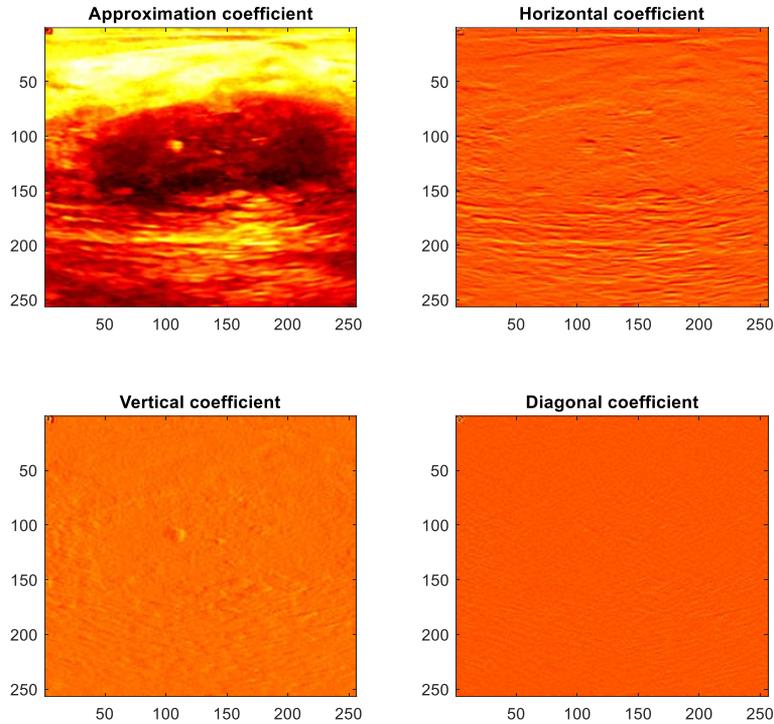
Figure 2. Sample breast ultrasound images of this study

Table 2. Images used in the proposed research work

Class	Total (100%)	Images	
		Training (80%)	Validation (20%)
Benign	500	400	100
Malignant	500	400	100



(a) Benign



(b) Malignant

Figure 3. DWT enhanced breast ultrasound pictures

3.3 ResUNet

The segmentation task is widely employed in the medical image evaluation domain to mine and assess the abnormality. However, the recent works confirm that the CNN segmentation helps get better results than the conventional abnormality extraction tasks. Hence, this work implemented the ResUNet scheme to segment the breast tumor with better accuracy. In ResUNet, a decoder-encoder based on the ResNet18 scheme is implemented to learn the lesion section from the test image and the GT. After appropriate learning, the lesion section is mined with better accuracy, and this mined lesion is then considered to get the DWT and GLCM features as discussed in studies [29-35]. The earlier segmentation works which employ the ResUNet segmentation can be accessed from studies [36, 37].

feature improves classification accuracy when used with the Mayfly Algorithm. To get the necessary information from the breast UI, DWT [38-40], GLCM [41, 42], and LBP [43, 44] are considered. This scheme extracts not just these features, but also deep features. Figures 3 and 4 show DWT and LBP enhanced images, and you can find the theoretical background in studies [39, 40, 43, 44]. LBP in this work is based on the methodology proposed by Gudigar et al. [43], in which the weight (W) is varied from 1 to 4. Figure 3(a) and (b) present the DWT enhanced images from which the DWT features of dimension $1 \times 1 \times 160$ are extracted from all the images (Approximation, Horizontal, Vertical and Diagonal coefficients). Later $1 \times 1 \times 236$ features are then extracted from the LBP images shown in Figure 4. Along with these features, $1 \times 1 \times 23$ GLCM features are also computed.

The handcrafted and deep-features mined in his work is presented below as in Eqs. (1) to (8):

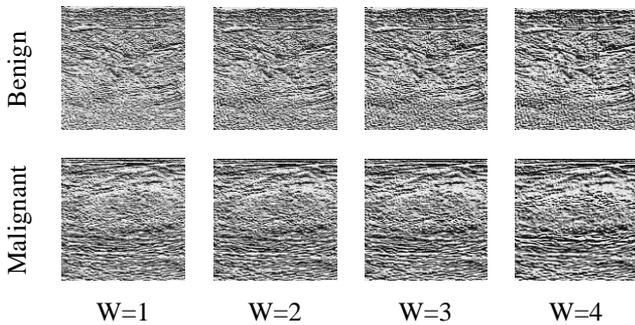


Figure 4. LBP enhanced breast ultrasound pictures

3.4 Feature extraction

Based on the features considered during the classification task, automatic disease diagnosis performs better. A recent study by Vijayakumar et al. [24] shows that the DWT and LBP

$$\begin{aligned} & \textit{Approximation}_{(1 \times 1 \times 40)} \\ & = DWT1_{(1,1)}, DWT1_{(1,2)}, \dots, DWT1_{(1,40)} \end{aligned} \quad (1)$$

$$\begin{aligned} & \textit{Horizontal}_{(1 \times 1 \times 40)} \\ & = DWT2_{(1,1)}, DWT2_{(1,2)}, \dots, DWT2_{(1,40)} \end{aligned} \quad (2)$$

$$\begin{aligned} & \textit{Vertical}_{(1 \times 1 \times 40)} \\ & = DWT3_{(1,1)}, DWT3_{(1,2)}, \dots, DWT3_{(1,40)} \end{aligned} \quad (3)$$

$$\begin{aligned} & \textit{Diagonal}_{(1 \times 1 \times 40)} \\ & = DWT4_{(1,1)}, DWT4_{(1,2)}, \dots, DWT4_{(1,40)} \end{aligned} \quad (4)$$

$$\begin{aligned} DWT_{(1 \times 1 \times 160)} & = DWT1_{(1 \times 1 \times 40)} + DWT2_{(1 \times 1 \times 40)} \\ & \quad + DWT3_{(1 \times 1 \times 40)} \\ & \quad + DWT4_{(1 \times 1 \times 40)} \end{aligned} \quad (5)$$

$$LBP_{(1 \times 1 \times 236)} = LBP1_{(1 \times 1 \times 59)} + LBP2_{(1 \times 1 \times 59)} + LNBP3_{(1 \times 1 \times 59)} + LBP4_{(1 \times 1 \times 59)} \quad (6)$$

$$Handcrafted\ features_{(1 \times 1 \times 399)} = DWT_{(1 \times 1 \times 160)} + LBP_{(1 \times 1 \times 236)} + GLCM_{(1 \times 1 \times 23)} \quad (7)$$

$$Deep\ features_{(1 \times 1 \times 1000)} = DL_{(1,1)}, DL_{(1,2)}, \dots, DL_{(1,1000)} \quad (8)$$

Eqs. (1) to (4) presents the DWT features and Eq. (5) shows the combined DWT features. Eq. (6) presents the LBP feature and Eqs. (7) and (8) shows the overall handcrafted and deep features of this research. These features are then optimized with FA and the optimized features are consecutively integrated to acquire the hybrid feature vector.

3.5 Feature reduction

FA is a nature-inspired heuristic method developed by Yang and published in 2008 [45]. As a consequence of its merit and high level of optimization accuracy, FA has been widely adopted and adopted by researchers to find solutions to several optimization problems [46-48].

The mathematical expression of the FA is depicted below:

Let us consider, in a search space, there exist two groups of fireflies, like i and j . Due to its attractiveness, the firefly i will move close to j , and this procedure can be demoted as follows:

$$X_i^{t+1} = X_i^t + \beta_0 e^{-\gamma d_{ij}^t} (X_j^t - X_i^t) + LF \quad (9)$$

where, X_i^t =early place of firefly i , X_j^t =early place of firefly j , LF =Levy walk, β_0 =attractiveness coefficient, γ =light absorption coefficient and d_{ij}^t =Cartesian distance (CD) between flies.

The fireflies try to find the CD between benign and malignant images during the feature optimization task. It's important to consider the feature with the maximal CD, and to discard the ones with a minimal CD. Figure 5 illustrates the concept of FA-supported feature reduction, and similar procedures can be found elsewhere [49, 50]. There are 30 flies in this work, 1500 iterations, and maximum CD as guiding parameter.

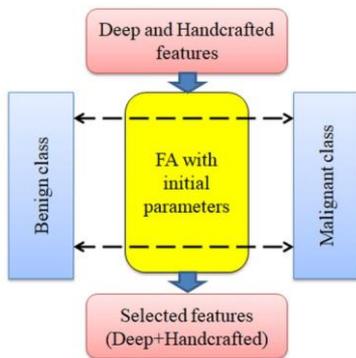


Figure 5. FA based feature reduction

In this work, the FA-supported reduction provided $1 \times 1 \times 173$ handcrafted and $1 \times 1 \times 411$ depth features, and the serial combination of these values provided the hybrid feature vector as presented in Eq. (10):

$$Hybrid\ features_{(1 \times 1 \times 584)} = Handcrafted_{(1 \times 1 \times 173)} + Deep_{(1 \times 1 \times 411)} \quad (10)$$

This feature vector is then used for training the classifier.

3.6 Classification and validation

UI sections are segmented with ResUNet and compared with GTs. Hybrid features are applied for UI classification. To classify benign/malignant UI sections, the SoftMax classifier is applied first. The merit of implemented scheme is validated with available binary classifiers, like Naïve-Bayes (NB), Decision-Tree (DT), Random-Forest (RF), K-Nearest Neighbor (KNN), and Support-Vector-Machine (SVM with linear/RBF kernels). The primary metrics, like True-Positive (TP), False-Negative (FN), True-Negative (TN), and False-Positive (FP) are initially obtained. To appraise the performance of this scheme, Jaccard (JA), Dice (DI), Accuracy (AC), Precision (PR), Sensitivity (SE), Specificity (SP), and F1-Score (FS) are computed as in Eqs. (11) to (17) [51-57]:

$$JA = \frac{TP}{TP+FP+FN} \quad (11)$$

$$DI = \frac{2TP}{2TP+FP+FN} \quad (12)$$

$$AC = \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

$$PR = \frac{TP}{TP+FP} \quad (14)$$

$$SE = \frac{TP}{TP+FN} \quad (15)$$

$$SP = \frac{TN}{TN+FP} \quad (16)$$

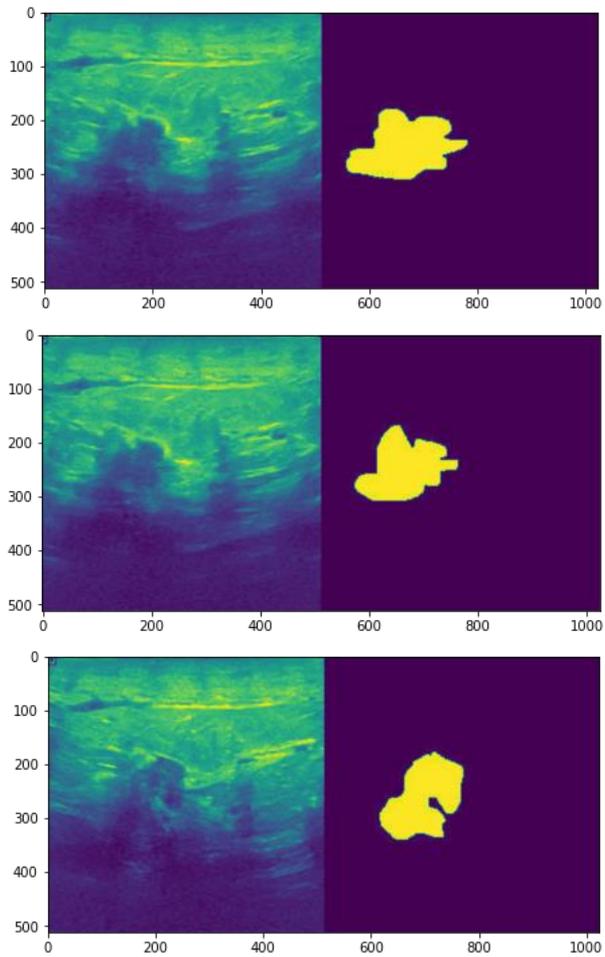
$$FS = \frac{2TP}{2TP+FN+FP} \quad (17)$$

4. EXPERIMENTAL RESULTS

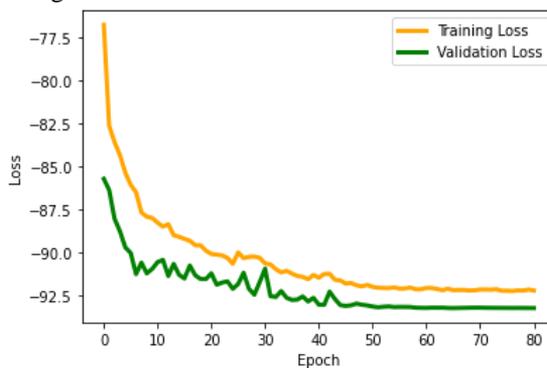
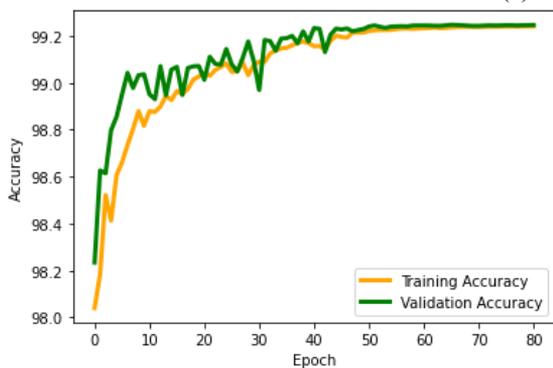
Investigational examination of planned framework using Intel i5 processor, 16GB RAM, and 2GB VRAM set with Python® is conducted.

In this scheme, CNN segmentation is done using a pre-trained ResUNet scheme, and CNN classification is done with the extracted breast tumor to mine GLCM features. The result is shown in Figure 6. In Figure 6, the sample images and GT that were used to train the scheme are presented (a), and the accuracy and loss values are shown (b) and (c). This system was trained and validated to be >99% accurate and Fig 6(b) shows the mined tumor and its image.

Following mining of the sample images and segmentation of the tumor, Table 3 provides the necessary metrics for computation of the metrics listed in Table 3. As a result of the proposed scheme, the tumor section can be mined with greater accuracy. Table 3 indicates that segmentation accuracy is greater than 97%. These images are considered for the GLCM features analysis. Figure 7 presents segmented results achieved with few chosen sample images (Figure 7(a) to (d)).

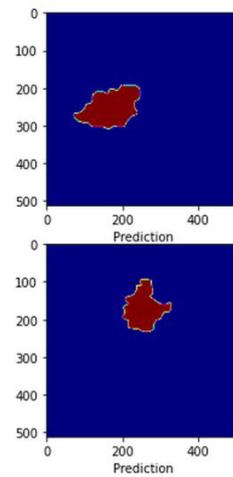
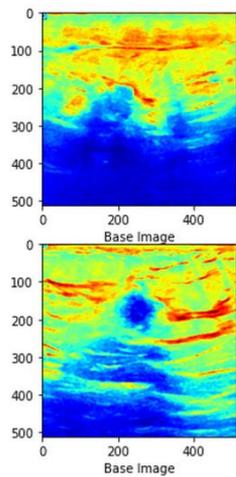


(a) Training



(b) Accuracy

(c) Loss



(d) ResUNet segmentation

Figure 6. CNN Segmentation

Table 3. Segmentation outcome achieved with ResUNet

Method	TP	FN	TN	FP	JA	DI	AC	PR	SE	SP
Im1	86577	2668	169683	3216	93.6362	96.7135	97.7554	96.4184	97.0105	98.1400
Im2	15595	1360	244275	914	87.2741	93.2046	99.1325	94.4636	91.9788	99.6272
Im3	14176	1947	245580	441	85.5832	92.2316	99.0891	96.9830	87.9241	99.8207
Im4	15276	2010	244144	714	84.8667	91.8139	98.9609	95.5347	88.3721	99.7084

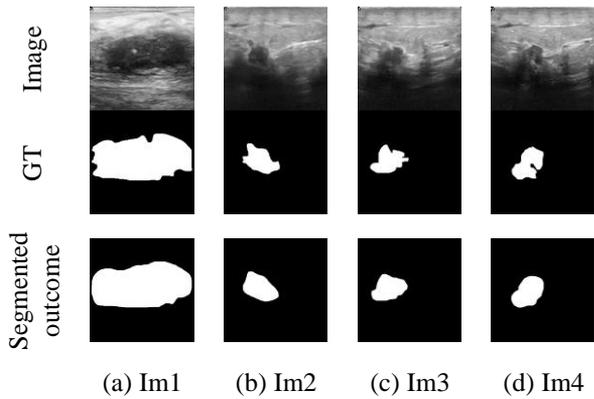


Figure 7. Sample images and the achieved outcome

With the SM classifier and deep features, Table 4 presents the classification performance of the considered deep learning schemes. Using a 5-fold cross-validation, this study confirms that ResNet18 provides a higher detection accuracy (91%) than other methods. As a result, ResNet18 is considered in this study for assessment, and other DL schemes are excluded. The ResNet18 features are then optimized using FA, and then combined with handcrafted features to produce new hybrid features, as shown in equation 1. Based on this feature vector, the proposed scheme is used to examine BC detection and the results are depicted in Table 5.

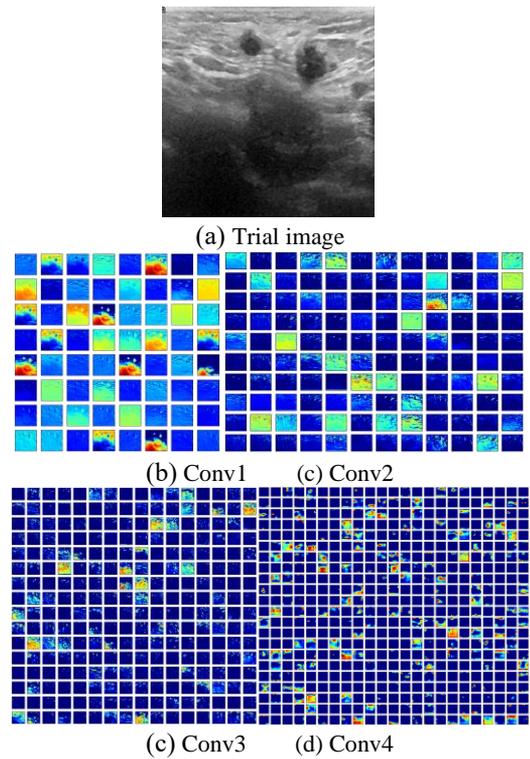


Figure 8. Convolutional layer outcomes of ResNet18 for a chosen sample test image

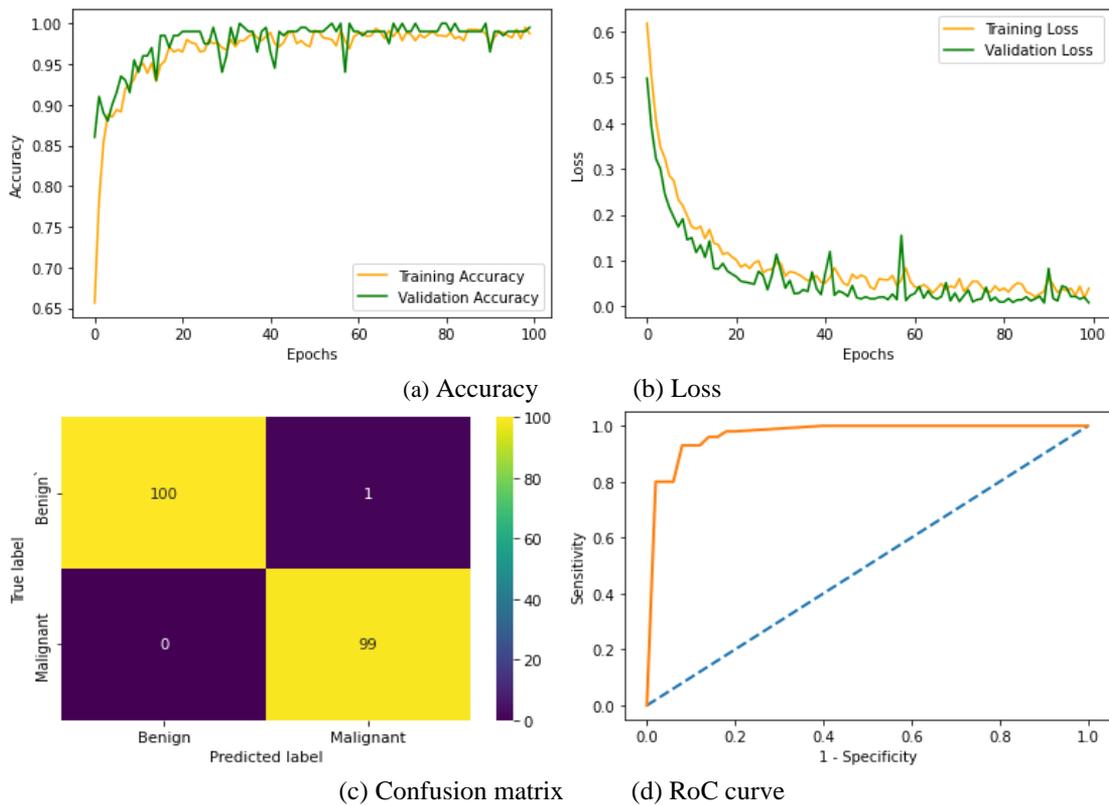


Figure 9. Obtained result with the hybrid image features

Table 4. Performance evaluation of pre-trained deep-learning schemes

Method	TP	FN	TN	FP	AC	PR	SE	SP	FS
AlexNet	89	11	87	13	88.00	87.25	89.00	87.00	88.12
VGG16	90	10	88	12	89.00	88.23	90.00	88.00	89.11
VGG19	89	11	89	11	89.00	89.00	89.00	89.00	89.00
ResNet18	90	10	92	08	91.00	91.83	90.00	92.00	90.91
ResNet50	88	12	91	09	89.50	90.72	88.00	91.00	89.34
ResNet101	89	11	88	12	88.50	88.12	89.00	88.00	88.56

Table 5. Performance assessment of hybrid feature based classification of UI to detect BC

Method	TP	FN	TN	FP	AC	PR	SE	SP	FS
SM	92	8	94	6	93.00	93.88	92.00	94.00	92.93
NB	91	9	95	5	93.00	94.79	91.00	95.00	92.86
DT	94	6	96	4	95.00	95.91	94.00	96.00	94.95
RF	95	5	97	3	96.00	96.94	95.00	97.00	95.96
KNN	99	0	100	1	99.50	99.01	100	99.01	99.50
SVM-L	98	2	98	2	98.00	98.00	98.00	98.00	98.00
SVM-RBF	97	3	99	1	98.00	98.98	97.00	99.00	97.98

Table 6. Performance comparison of the proposed scheme with existing works

Method	Accuracy (%)
Meraj et al. [21]	98.61
Irfan et al. [22]	98.97
Jabeen et al. [23]	99.10
Sahu et al. [58]	98.13
Cruz-Ramos et al. [59]	97.60
Raza et al. [60]	99.35
Proposed work	99.50

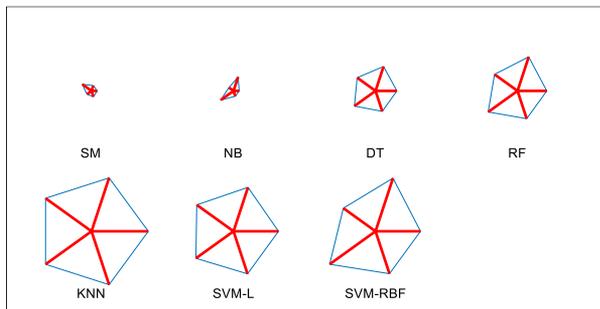
**Figure 10.** Glyph-Plot to show the overall performance of classifiers

Figure 8 shows the processed images collected from each convolution layer. Figure 8(a) shows the sample test image for demonstration, and Figure 8(b) to (e) shows the results achieved from each convolution layer. The KNN classifier's outcome is shown in Figure 9. Figure 9(a) and (b) show the accuracy and loss function during training and validation. This result confirms the merit of the proposed scheme. Figure 9(c) shows the confusion matrix, and Figure 9(d) shows the RoC curve. This table also demonstrates the effectiveness of the proposed scheme in detecting BC from UI with improved overall performance when compared with other binary classifiers. As shown in Figure 10, the KNN performs better than other study classifiers. The significance of the proposed methodology is confirmed based on the comparative assessment of earlier works in the literature.

A comparison of the best results from the present study (KNN-classifier) against the best results of the earlier works is shown in Table 6, and this comparison confirms that the proposed scheme provides better result compared to the

chosen earlier studies. In the future, the proposed research work can be considered to examine the breast UI collected from the hospitals.

5. DISCUSSIONS

Using the ResUNet supported scheme to examine BC data from the user interface, the proposed research integrates CNN segmentation and classification to improve the accuracy of the analysis. As a first step, CNN segmentation is implemented in Figure 6 and the results are presented. The proposed scheme results in a better training and validation outcome, since it extracts the BC section from the user interface in both the benign and malignant classes accurately, as shown in Figure 7. A binary classification accuracy of 99.50% is achieved when the KNN classifier is used in the classification task implemented with the ResNet18 scheme using the deep, DWT, GLCM and LBP features, as shown in Figure 8, for the classification task implemented with the ResNet18 scheme. With other related works as illustrated in Table 6, which is also a confirmation of the merit of the proposed research work, the merit of the proposed scheme can be verified. This scheme has some limitations, such as the need to resize each image and GT to a certain size in order to fit into the CNN approach that has been considered. A major advantage of the proposed scheme is the fact that it implements a pretrained CNN model, which is a simpler method than modelling a CNN from scratch. Also, as compared to the existing schemes, the achieved results are much better.

6. CONCLUSIONS

The proposed research aims to develop an accurate scheme to classify the breast Uis into benign and malignant class using the pretrained deep learning scheme. This work integrated the CNN segmentation implemented using ResUNet and CNN classification with the ResNet18 features to achieve an improved classification accuracy. Along with the deep-features, this work also considered the well known handcrafted features, like LBP, DWT and GLCM and these features are optimized using the FA to avoid the overfitting issue. The experimental investigation of this provides a classification accuracy of 99.50% when the KNN classifier along with the

hybrid features are implemented. This study confirms that the integration of the CNN segmentation and classification helps to achieve a better classification accuracy and in future, the developed tool can be tested and verified using the real clinical images.

REFERENCES

- [1] Rajinikanth, V., Kadry, S., Nam, Y. (2021). Convolutional-neural-network assisted segmentation and SVM classification of brain tumor in clinical MRI slices. *Information Technology and Control*, 50(2): 342-356. <https://doi.org/10.5755/j01.itc.50.2.28087>
- [2] Maryada, S.K.R., Booker, W.L., Danala, G., Ha, C.A., Mudduluru, S., Hougen, D.F., Zheng, B. (2022). Applying a novel two-stage deep-learning model to improve accuracy in detecting retinal fundus images. In *Medical Imaging 2022: Computer-Aided Diagnosis*, pp. 135-142.
- [3] Chen, X., Wang, X., Zhang, K., Fung, K.M., Thai, T.C., Moore, K., Qiu, Y. (2022). Recent advances and clinical applications of deep learning in medical image analysis. *Medical Image Analysis*, 79: 102444. <https://doi.org/10.1016/j.media.2022.102444>
- [4] Fernandes, S.L., Tanik, U.J., Rajinikanth, V., Karthik, K.A. (2020). A reliable framework for accurate brain image examination and treatment planning based on early diagnosis support for clinicians. *Neural Computing and Applications*, 32(20): 15897-15908. <https://doi.org/10.1007/s00521-019-04369-5>
- [5] Li, Y., Yuan, W., Fan, M., Zheng, B., Li, L. (2022). Prediction of short-term breast cancer risk with fusion of CC-and MLO-based risk models in four-view mammograms. *Journal of Digital Imaging*, 35(4): 910-922. <https://doi.org/10.1007/s10278-019-00266-4>
- [6] <https://www.who.int/news-room/fact-sheets/detail/cancer>.
- [7] <https://www.who.int/news/item/20-10-2021-who-launches-women-s-health-chatbot-with-messaging-on-breast-cancer>.
- [8] Aghaei, F., Mirniaharikandehei, S., Hollingsworth, A.B., Stoug, R.G., Pearce, M., Liu, H., Zheng, B. (2018). Association between background parenchymal enhancement of breast MRI and BIRADS rating change in the subsequent screening. In *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*, p. 105790R. <https://doi.org/10.1117/12.2288001>
- [9] Kadry, S., Damaševičius, R., Taniar, D., Rajinikanth, V., Lawal, I.A. (2021). Extraction of tumour in breast MRI using joint thresholding and segmentation—A study. In *2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII)*, pp. 1-5. <https://doi.org/10.1109/ICBSII51839.2021.9445152>
- [10] Raja, N., Rajinikanth, V., Fernandes, S.L., Satapathy, S.C. (2017). Segmentation of breast thermal images using Kapur's entropy and hidden Markov random field. *Journal of Medical Imaging and Health Informatics*, 7(8): 1825-1829.
- [11] EtehadTavakol, M., Chandran, V., Ng, E.Y.K., Kafieh, R. (2013). Breast cancer detection from thermal images using bispectral invariant features. *International Journal of Thermal Sciences*, 69: 21-36. <https://doi.org/10.1016/j.ijthermalsci.2013.03.001>
- [12] Mirniaharikandehei, S., VanOsdol, J., Heidari, M., Danala, G., Sethuraman, S.N., Ranjan, A., Zheng, B. (2019). Developing a quantitative ultrasound image feature analysis scheme to assess tumor treatment efficacy using a mouse model. *Scientific reports*, 9(1): 1-10. <https://doi.org/10.1038/s41598-019-43847-7>
- [13] MacMahon, B. (2006). Epidemiology and the causes of breast cancer. *International Journal of Cancer*, 118(10): 2373-2378. <https://doi.org/10.1002/ijc.21404>
- [14] Key, T.J., Verkasalo, P.K., Banks, E. (2001). Epidemiology of breast cancer. *The Lancet Oncology*, 2(3): 133-140. [https://doi.org/10.1016/S1470-2045\(00\)00254-0](https://doi.org/10.1016/S1470-2045(00)00254-0)
- [15] Danala, G., Patel, B., Aghaei, F., Heidari, M., Li, J., Wu, T., Zheng, B. (2018). Classification of breast masses using a computer-aided diagnosis scheme of contrast enhanced digital mammograms. *Annals of Biomedical Engineering*, 46(9): 1419-1431. <https://doi.org/10.1007/s10439-018-2044-4>
- [16] Rajinikanth, V., Kadry, S., Taniar, D., Damaševičius, R., Rauf, H.T. (2021). Breast-cancer detection using thermal images with marine-predators-algorithm selected features. In *2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*, pp. 1-6. <https://doi.org/10.1109/ICBSII51839.2021.9445166>
- [17] Dey, N., Rajinikanth, V., Hassaniien, A.E. (2021). An examination system to classify the breast thermal images into early/acute DCIS class. In *Proceedings of International Conference on Data Science and Applications*, pp. 209-220. [10.1007/978-981-15-7561-7_17](https://doi.org/10.1007/978-981-15-7561-7_17)
- [18] Fernandes, S.L., Rajinikanth, V., Kadry, S. (2019). A hybrid framework to evaluate breast abnormality using infrared thermal images. *IEEE Consumer Electronics Magazine*, 8(5): 31-36. <https://doi.org/10.1109/MCE.2019.2923926>
- [19] Thanaraj, R., Anand, B., Allen Rahul, J., Rajinikanth, V. (2020). Appraisal of breast ultrasound image using Shannon's thresholding and level-set segmentation. In *Progress in Computing, Analytics and Networking*, 621-630. https://doi.org/10.1007/978-981-15-2414-1_62
- [20] Ilesanmi, A.E., Chaumrattanakul, U., Makhnov, S.S. (2021). A method for segmentation of tumors in breast ultrasound images using the variant enhanced deep learning. *Biocybernetics and Biomedical Engineering*, 41(2): 802-818. <https://doi.org/10.1016/j.bbe.2021.05.007>
- [21] Meraj, T., Alosaimi, W., Alouffi, B., Rauf, H.T., Kumar, S.A., Damaševičius, R., Alyami, H. (2021). A quantization assisted U-Net study with ICA and deep features fusion for breast cancer identification using ultrasonic data. *PeerJ Computer Science*, 7: e805.
- [22] Irfan, R., Almazroi, A.A., Rauf, H.T., Damaševičius, R., Nasr, E.A., Abdelgawad, A.E. (2021). Dilated semantic segmentation for breast ultrasonic lesion detection using parallel feature fusion. *Diagnostics*, 11(7): 1212. <https://doi.org/10.3390/diagnostics11071212>
- [23] Jabeen, K., Khan, M.A., Alhaisoni, M., Tariq, U., Zhang, Y.D., Hamza, A., Damaševičius, R. (2022). Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion. *Sensors*, 22(3): 807. <https://doi.org/10.3390/s22030807>
- [24] Vijayakumar, K., Rajinikanth, V., Kirubakaran, M.K.

- (2022). Automatic detection of breast cancer in ultrasound images using Mayfly algorithm optimized handcrafted features. *Journal of X-Ray Science and Technology*, (Preprint), 1-16. <https://doi.org/10.3233/XST-221136>
- [25] Zhou, Y., Chen, H., Li, Y., Liu, Q., Xu, X., Wang, S., Yap, P.T., Shen, D. (2021). Multi-task learning for segmentation and classification of tumors in 3D automated breast ultrasound images. *Medical Image Analysis*, 70: 101918. <https://doi.org/10.1016/j.media.2020.101918>
- [26] Lei, Y., He, X., Yao, J., Wang, T., Wang, L., Li, W., Curran, W.J., Liu, T., Xu, D., Yang, X. (2021). Breast tumor segmentation in 3D automatic breast ultrasound using Mask scoring R-CNN. *Medical Physics*, 48(1): 204-214. <https://doi.org/10.1002/mp.14569>
- [27] Sehgal, C.M., Weinstein, S.P., Arger, P.H., Conant, E.F. (2006). A review of breast ultrasound. *Journal of Mammary Gland Biology and Neoplasia*, 11(2): 113-123. <https://doi.org/10.1007/s10911-006-9018-0>
- [28] Al-Dhabyani, W., Gomaa, M., Khaled, H., Fahmy, A. (2020). Dataset of breast ultrasound images. *Data in Brief*, 28: 104863. <https://doi.org/10.1016/j.dib.2019.104863>
- [29] Jones, M.A., Pham, H., Gai, T., Zheng, B. (2022). Fusion of handcrafted and deep transfer learning features to improve performance of breast lesion classification. In *Medical Imaging 2022: Computer-Aided Diagnosis*, 12033: 682-693. <https://doi.org/10.1117/12.2611607>
- [30] Jones, M.A., Faiz, R., Qiu, Y., Zheng, B. (2022). Improving mammography lesion classification by optimal fusion of handcrafted and deep transfer learning features. *Physics in Medicine & Biology*, 67(5): 054001. <https://doi.org/10.1088/1361-6560/ac5297>
- [31] Dey, N., Rajinikanth, V., Fong, S.J., Kaiser, M.S., Mahmud, M. (2020). Social group optimization-assisted Kapur's entropy and morphological segmentation for automated detection of COVID-19 infection from computed tomography images. *Cognitive Computation*, 12(5): 1011-1023. <https://doi.org/10.1007/s12559-020-09751-3>
- [32] Kadry, S., Rajinikanth, V., González Crespo, R., Verdú, E. (2022). Automated detection of age-related macular degeneration using a pre-trained deep-learning scheme. *The Journal of Supercomputing*, 78(5): 7321-7340. <https://doi.org/10.1007/s11227-021-04181-w>
- [33] Kadry, S., Rajinikanth, V., Taniar, D., Damaševičius, R., Valencia, X.P.B. (2022). Automated segmentation of leukocyte from hematological images-A study using various CNN schemes. *The Journal of Supercomputing*, 78(5): 6974-6994. <https://doi.org/10.1007/s11227-021-04125-4>
- [34] Wang, Y., Chen, Y., Yang, N., Zheng, L., Dey, N., Ashour, A.S., Rajinikanth, V., Tavares, J.M.R.S., Shi, F. (2019). Classification of mice hepatic granuloma microscopic images based on a deep convolutional neural network. *Applied Soft Computing*, 74: 40-50. <https://doi.org/10.1016/j.asoc.2018.10.006>
- [35] Wang, Y., Shi, F., Cao, L., Dey, N., Wu, Q., Ashour, A. S., Sherratt, R.S., Rajinikanth, V., Wu, L. (2019). Morphological segmentation analysis and texture-based support vector machines classification on mice liver fibrosis microscopic images. *Current Bioinformatics*, 14(4): 282-294. <https://doi.org/10.2174/1574893614666190304125221>
- [36] Diakogiannis, F.I., Waldner, F., Caccetta, P., Wu, C. (2020). ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162: 94-114. <https://doi.org/10.1016/j.isprsjprs.2020.01.013>
- [37] Jha, D., Smedsrud, P.H., Riegler, M.A., Johansen, D., De Lange, T., Halvorsen, P., Johansen, H.D. (2019). Resunet++: An advanced architecture for medical image segmentation. In *2019 IEEE International Symposium on Multimedia (ISM)*, pp. 225-225. <https://doi.org/10.1109/ISM46123.2019.00049>
- [38] Ramkumar, M., Babu, C.G., Kumar, K.V., Hepsiba, D., Manjunathan, A., Kumar, R.S. (2021). ECG cardiac arrhythmias classification using DWT, ICA and MLP neural networks. In *Journal of Physics: Conference Series*, 1831(1): 012015. <https://doi.org/10.1088/1742-6596/1831/1/012015>
- [39] Dhingra, G., Kumar, V., Joshi, H.D. (2021). Quality assessment of leaves quality using texture and DWT based local feature extraction analysis. *Chemometrics and Intelligent Laboratory Systems*, 208: 104195. <https://doi.org/10.1016/j.chemolab.2020.104195>
- [40] Haweel, R., Shalaby, A., Mahmoud, A., Seada, N., Ghoniemy, S., Ghazal, M., Casanova, M.F., Barnes, G.N., El-Baz, A. (2021). A robust DWT-CNN-based CAD system for early diagnosis of autism using task-based fMRI. *Medical Physics*, 48(5): 2315-2326. <https://doi.org/10.1002/mp.14692>
- [41] Muhathir, M., Santoso, M.H., Larasati, D.A. (2021). Wayang image classification using SVM method and GLCM feature extraction. *Journal of Informatics and Telecommunication Engineering*, 4(2): 373-382. <https://doi.org/10.31289/jite.v4i2.4524>
- [42] Yogeshwari, M., Thailambal, G. (2021). Automatic feature extraction and detection of plant leaf disease using GLCM features and convolutional neural networks. *Materials Today: Proceedings*, 81(2): 530-536. <https://doi.org/10.1016/j.matpr.2021.03.700>
- [43] Gudigar, A., Raghavendra, U., Devasia, T., Nayak, K., Danish, S.M., Kamath, G., Samanth, J., Pai, U.M., Nayak, V., Tan, R.S., Ciaccio, E.J., Acharya, U.R. (2019). Global weighted LBP based entropy features for the assessment of pulmonary hypertension. *Pattern Recognition Letters*, 125: 35-41. <https://doi.org/10.1016/j.patrec.2019.03.027>
- [44] Lakshmi, D., Ponnusamy, R. (2021). Facial emotion recognition using modified HOG and LBP features with deep stacked autoencoders. *Microprocessors and Microsystems*, 82: 103834. <https://doi.org/10.1016/j.micpro.2021.103834>
- [45] Yang, X.S. (2010). Firefly algorithm, stochastic test functions and design optimisation. *International Journal of Bio-Inspired Computation*, 2(2): 78-84. <https://doi.org/10.1504/IJBIC.2010.032124>
- [46] Rajinikanth, V., Raja, N.S.M., Kamalanand, K. (2017). Firefly algorithm assisted segmentation of tumor from brain MRI using Tsallis function and Markov random field. *Journal of Control Engineering and Applied Informatics*, 19(3): 97-106.
- [47] Raja, N.S.M., Manic, K.S., Rajinikanth, V. (2013). Firefly algorithm with various randomization parameters: an analysis. In *International Conference on Swarm, Evolutionary, and Memetic Computing*, pp. 110-121.

- https://doi.org/10.1007/978-3-319-03753-0_11
- [48] Bakiya, A., Kamalanand, K., Rajinikanth, V. (2021). Automated diagnosis of amyotrophic lateral sclerosis using electromyograms and firefly algorithm based neural networks with fractional position update. *Physical and Engineering Sciences in Medicine*, 44(4): 1095-1105. [10.1007/s13246-021-01046-7](https://doi.org/10.1007/s13246-021-01046-7)
- [49] Khan, M.A., Rajinikanth, V., Satapathy, S.C., Taniar, D., Mohanty, J.R., Tariq, U., Damaševičius, R. (2021). VGG19 network assisted joint segmentation and classification of lung nodules in CT images. *Diagnostics*, 11(12): 2208. <https://doi.org/10.3390/diagnostics11122208>
- [50] Arunmozhi, S., Sarojini, V.S.S., Pavithra, T., Varghese, V., Deepti, V., Rajinikanth, V. (2021). Automated detection of COVID-19 lesion in lung CT slices with VGG-UNet and handcrafted features. In *Digital Future of Healthcare*, pp. 185-200. <https://doi.org/10.1201/9781003198796-11>
- [51] Rajakumar, M.P., Sonia, R., Uma Maheswari, B., Karuppiah, S.P. (2021). Tuberculosis detection in chest X-ray using Mayfly-algorithm optimized dual-deep-learning features. *Journal of X-Ray Science and Technology*, 29(6): 961-974. <https://doi.org/10.3233/XST-210976>
- [52] Danala, G., Ray, B., Desai, M., Heidari, M., Mirniaharikandehi, S., Maryada, S.K.R., Zheng, B. (2022). Developing new quantitative CT image markers to predict prognosis of acute ischemic stroke patients. *Journal of X-Ray Science and Technology*, 30(3): 459-475. <https://doi.org/10.3233/XST-221138>
- [53] Zuo, T., Zheng, Y., He, L., Chen, T., Zheng, B., Zheng, S., You, J.H., Li, X.Y., Liu, R., Bai, J.J., Si, S.X., Wang, Y.Y., Zhang, S.Y., Wang, L.L., Chen, J. (2021). Automated classification of papillary renal cell carcinoma and chromophobe renal cell carcinoma based on a small computed tomography imaging dataset using deep learning. *Frontiers in Oncology*, 11: 4820. <https://doi.org/10.3389/fonc.2021.746750>
- [54] Ghani, M.U., Fajardo, L.L., Omoumi, F., Yan, A., Jenkins, P., Wong, M., Li, Y., Peterson, M.E., Callahan, E.J., Hillis, S.L., Zheng, B., Wu, X.Z., Liu, H. (2021). A phase sensitive x-ray breast tomosynthesis system: Preliminary patient images with cancer lesions. *Physics in Medicine & Biology*, 66(21): 21LT01. <https://doi.org/10.1088/1361-6560/ac2ea6>
- [55] Ahuja, S., Panigrahi, B.K., Dey, N., Rajinikanth, V., Gandhi, T.K. (2021). Deep transfer learning-based automated detection of COVID-19 from lung CT scan slices. *Applied Intelligence*, 51(1): 571-585. <https://doi.org/10.1007/s10489-020-01826-w>
- [56] Yurttakal, A.H., Erbay, H., İközceli, T., Karaçavuş, S. (2020). Detection of breast cancer via deep convolution neural networks using MRI images. *Multimedia Tools and Applications*, 79(21): 15555-15573. [10.1007/s11042-019-7479-6](https://doi.org/10.1007/s11042-019-7479-6)
- [57] Yurttakal, A.H., Erbay, H., İközceli, T., Karaçavuş, S., Biçer, C. (2022). Diagnosing breast cancer tumors using stacked ensemble model. *Journal of Intelligent & Fuzzy Systems*, 42(1): 77-85. [10.3233/JIFS-219176](https://doi.org/10.3233/JIFS-219176)
- [58] Sahu, A., Das, P.K., Meher, S. (2023). High accuracy hybrid CNN classifiers for breast cancer detection using mammogram and ultrasound datasets. *Biomedical Signal Processing and Control*, 80: 104292. <https://doi.org/10.1016/j.bspc.2022.104292>
- [59] Cruz-Ramos, C., García-Avila, O., Almaraz-Damian, J. A., Ponomaryov, V., Reyes-Reyes, R., Sadovnychiy, S. (2023). Benign and malignant breast tumor classification in ultrasound and mammography images via fusion of deep learning and handcraft features. *Entropy*, 25(7): 991. <https://doi.org/10.3390/e25070991>
- [60] Raza, A., Ullah, N., Khan, J.A., Assam, M., Guzzo, A., Aljuaid, H. (2023). DeepBreastCancerNet: A novel deep learning model for breast cancer detection using ultrasound images. *Applied Sciences*, 13(4): 2082. <https://doi.org/10.3390/app13042082>