

Multiscale Analysis Domain Interpretable Deep Neural Network for Detection of Breast Cancer Using Thermogram Images

Anurag Gade^{ID}, Dinesh Kumar Dash^{ID}, Member, IEEE, T. Mita Kumari^{ID}, Samit Kumar Ghosh^{ID}, Rajesh Kumar Tripathy^{ID}, Member, IEEE, and Ram Bilas Pachori^{ID}, Senior Member, IEEE

Abstract—Breast cancer is the most prevalent cancer among women, with a high mortality rate. The early detection of breast cancer using medical imaging techniques helps reduce the number of deaths caused by this disease. Thermogram imaging is safer and less expensive than mammography for diagnosing breast cancer. The automated analysis of thermogram images using artificial intelligence (AI) methods is an interesting approach to detect breast cancer. This article proposes a novel multiscale analysis domain interpretable deep learning (MSADIDL) approach for automatically detecting breast cancer using thermogram images. The 2D empirical wavelet transform (2DEWT) with fixed boundary points (FBPs) is employed for the multiscale analysis of thermogram images and evaluation of modes or subbands. All the modes of the thermogram images are used as the input to the MSADIDL model for the automated detection of breast cancer. The MSADIDL architecture comprises seven individual deep neural networks (DNNs) connected in parallel. The outputs of the individual DNNs are concatenated and then used as the input to the dense layers, after which the output layer evaluates the probability score for the automated categorization of normal versus cancerous classes. A publicly available thermogram imaging dataset is utilized to evaluate the performance of the proposed MSADIDL approach. The results show that the proposed MSADIDL approach has obtained an accuracy value of 99.54% for both fivefold cross-validation (CV) and hold-out validation cases using all seven modes of thermogram images. The MSADIDL model has achieved an accuracy higher than all of the transfer learning-based breast cancer detection techniques using thermogram images. The suggested MSADIDL model has shown higher accuracy when compared with different existing methods to detect breast cancer using thermogram images.

Index Terms—Accuracy, breast cancer, deep learning (DL), multiscale analysis, thermogram images.

Manuscript received 23 May 2023; revised 5 August 2023; accepted 7 September 2023. Date of publication 22 September 2023; date of current version 3 October 2023. The Associate Editor coordinating the review process was Dr. Carlos G. Juan. (*Corresponding author:* Rajesh Kumar Tripathy.)

Anurag Gade and Rajesh Kumar Tripathy are with the Department of Electronics and Electrical Engineering, BITS Pilani, Hyderabad 500078, India (e-mail: f20201337@hyderabad.bits-pilani.ac.in; rajeshiiitg13@gmail.com).

Dinesh Kumar Dash and T. Mita Kumari are with the Department of Electronics and Telecommunication Engineering, Parala Maharaja Engineering College, Berhampur, Odisha 761003, India (e-mail: dinesh.etc@pmec.ac.in; mitat2006@gmail.com).

Samit Kumar Ghosh is with the Department of Computer Application, Techno India University, Kolkata, West Bengal 700091, India (e-mail: samitnitrkl@gmail.com).

Ram Bilas Pachori is with the Department of Electrical Engineering, Indian Institute of Technology Indore, Indore 453552, India (e-mail: pachori@iiti.ac.in).

Digital Object Identifier 10.1109/TIM.2023.3317913

I. INTRODUCTION

BREAST cancer occurs when the breast cells grow abnormally and create a tumor mass [1]. The symptoms of this disease are changes in the skin of the breast, the feeling of a lump, and variations in the size of the breast [2]. The early detection of breast cancer helps to decide the treatment for the better management of this disease [3]. Various imaging techniques, such as mammography [4], sonography (SG) [5], magnetic resonance imaging (MRI) [6], nuclear medicine imaging (NMR) [7], and thermography [8], are used to diagnose breast cancer. Mammography uses low-energy X-rays to produce images that have been used to detect calcifications, cysts, and other breast abnormalities [4], [9]. Mammography is considered the initial screening for the diagnosis of breast cancer. Regular mammography tests cause radiation exposure to the breast tissue and thereby increase the risk of breast cancer [10]. Also, the mammography imaging technique causes discomfort to the patients as it involves the compression of breasts during the test [10]. Other imaging techniques, such as SG, MRI, and NMR, are costly and not readily available in medical centers for the initial screening of breast cancer. On the other hand, thermography is a noncontact and noninvasive diagnosis method used to detect breast cancer [8]. Thermography is based on the change in the temperature on the breast skin surface, producing a thermogram image used to detect breast cancer [8]. The healthcare professional manually inspects the thermogram image changes for the early diagnosis of breast cancer. This manual procedure may result in a diagnostic error; hence, artificial intelligence (AI) based methods are used for automated breast cancer diagnosis from thermogram images [11]. The AI-based techniques for diagnosing breast cancer are formulated using either image processing coupled with machine learning (ML) or deep learning (DL) approaches [11]. The DL-based approach directly used the thermogram image as the input to the network for detecting breast cancer. The major drawback of the DL-based approach is that it is still unknown to get the type of learned features in the hidden layers. There is a lack of interpretability of the DL-based model for the automated detection of breast cancer using medical images. The development of novel interpretable DL-based techniques for the automated detection of breast cancer from thermogram images is helpful for medical image processing applications.

In the last decade, various approaches have been reported to automatically detect breast cancer using thermogram images [11], [12]. Milosevic et al. [13] have extracted gray-level co-occurrence matrices (GLCM) based texture features from thermogram images and used the K-nearest neighbor (KNN) classifier to detect breast cancer. In another study, Pramanik et al. [14] have evaluated statistical and wavelet domain features from thermogram images and used the multilayer perceptron (MLP) model to detect breast cancer. Similarly, Acharya et al. [15] have computed the texture features from thermogram images. They have used the support vector machine (SVM) classifier to automatically detect breast cancer using the texture features. Francis et al. [16] have extracted statistical features using the curvelet transform domain representation of thermogram images and employed the SVM model for the automated identification of breast cancer. Abdel-Nasser et al. [17] have used representation learning and texture analysis-based techniques to extract features from the thermogram images. They have used the MLP-based classifier to detect breast cancer. Resmini et al. [18] have computed energy features from the wavelet subbands of normalized thermogram images and used the genetic algorithm to select features. The SVM classifier has been used to detect breast cancer using the selected features of thermogram images. Satish et al. [19] have extracted DWT domain local energy features from normalized thermogram images and used the random subset feature selection (RSFS) method to choose relevant features. The SVM classifier has been utilized to detect breast cancer using the RSFS-based reduced feature set of normalized thermogram images. The performance of the aforementioned automated methods for detecting breast cancer is highly dependent on various steps, such as extraction of features from the thermogram images, selecting relevant features using different algorithms, and using various supervised learning techniques. In recent years, various DL-based methods have been explored for the automated identification of breast cancer using thermogram images [8], [20]. Among them, Dey et al. [21] have applied the DenseNet121-based deep transfer learning model in the edge-detected thermogram images for the automated identification of breast cancer. Similarly, Zuluaga-Gomez et al. [22] have considered a deep CNN model to detect breast cancer using thermogram images. They have compared the performance of their suggested deep CNN model with various transfer learning models in terms of accuracy measures for breast cancer detection. The shortcomings of these approaches in detecting breast cancer using thermogram images are that these techniques are not interpretable regarding the type of features produced in different layers of the DL-based model. Therefore, there is scope to develop novel interpretable DL-based methods for automated breast cancer screening using thermogram images.

The 2D empirical wavelet transform (2DEWT) is a multi-scale domain analysis technique to evaluate modes or subbands from the original image [23]. The modes obtained in the 2DEWT stage capture the segregated information of the original image at different frequency scales. In 2DEWT, the modes are computed based on the filter bank designed

using the boundary points of the mean spectral radius vector of the original image's pseudo-polar fast Fourier transform (PPFFT) [23]. Muralidharan et al. [24] have suggested the fixed boundary points (FBPs) to segregate the mean spectral radius vector of the X-ray image into the contiguous segment to design the 2DEWT filter bank. The FBPs helped reduce the computational complexity of the EWT algorithms for evaluating modes from signals and images [24], [25]. Furthermore, Muralidharan et al. [24] also designed a deep neural network (DNN) model using various combinations of the 2DEWT domain modes of X-ray images to detect COVID-19. They have verified that the multiscale information of X-ray images helps to achieve a better performance of the DNN for COVID-19 detection. The 2DEWT with FBPs has not been employed for the multiscale analysis of thermogram images to detect breast cancer. The novelty of this work is to employ different scale information of thermogram images to develop an interpretable DNN for the automated detection of breast cancer. The salient contributions of this work are highlighted as follows.

- 1) The 2DEWT with FBPs is introduced in this work for the multiscale decomposition of the thermogram image into modes or subbands.
- 2) The mode-based deep feature extractor block (DFEB) is proposed to extract local learnable features from the multiscale information of the thermogram image.
- 3) The local learnable features extracted using the proposed mode-based DFEB blocks are concatenated and used as input to the different dense layers, followed by the softmax layer to detect breast cancer.
- 4) The proposed multiscale analysis domain interpretable DL (MSADIDL) method is compared with various transfer learning models and existing methods to detect breast cancer using thermogram images.

The remaining sections of this article are structured as follows. In Section II, we have described the thermogram image database used in the proposed work. The proposed MSADIDL approach is described in Section III. The results and discussion of the suggested approach are presented in Section IV. In Section V, the conclusions of this article are written.

II. THERMOGRAM IMAGE DATABASE

In this study, we have considered a publicly available thermogram image database [26] to evaluate the proposed MSADIDL approach performance for breast cancer detection. The link for downloading the thermogram image database can be found in DMRIR.¹ This database contains 1520 thermogram images obtained from 56 subjects. A total of 20 sequential thermogram images are taken from each subject at a 15-s interval. A room temperature between 20 °C and 22 °C has been maintained while capturing breast thermogram images from the subjects. The FLIR thermal camera (model: SC 620) has been used to capture the thermogram images [26]. The sensitivity and temperature range in the FLIR thermal camera is in the range of <0.04 °C and –40 °C to 500 °C, respectively. The size of each thermogram image in the

¹<http://visual.ic.uff.br/dmi/>

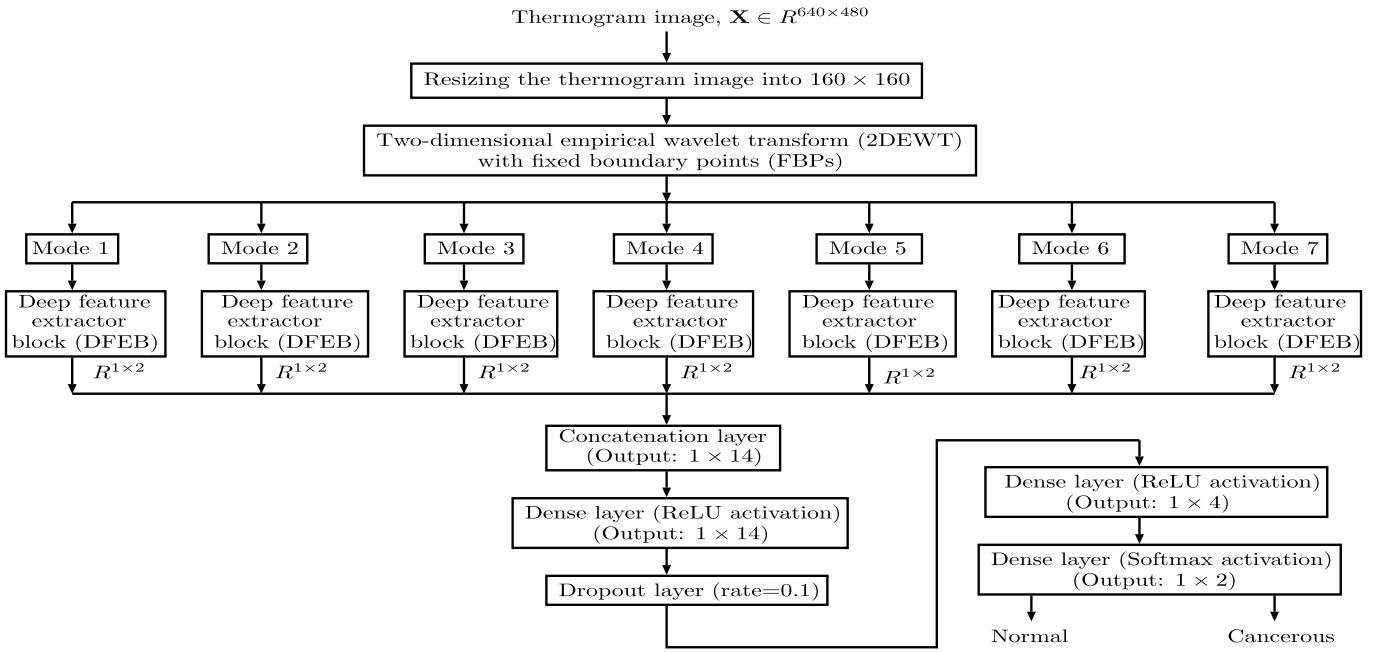


Fig. 1. Flowchart of the proposed MSADIDL model for breast cancer detection using thermogram images (the layers in DFEB are shown in Fig. 4).

database is 640×480 pixels. The region of interest (ROI) for each thermogram image has been given in the database. In this work, we have used 760 normal and 760 cancerous thermogram images to evaluate the performance of the proposed approach. In the database, for the normal class, the segmented thermogram images of both left and right breasts (380 images each) are given for 19 subjects. Similarly, for the cancerous class, 720 images are provided in the database for 36 subjects. For one subject, 40 thermogram images are given for the cancerous class.

Moreover, we have also used a publicly available ultrasound image database² to verify the effectiveness of the proposed MSADIDL approach to detect benign and malignant tumors. A total of 2000 ultrasound images (1000 images from the benign class and 1000 images from the malignant class) are used in this work. Each ultrasound image has a size of 224×224 . The ultrasound images are resized to 160×160 for the classification of benign and malignant classes using the MSADIDL model.

III. PROPOSED METHOD

The flowchart of the suggested multiscale domain DNN approach to detect breast cancer using the thermogram image is shown in Fig. 1. It consists of resizing the thermogram image, multiscale decomposition of the thermogram image into modes, DFEB, and additional dense layers. The size of the thermogram image in the database is 640×480 . In this study, we have resized each thermogram image into 160×160 image for further processing. The resizing of the thermogram image helps reduce the computational complexity of the MSADIDL model for breast cancer detection.

²<https://www.kaggle.com/datasets/vuppalaadithyasairam/ultrasound-breast-images-for-breast-cancer?resource=download>.

A. Multiscale Analysis of Thermogram Images

In this work, the 2DEWT with FBPs is used for the multiscale analysis of the thermogram images. The multiscale domain components or modes of thermogram images are evaluated using five steps in 2DEWT. The first step is to evaluate the thermogram image's PPFFT. The PPFFT of the thermogram image, $X(n_1, n_2)$ is denoted as $Z(\theta, |f|)$, where θ and $|f|$ are phase angle and spectral radius, respectively [23]. The n_1 and n_2 vary in the range as $n_1 = [0, 1, 2, \dots, H]$, and $n_2 = [0, 1, 2, \dots, W]$, respectively. The parameters H and W are the height and width of the thermogram image. In the next step, the mean spectral radius in the PPFFT domain of the thermogram image is evaluated, and it is denoted as $\tilde{z}(|f|)$. The mean spectral radius vector is given as follows [24]:

$$\tilde{z}(|f|) = \frac{1}{K} \sum_{i=0}^K z(\theta_i, |f|) \quad (1)$$

where K is interpreted as the total number of phase angles in the PPFFT matrix of the thermogram image. The normalized spectral radius vector with fixed frequency points of the thermogram image is shown in Fig. 2. The FBPs segregate the mean spectral radius into contiguous segments, and furthermore, these segments are used to design 2-D empirical wavelets to evaluate modes from the thermal images [23], [27]. The FBPs are evaluated from the frequency points in the mean spectral radius vector, $\tilde{z}(|f|)$. In this study, we have used the frequency points (f_p) such as $[4, 8, 16, 24, 32, 40]$ in the mean spectral radius of the thermogram image. The FBPs (B_p) are evaluated as $B_p = ((2\pi f_p)/N)$, where N is the length of the mean spectral radius vector of the thermal image and f_p is the p th frequency point. In the fourth step, the 2-D empirical scaling and wavelet functions are computed using the contiguous segments extracted from the mean spectral radius vector. The first segment, $[0, B_1]$ is used to design the

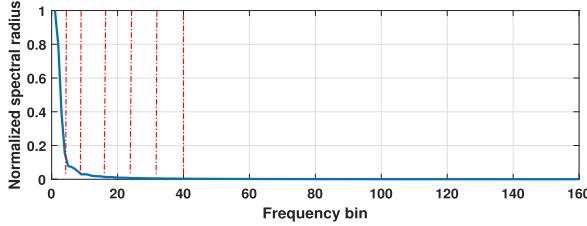


Fig. 2. Normalized mean spectral radius vector with FBP for thermogram image.

empirical scaling function, and it is given as follows [23]:

$$\phi_s^1(f) = \begin{cases} 1; & \text{if } |f| \leq (1 - \eta)\beta_1 \\ \cos\left[\frac{\pi}{2}\alpha\left(\frac{1}{2\eta\beta_1}(|f| - (1 - \eta)\beta_1)\right)\right] & \text{if } (1 - \eta)\beta_1 \leq |f| \leq (1 + \eta)\beta_1 \\ 0; & \text{elsewhere} \end{cases} \quad (2)$$

where, $|f| = (f_1^2 + f_2^2)^{1/2}$, $f_1 = ((2\pi(k_1 - (H/2))/H)$, and $f_2 = ((2\pi(k_2 - (W/2))/W)$, respectively. The f_1 , and f_2 are the frequency bins to evaluate 2-D empirical scaling function. The index k_1 , and k_2 are given as $k_1 = [0, 1, 2, \dots, W - 1]$, $k_2 = [0, 1, 2, \dots, H - 1]$, respectively. The factor $|f|$ is the spectral radius which is the same as that evaluated using PPFFFT of the thermogram image. Similarly, the empirical wavelet function designed using j th boundary range ($j = 2, 3, \dots, N - 1$) are given as follows [23]:

$$\psi_w^j(f) = \begin{cases} 1; & \text{if } (1 + \eta)\beta_j \leq |f| \leq (1 - \eta)\beta_{j+1} \\ \cos\left[\frac{\pi}{2}\alpha\left(\frac{1}{2\eta\beta_{j+1}}(|f| - (1 - \eta)\beta_{j+1})\right)\right] & \text{if } (1 - \eta)\beta_{j+1} \leq |f| \leq (1 + \eta)\beta_{j+1} \\ \sin\left[\frac{\pi}{2}\alpha\left(\frac{1}{2\eta\beta_j}(|f| - (1 - \eta)\beta_j)\right)\right] & \text{if } (1 - \eta)\beta_j \leq |f| \leq (1 + \eta)\beta_j \\ 0; & \text{elsewhere.} \end{cases} \quad (3)$$

Similarly, the empirical wavelet function defined in the boundary range $[\beta_N, \pi]$ is given as follows:

$$\psi_w^N(f) = \begin{cases} 1; & \text{if } (1 + \eta)\beta_N \leq |f| \\ \sin\left[\frac{\pi}{2}\alpha\left(\frac{1}{2\eta\beta_N}(|f| - (1 - \eta)\beta_N)\right)\right] & \text{if } (1 - \eta)\beta_N \leq |f| \leq (1 + \eta)\beta_N \\ 0; & \text{elsewhere} \end{cases} \quad (4)$$

where f spreads in both direction as f_1 , and f_2 in frequency domain. Mode 1 or subband image 1 is evaluated as follows [23], [24]:

$$M_1(n_1, n_2) = \text{FFT2}^{-1}[\text{FFT2}(X(n_1, n_2))\bar{\phi}_s^1(f)]. \quad (5)$$

Similarly, the j th detail subband image or mode is evaluated as follows [24]:

$$M_j(n_1, n_2) = \text{FFT2}^{-1}[\text{FFT2}(X(n_1, n_2))\bar{\psi}_w^j(f)]. \quad (6)$$

The last mode is computed as

$$M_N(n_1, n_2) = \text{FFT2}^{-1}[\text{FFT2}(X(n_1, n_2))\bar{\psi}_w^N(f)] \quad (7)$$

where FFT2 and FFT2^{-1} are the 2-D fast Fourier transform and its inverse, respectively. The normal and cancerous thermogram images are depicted in Fig. 3(a) and (i), respectively. The modes of the normal and cancerous thermogram images are shown in Fig. 3(b)–(h) and (j)–(p), respectively. It is observed that the morphology of thermal images is different for normal and cancerous classes. The modes computed using the 2DEWT phase capture the segregated information of thermogram images for normal and cancerous classes. The first three modes capture low-frequency information, whereas the mid-band frequency information has been captured using mode 4 and mode 5 of the thermogram images. Similarly, mode 6 and mode 7 contain the high-frequency information of the thermogram images. Investigating the information of each scale or local information of the thermogram image helps in the better diagnosis of breast cancer as compared to the global information of the thermogram image. Hence, the DL model developed using different mode-based information of thermogram images can be used to detect breast cancer.

B. Multiscale Domain Interpretable DNN

In this work, we have proposed the MSADIDL model to detect breast cancer using modes of thermogram images. The overall architecture of the model is shown in Fig. 1. A DFEB is employed from each mode of the thermogram image to evaluate learnable mode-based local features. Furthermore, these mode-based local features are concatenated and used as input to dense layers for feature vector size reduction, followed by the classification of normal and cancerous types in the softmax layers. The architecture of the DFEB used in each mode is shown in Fig. 4. It consists of three convolution layers, three max-pooling layers, one dropout layer, one batch normalization (BN) layer [28], one global average pooling (GAP) layer [29], one gated recurrent unit (GRU) followed by attention layers, and three dense layers, respectively. The feature map produced in the convolution layer is given as follows [24]:

$$F_{l_1, l_2}^t = h\left[\left(\sum_{p=1}^u \sum_{q=1}^v S_{p,q}^t Z_{l_1+p-1, l_2+q-1}\right) + b^t\right] \quad (8)$$

where Z is the input matrix to the 2-D convolution layer and $S_{p,q}^t$ is the filter or kernel to evaluate t th feature map F_{l_1, l_2}^t . The $h(\cdot)$ stands for the tanh activation function. Similarly, the pooling layer feature map in DFEB is evaluated as follows [24]:

$$\tilde{F}_{\bar{l}_1, \bar{l}_2}^t = \text{max-pooling}(F_{l_1, l_2}^t) \quad (9)$$

where \tilde{F}^t stands for the t th feature map obtained in the pooling-layer. Similarly, for the GAP layer, the mathematical expression to evaluate the feature vector ($\tilde{\mathbf{g}} = \tilde{g}(1, t)$ with $t = 1, 2, 3, \dots, T$ and T is the total number of feature maps) from the tensor representation ($\mathbf{G} = G(l_1, l_2, t)$) is given as follows [29]:

$$\tilde{g}(1, t) = \frac{1}{pq} \sum_{l_1=1}^p \sum_{l_2=1}^q G(l_1, l_2, t). \quad (10)$$

Similarly, we have used the GRU layer with 16 hidden neurons for each scale DFEB. The GRU layer comprises GRU

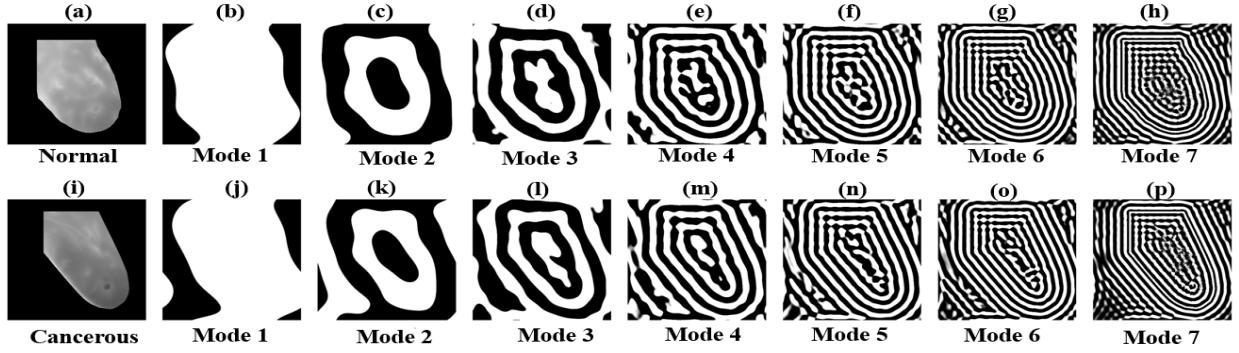


Fig. 3. (a) Thermogram image for normal class. (b)–(h) Seven modes of thermogram image obtained using 2DEWT with FBPs for normal class. (i) Thermogram image for cancerous class. (j)–(p) Seven modes of thermogram image obtained using 2DEWT with FBPs for cancerous class.

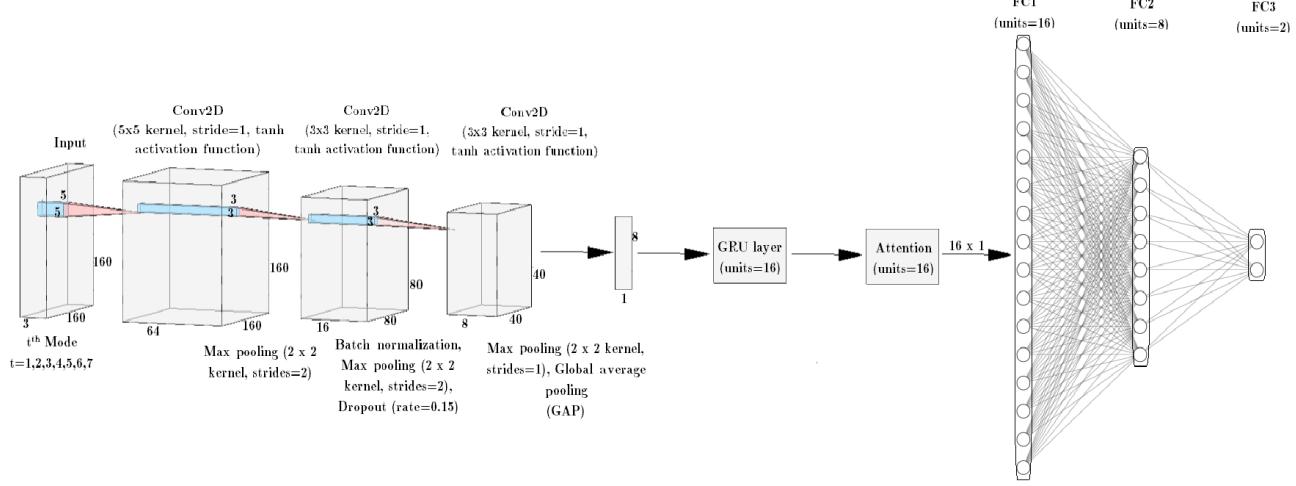


Fig. 4. Architecture or the layers in DFEB for each mode case in the proposed MSADIDL model (DFEB is placed in the fifth row of the MSADIDL model in Fig. 1).

cells that contain reset (R) and update gates (U). The GRU cell processes the present and past time step information to evaluate the memory value, $mv(k)$, for k th time-step. The equations to evaluate the responses for reset gate, forget gate, and memory value for a single GRU cell are given as follows [30], [31]:

$$R(k) = h(W_R[mv(k-1), f(k)] + b_R) \quad (11)$$

$$V(k) = h(W_V[mv(k-1), f(k)] + b_V) \quad (12)$$

$$\tilde{m}(k) = \tanh(W_m[R(k) \otimes mv(k-1), f(k)] + b_m) \quad (13)$$

$$mv(k) = (1 - V(k)) \otimes mv(k-1) + V(k) \otimes \tilde{m}(k) \quad (14)$$

where (W_R, W_V) , (b_R, b_V) are weight matrices and bias vectors for reset and update gates. In this work, we have evaluated the vector of size as 8×1 after the GAP layer, as shown in Fig. 4. The vector of size 8×1 is used as the input to the GRU layer. The number of time steps in the GRU layer is considered as 8. The output of the GRU layer has the size as 8×16 , where 16 is the number of GRU units. After the GRU layer for each scale of the thermogram image, we have used the attention layer. The attention layer helps assign higher weights to the relevant features while lesser weight values to other features [32]. In this work, intra-scale attention helps select discriminative features from the modes of thermogram images for breast cancer detection. The softmax activation function is used in the output layer of the proposed

MSADIDL model. The training parameters include the learning rate of 0.001, Adam optimizer, and the number of epochs as 100 for the MSADIDL model. The early stopping criteria with a patience of 20 epochs are used during the training of the proposed model [33]. The hold-out validation and fivefold cross-validation (CV) techniques are employed for the selection of the training and test instances of the MSADIDL model [34]. For hold-out validation, 68%, 12%, and 20% of the thermogram images are employed for the training, validation, and testing of the proposed MSADIDL model. The metrics such as accuracy, precision, recall, F1-score, and Kappa score [35] are utilized for the evaluation of the performance of the MSADIDL model in both hold-out validation and fivefold CV for the detection of breast cancer using thermogram images. Transfer learning techniques have been widely used to detect breast cancer using images from different modalities such as mammography, ultrasound, and thermogram, respectively [36], [37], [38]. The advantages of these models are that they have fewer trainable parameters and can be used as feature extractors. Aidossov et al. [38] have compared four transfer learning models, such as Xception, MobileNet, ResNet50, and VGG16, to detect breast cancer from thermogram images. They obtained an accuracy value of 93.8% using the combined thermogram images from two different databases. In this work, the classification performance of the proposed MSADIDL model is compared with various

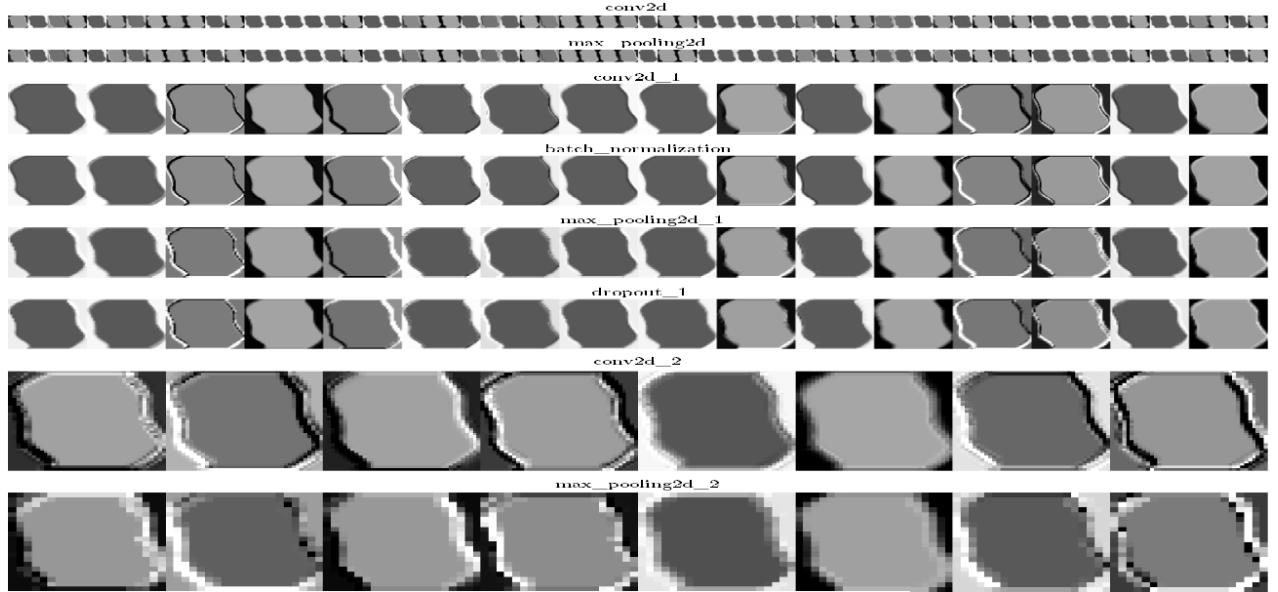


Fig. 5. Feature-map plots evaluated in the different layers of the DFEB of mode 1 of the thermogram image.

transfer learning models [39], [40] using the thermogram images from the same database to detect breast cancer. For each transfer learning model, we have used all layers before the fully connected layer as feature extractors or frozen blocks. Then, the final layer of the frozen block is connected to the new output layer with two output neurons corresponding to normal and cancerous classes. Fine-tuning is only performed between the weight parameters between the final layer of the frozen block and the new output layer to detect breast cancer using thermogram images. The training parameters, such as the learning rate, optimizer, and the number of epochs of each transfer learning model, are the same as those of the MSADIDL model. The transfer learning model requires the input as a third-order tensor [41]. In this work, we have considered the third-order tensor for both MSADIDL and transfer learning-based methods. The third-order tensor of size $160 \times 160 \times 3$ is formulated by appending the same mode image three times in the input layer of Fig. 4.

IV. RESULTS AND DISCUSSION

In this work, we have shown the results obtained using the proposed MSADIDL approach and discussed the results for detecting breast cancer using thermogram images.

A. Interpretability of the Proposed MSADIDL Model

The DL models that consider the input as the direct thermogram image are called the black-box model [22]. These models are complex in understanding the reason behind the output for the given test input image. In this work, we have designed the MSADIDL model in such a way that it can be interpretable to detect breast cancer using thermogram images. First, the 2DEWT with FBPs obtains interpretable components or modes from the thermogram image. The frequency ranges for mode 1, mode 2, mode 3, mode 4, mode 5, mode 6, and mode 7 in the mean spectral radius vector of the thermogram

image are [0–4 bin], [4–8 bin], [8–16 bin], [16–24 bin], [24–32 bin], [32–40 bin], and [40–160 bin], respectively. The DFEB block is applied to each mode of the thermogram image to evaluate a 2-D feature vector. The feature map plots obtained in the different layers of the mode 1 DFEB is depicted in Fig. 5. It is observed that the feature maps obtained in the layers of mode 1 DFEB are associated in the frequency range of [0–4 bin] in the thermogram image. The feature maps produced in the layers of the DFEB block for all seven modes are different as these modes are associated with different frequency ranges in the mean spectral radius vector of the thermogram image. The 2DEWT helps to provide interpretability in terms of the frequency range for the feature maps produced in the layers of the DFEB block for each mode of thermogram image. The 2-D feature vectors of all seven modes are merged in the concatenation layer, and the dimension of this feature vector is 1×14 . After the concatenation layer, the dense layers obtain the reduced-size feature vector, which captures the reduced all-scale information of thermogram images. The probability values or scores evaluated in the output layer (using softmax-based activation function) of the MSADIDL model for normal and cancerous classes are shown in Fig. 6(a). It is observed that out of 1216 training instances, the softmax layer's score 1 and score 2 values of 1210 instances fall in the range of [0–0.5] and [0.5–1] for normal and cancerous classes after the training of the proposed MSADIDL model. Therefore, the training accuracy of the model is evaluated as 99.51%. The plot for the accuracy versus epoch obtained using the proposed MSADIDL model with all seven modes is shown in Fig. 6(b) for hold-out validation. We have selected the early stopping with patience for the proposed model as 20 epochs. Hence, after 100 epochs, the training has been terminated as there is no further improvement in the validation accuracy value of the MSADIDL model. Similarly, we have shown the confusion matrix obtained using the MSADIDL model for test thermogram images in Fig. 6(c). The proposed method has resulted in only one false positive value, and the testing

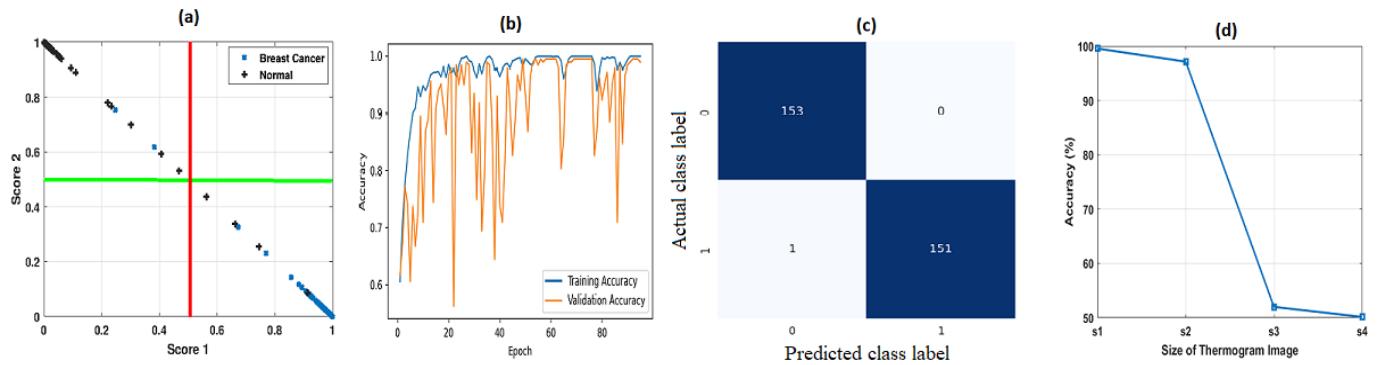


Fig. 6. (a) Plot for probability or score evaluated at the softmax layer of the proposed MSADIDL model for normal and cancerous classes (red and green color lines are interpreted as the thresholds for score 1 and score 2). (b) Accuracy versus epoch plot of the proposed MSADIDL model with all seven modes of thermogram images for hold-out validation. (c) Confusion matrix plot for the proposed MSADIDL model with all seven modes of thermogram images. (d) Variation of accuracy with different sizes of thermogram images (s1 as 160×160 , s2 as 240×180 , s3 as 320×240 , and s4 as 640×480).

accuracy is obtained as 99.67% for one single run of hold-out validation.

The variations of the accuracy of the MSADIDL model with different sizes of thermogram images are shown in Fig. 6(d). It is observed that the highest accuracy is obtained for the thermogram image with size as 160×160 using the MSADIDL model. When the thermogram image of size as 640×480 is used, the proposed model has produced an accuracy value of 50.16%, which is less than the MSADIDL model with thermogram images of size as 160×160 . The difference in the mean values of both scores (score 1 and score 2) between normal and cancerous classes is high using the thermogram images of sizes 160×160 as compared to the other sizes of the thermogram images. For 320×240 and 640×480 sizes-based thermogram images, the feature maps produced in the different layers of the proposed MSADIDL model for normal and cancerous classes are correlated. Due to this reason, the differences in the mean values of both scores (softmax layer probability values) between normal and cancerous classes are less for the thermogram images with sizes as 320×240 and 640×480 using the proposed approach to detect breast cancer. The computational complexity is enhanced with the increase in the size of the thermogram images using the proposed model to detect breast cancer. Therefore, the optimal size of thermogram images is obtained as 160×160 for the automated detection of breast cancer using thermogram images. We have evaluated the classification performance of the proposed MSADIDL model using the single output neuron and binary cross-entropy as the cost function. The classification accuracy of the MSADIDL model with one output neuron with binary cross-entropy and the same hyper-parameters is obtained as 50.16%, which is less than the MSADIDL model with two output neurons with categorical cross-entropy cases. In the two-output neuron case, the network simultaneously maximizes one output and minimizes other output for better generalization of the model. Also, the two-output neurons provide additional gradient signals in the weight updation stage to constrain the model better and avoid overfitting. Due to this reason, we have used two output neurons with categorical cross-entropy in the proposed MSADIDL model to detect breast cancer using thermogram images.

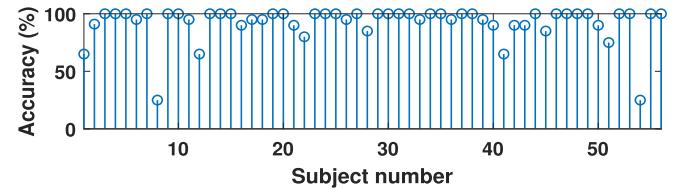


Fig. 7. Accuracy versus the number of subjects for MSADIDL model using LOOCV method.

We have used grid search to evaluate the number of units in the GRU block for the proposed DNN. The GRU units grid is initialized as [4, 8, 12, 16, 20, 32]. The accuracy value of the proposed MSADIDL model is evaluated using each GRU unit in the grid. The optimal number of GRU units is selected based on the maximum accuracy of the MSADIDL model to detect breast cancer using thermogram images. It is observed that the accuracy values are obtained as 98.95%, 97.77%, 99.01%, 99.54%, 99.01%, and 96.19% using the MSADIDL models the number of GRU units as 4, 8, 12, 16, and 20, respectively. The highest accuracy of the MSADIDL model is found for the GRU unit at 16. Due to this reason, we have selected the number of GRU units in the proposed MSADIDL model as 16 for detecting breast cancer using thermogram images. The classification accuracy values of the MSADIDL model evaluated using the leave-one-out CV (LOOCV) scheme (56-fold CV) for all 56 subjects are shown in Fig. 7. The 20 thermogram images for each subject are used in the MSADIDL model testing phase for each LOOCV fold. It is observed that the proposed model has produced 100% accuracy for 32 subjects. The average accuracy of the MSADIDL model for all 56 subjects is 92.16% using LOOCV. The MSADIDL model has obtained an average accuracy of below 25% for subject 8 and subject 54 using LOOCV. Similarly, for subject 1, subject 12, and subject 41, the MSADIDL model has obtained an average accuracy of less than 70%. We have evaluated the correlations between the 20 thermogram images recorded in a 15-s interval for subject 8 and subject 54 cases. It is observed that the average correlation coefficient (ACC) below 0.3 is observed between the thermogram images for these two subjects (subject 8 and subject 54). This may happen due to the variations in the body temperature during

TABLE I
CLASSIFICATION RESULTS OF MSADIDL MODEL WITH DIFFERENT MODE SELECTION CASES FOR HOLD-OUT VALIDATION

Modes Selection	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Kappa (%)	Concatenated layer (dimension)
M1 and M2	96.78 ± 1.88	96.94 ± 1.80	96.79 ± 1.88	96.78 ± 1.88	93.57 ± 3.76	1 × 4
M1, M2 and M3	98.49 ± 2.38	98.49 ± 2.38	98.49 ± 2.38	98.49 ± 2.38	96.98 ± 4.77	1 × 6
M1, M2, M3 and M4	99.21 ± 0.57	99.21 ± 0.56	99.21 ± 0.57	99.21 ± 0.57	98.42 ± 1.14	1 × 8
M1, M2, M3, M4 and M5	97.83 ± 1.18	97.86 ± 1.16	97.83 ± 1.18	97.83 ± 1.18	95.67 ± 2.36	1 × 10
M1, M2, M3, M4, M5 and M6	99.01 ± 0.68	99.01 ± 0.68	99.01 ± 0.68	99.01 ± 0.687	98.03 ± 1.37	1 × 12
M2 and M3	98.55 ± 1.32	98.60 ± 1.27	98.55 ± 1.32	98.55 ± 1.32	97.11 ± 2.64	1 × 2
M2, M3 and M4	96.32 ± 3.53	96.38 ± 3.52	96.32 ± 3.54	96.32 ± 3.54	92.65 ± 7.08	1 × 6
M2, M3, M4 and M5	98.16 ± 1.41	98.18 ± 1.38	98.16 ± 1.41	98.16 ± 1.41	96.32 ± 2.83	1 × 8
M2, M3, M4, M5 and M6	95.67 ± 3.58	95.69 ± 3.59	95.67 ± 3.58	95.67 ± 3.58	91.34 ± 7.16	1 × 10
M2, M3, M4, M5, M6 and M7	97.04 ± 0.85	97.10 ± 0.86	97.05 ± 0.85	97.04 ± 0.85	94.09 ± 1.70	1 × 12
M3 and M4	97.90 ± 0.67	97.92 ± 0.65	97.00 ± 0.67	97.00 ± 0.67	96.06 ± 1.71	1 × 4
M3, M4 and M5	95.93 ± 1.99	95.97 ± 1.96	95.93 ± 1.99	95.93 ± 1.99	91.86 ± 3.99	1 × 6
M3, M4, M5 and M6	96.98 ± 2.77	97.03 ± 2.67	96.98 ± 2.77	96.98 ± 2.78	93.96 ± 5.55	1 × 8
M3, M4, M5, M6 and M7	95.34 ± 1.67	95.46 ± 1.64	95.34 ± 1.67	95.34 ± 1.67	90.68 ± 3.35	1 × 10
M4 and M5	83.08 ± 12.33	83.30 ± 12.12	83.09 ± 12.32	83.00 ± 12.43	66.17 ± 24.66	1 × 4
M4, M5 and M6	90.62 ± 7.44	91.01 ± 6.84	90.62 ± 7.42	90.55 ± 7.57	81.25 ± 14.87	1 × 6
M4, M5, M6 and M7	84.52 ± 17.22	79.51 ± 27.24	84.49 ± 17.28	81.17 ± 23.91	68.98 ± 34.57	1 × 8
M5 and M6	73.50 ± 19.12	63.59 ± 31.48	73.44 ± 19.20	66.79 ± 27.30	46.88 ± 38.41	1 × 4
M5, M6 and M7	91.40 ± 1.90	91.77 ± 1.77	91.41 ± 1.90	91.38 ± 1.91	82.82 ± 3.81	1 × 6
M6 and M7	89.70 ± 2.91	90.27 ± 2.34	89.70 ± 2.91	89.65 ± 2.96	79.40 ± 5.80	1 × 4
All seven modes	99.54 ± 0.39	99.54 ± 0.39	99.54 ± 0.39	99.54 ± 0.39	99.08 ± 0.78	1 × 14

the recording of thermogram images using the FLIR thermal camera. Due to this reason, the MSADIDL model has obtained less accuracy for these subjects as compared to other subjects using LOOCV.

B. Mode Selection Results

The results evaluated using the MSADIDL model for different combinations of modes of thermogram images with hold-out validation are shown in Table I. The average and standard deviation values are shown for five-run random hold-out validation cases to detect breast cancer using the proposed model. For selected modes, the dimension of the concatenation layer in the proposed MSADIDL model is different. We have shown the dimensions of the concatenation layer for different mode selection cases in Table I. It is noted that when Mode 1 (M1) and Mode 2 (M2) of thermogram images are utilized, the MSADIDL has produced an accuracy value of 96.78%. The precision, recall, and F1-score values are more than 96%. Similarly, when the first three modes (M1, M2, and M3) of the thermogram images are considered, the classification accuracy of the MSADIDL is improved to 98.41%. It is evident from Table I that the accuracy of the MSADIDL model when using the first four modes of the thermogram images is 99.21%. The first four modes capture the major spectral energy in the mean spectral vector of the thermogram image evaluated using PPFFT. Due to this reason, the MSADIDL model has obtained higher accuracy using the first four modes as compared to the first two modes and the first three modes of the thermogram images using hold-out validation. Furthermore, when the first five modes of the thermogram images coupled with MSADIDL and the first six modes coupled with the MSADIDL model are utilized, the accuracy value is reduced compared to that of the first four modes coupled with the MSADIDL model. The mode 5, mode 6, and mode 7 capture the high-frequency components of the thermogram image. These components

capture edge information and also noise information. Due to this reason, the accuracy values of the MSADIDL model are decreased when the first five and first six modes of thermogram images are used. We have also evaluated the ACC values between mode 4 and mode 5, ACC between mode 5 and mode 6, and also between mode 6 and mode 7. The ACC values are high for these three mode combination cases; hence, these modes are statistically dependent. Due to this correlated information, there is a decrease in the accuracy of the MSADIDL model with input as the first five modes of the thermogram images. Similarly, for other mode combinations, the accuracy and other metrics are less than the first four mode combinations using the MSADIDL model for detecting breast cancer using thermogram images. When all the seven modes of the thermogram images are considered, the MSADIDL has produced accuracy, precision, recall, F1-score, and specificity values greater than 99.50% compared to all other mode combinations of the thermogram images.

The classification performance of the MSADIDL model computed using fivefold CV is shown in Table II for breast cancer detection using thermogram images. It is noted that the accuracy value of the MSADIDL model is obtained as 99.34% using the first six modes of the thermogram images. When the first five modes of the thermogram images are used, the accuracy value of the MSADIDL model is found as 99.40%, which is higher than the first two, first three, first four, and first six modes of the thermal images with the MSADIDL model using fivefold CV for breast cancer detection. The CV method selects the training and test modes of the thermogram images from different positions of the dataset at each fold to detect breast cancer using thermogram images. The overall classification performance is evaluated using the average of all fold results evaluated using the MSADIDL model. Due to the correlated information associated with the higher frequency component-based modes, there are variations in the classification accuracy values of the MSADIDL model at each

TABLE II
CLASSIFICATION RESULTS OF MSADIDL MODEL WITH DIFFERENT MODE SELECTION CASES FOR FIVEFOLD CV

Mode Selection	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	Kappa (%)
M1 and M2	97.90 \pm 1.92	97.92 \pm 1.93	97.90 \pm 1.93	97.90 \pm 1.92	95.80 \pm 3.85
M1, M2 and M3	99.14 \pm 0.60	99.15 \pm 0.60	99.14 \pm 0.60	99.14 \pm 0.60	98.29 \pm 1.21
M1, M2, M3 and M4	98.88 \pm 1.21	98.89 \pm 1.21	98.88 \pm 1.21	98.88 \pm 1.21	97.77 \pm 2.43
M1, M2, M3, M4 and M5	99.40 \pm 0.56	99.41 \pm 0.56	99.40 \pm 0.56	99.40 \pm 0.56	98.81 \pm 1.12
M1, M2, M3, M4, M5 and M6	99.34 \pm 0.29	99.34 \pm 0.29	99.34 \pm 0.29	99.34 \pm 0.29	98.68 \pm 0.58
M2 and M3	88.91 \pm 19.55	83.96 \pm 29.53	88.94 \pm 19.48	85.60 \pm 26.18	77.90 \pm 38.97
M2, M3 and M4	99.54 \pm 0.60	99.54 \pm 0.59	99.54 \pm 0.60	99.54 \pm 0.60	99.08 \pm 1.21
M2, M3, M4 and M5	98.22 \pm 1.80	98.28 \pm 1.73	98.22 \pm 1.80	98.22 \pm 1.80	96.45 \pm 3.60
M2, M3, M4, M5 and M6	98.55 \pm 1.57	98.58 \pm 1.55	98.55 \pm 1.57	98.55 \pm 1.57	97.11 \pm 3.14
M2, M3, M4, M5, M6 and M7	96.72 \pm 3.85	96.72 \pm 3.84	96.72 \pm 3.85	96.72 \pm 3.85	93.44 \pm 7.70
M3 and M4	98.49 \pm 0.94	98.50 \pm 0.93	98.49 \pm 0.94	98.49 \pm 0.94	96.98 \pm 1.88
M3, M4 and M5	98.49 \pm 1.33	98.51 \pm 1.32	98.49 \pm 1.33	98.49 \pm 1.33	96.98 \pm 2.67
M3, M4, M5 and M6	97.83 \pm 2.26	97.89 \pm 2.17	97.83 \pm 2.26	97.83 \pm 2.27	95.67 \pm 4.53
M3, M4, M5, M6 and M7	95.86 \pm 3.27	95.97 \pm 3.17	95.86 \pm 3.27	95.86 \pm 3.28	91.73 \pm 6.55
M4 and M5	96.00 \pm 2.65	96.09 \pm 2.59	95.99 \pm 2.65	95.99 \pm 2.65	91.99 \pm 5.30
M4, M5 and M6	95.34 \pm 1.81	95.56 \pm 1.55	95.34 \pm 1.80	95.33 \pm 1.82	90.68 \pm 3.62
M4, M5, M6 and M7	95.93 \pm 2.92	96.00 \pm 2.86	95.93 \pm 2.92	95.93 \pm 2.92	91.86 \pm 5.84
M5 and M6	92.45 \pm 3.98	93.09 \pm 3.12	92.45 \pm 3.99	92.40 \pm 4.07	84.91 \pm 7.97
M5, M6 and M7	92.45 \pm 2.63	92.81 \pm 2.37	92.45 \pm 2.62	92.43 \pm 2.64	84.91 \pm 5.25
M6 and M7	90.81 \pm 2.47	90.99 \pm 2.49	90.81 \pm 2.47	90.80 \pm 2.47	81.63 \pm 4.94
All seven modes	99.54 \pm 0.49	99.54 \pm 0.48	99.54 \pm 0.49	99.54 \pm 0.49	99.08 \pm 0.98

fold. Hence, the classification performance of the MSADIDL is declined by incorporating mode 6 information in the fifth row of Table II. Furthermore, when Mode 2, Mode 3, and Mode 4 of the thermogram images are used, the MSADIDL has produced an accuracy value of 99.54% which is the same as that of all seven modes of the thermogram images, coupled with the MSADIDL classifier with fivefold CV for breast cancer detection. For other mode selection cases, the MSADIDL model has a lesser classification performance in all metrics for breast cancer detection.

We have evaluated the performance of the proposed approach using other multiscale decomposition techniques, such as 2D discrete wavelet transform (2DDWT) [42], singular value decomposition (SVD) [41], and 2D stationary wavelet transform (2DSWT) [43] with the MSADIDL model. These results are depicted in Table III. For 2DDWT and 2DSWT-based methods, only one level decomposition of the thermogram image is performed. Hence, four modes are evaluated from each thermogram image using 2DDWT and 2DSWT-based methods. Similarly, in SVD-based method, four modes are extracted based on the grouping of eigentriples of each thermogram image. It is observed that the MSADIDL model has obtained higher overall accuracy using 2DEWT with FBPs than other decomposition methods. The 2DEWT with FBPs efficiently extracts the local components from the thermogram image. The MSADIDL model has produced the highest overall accuracy in detecting breast cancer using these components or modes than other decomposition methods mentioned in Table III. The classification results of the proposed MSADIDL model are also evaluated using breast ultrasound images for the automated classification of benign and malignant tumors. These results are shown in Table IV. It is observed that the proposed approach has obtained an average accuracy value of 84.96% to classify benign and malignant classes using breast ultrasound images. The DL models have shown less classification performance using ultrasound images to detect breast cancer [44], [45]. The proposed

TABLE III
CLASSIFICATION PERFORMANCE OF MSADIDL NETWORK USING DIFFERENT MULTISCALE DECOMPOSITION METHODS

Decomposition Methods	Accuracy (%)
2DDWT (4 sub-bands)	98.85 \pm 0.63
SVD (4 modes)	71.27 \pm 0.39
2DSWT (4 modes)	76.91 \pm 2.05
Proposed work (4 modes)	99.21 \pm 0.57
Proposed work (7 modes)	99.54 \pm 0.39

TABLE IV
CLASSIFICATION PERFORMANCE OF MSADIDL MODEL USING ULTRASOUND IMAGES TO CLASSIFY BENIGN AND MALIGNANT CLASSES

Measures	Value ($\mu \pm \sigma$)
Accuracy (%)	84.96 \pm 1.84
Precision (%)	85.14 \pm 1.73
Recall (%)	84.98 \pm 1.83
F1-score (%)	84.95 \pm 1.85
Kappa (%)	69.90 \pm 3.60

MSADIDL approach decomposes the thermogram image into modes and uses the mode information in the DNN layers to detect breast cancer. The mode characteristics are different for normal and cancerous classes. Due to this reason, the MSADIDL has shown higher accuracy in the detection of breast cancer. However, for ultrasound image cases, there is not much difference in the characteristics of each mode for benign and malignant classes. Therefore, the MSADIDL model with modes of ultrasound images has obtained less classification accuracy than that of the MSADIDL model with modes of thermogram images.

We have evaluated the classification accuracy values of the proposed MSADIDL model by varying the number of layers and activation functions. These new results are shown in Table V. It is observed that the MSADIDL model with all layers has obtained a higher accuracy value as compared to other types of multiscale domain DL model architectures. Similarly, the MSADIDL model with convolution layer followed by tanh function has demonstrated higher accuracy than

TABLE V
ACCURACY VALUES OF MSADIDL MODEL WITH VARIATIONS
IN ACTIVATION FUNCTION AND OTHER LAYERS

Figure 4 Architecture variation	Accuracy(%)
Sigmoid activation function after convolution layers	98.68
ReLU activation function after convolution layers	97.70
Architecture without BN, dropout, GAP, GRU, and attention layers	98.03
Architecture without BN, and dropout layers	97.05
Proposed MSADIDL model architecture	99.54

TABLE VI
CLASSIFICATION PERFORMANCE OF MSADIDL
MODELS FOR IMBALANCED DATASET CASES

Database	Accuracy (%)	Precision (%)	Recall (%)
Database1 (Thermogram images)	98.10 ± 1.97	97.99 ± 2.07	98.16 ± 1.91
Database2 (Ultrasound images)	79.88 ± 4.49	79.88 ± 4.43	79.91 ± 4.15

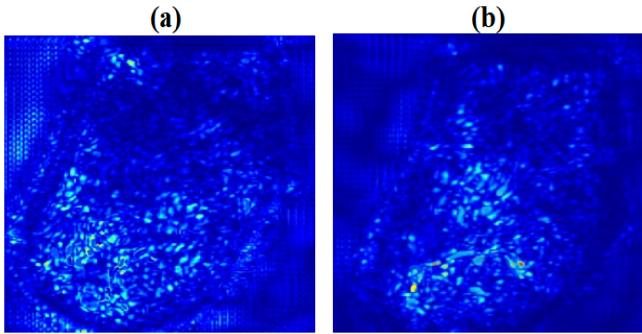


Fig. 8. (a) Saliency map for cancer class. (b) Saliency map for normal class.

the sigmoid and rectified linear unit (ReLU) based models to detect breast cancer using thermogram images. In Table VI, we have shown the performance of the MSADIDL approach for imbalanced dataset cases with hold-out validation to detect breast cancer. For database 1, the 760 and 560 thermogram images from the normal and cancerous classes are considered. Similarly, we have considered 1000 and 800 ultrasound images from benign and malignant classes for the database2 to evaluate the proposed MSADIDL approach. It is observed that the average accuracy of the MSADIDL approach is reduced to 98.17% for an imbalanced dataset case using thermogram images. The average accuracy value has reduced only by 1.37% when the imbalanced dataset is considered. This shows the robustness of the proposed MSADIDL model in detecting breast cancer using thermogram images. However, for ultrasound images in an imbalanced dataset case, the average accuracy of the MSADIDL model has been reduced to 79.88%. We have shown the saliency maps evaluated using the proposed approach in Fig. 8(a) and (b) for cancerous and normal classes, respectively. For the cancerous class, the intensity level in the saliency map is more spread compared to the normal class. We have evaluated the feature vector in the attention layer of each mode for the thermogram image used in the testing phase of the model. Significant differences are observed in the values of the attention layer neurons for

mode 1, mode 2, and mode 3 of thermogram images between the normal and cancerous classes. The proposed MSADIDL model gives more attention to the low-frequency information of the thermogram image to detect breast cancer.

C. Comparison With Transfer Learning-Based Methods

We have shown the classification results of different transfer learning models in Tables VII and VIII using thermogram images with hold-out validation and fivefold CV techniques to detect breast cancer. All transfer learning models have demonstrated average accuracy values below 80% for automatically detecting breast cancer using thermogram images. The proposed MSADIDL network has produced an average accuracy value of 99.54%, higher than all transfer learning-based models for detecting breast cancer using thermogram images in both validation cases. The proposed model captures the local components of the thermogram images in the 2DEWT phase and extracts the learnable features in the DFME block stage. Hence, the suggested model has extracted discriminative features from the thermogram images to obtain the highest accuracy for breast cancer detection. On the other hand, the transfer learning models cannot extract discriminative learnable features from thermogram images, thereby obtaining less classification performance than the proposed MSADIDL method for breast cancer detection. The trainable and non-trainable parameters in the proposed MSADIDL model are 122 038 and 224, respectively. The trainable parameters of each transfer learning model are shown in Table VII. The number of trainable parameters in the proposed MSADIDL model is higher than in the transfer learning models. However, the suggested MSADIDL model has demonstrated superior classification performance with more than 99% accuracy in detecting breast cancer. The total number of parameters in the proposed model is 122 262, which is lower than all transfer learning models. Due to fewer total parameters, the proposed MSADIDL model has the advantage of less inference time to detect breast cancer using thermogram images. The suggested MSADIDL model is suitable for edge computing-based applications due to the lower inference time to process thermogram images to detect breast cancer.

D. Comparison With Existing Methods

In Table IX, we have compared the classification accuracy of the suggested MSADIDL approach with various existing techniques to detect breast cancer disease using thermogram images. The existing methods in [13] and [15] have used GLCM-based frequency from the thermogram images and used ML-based techniques to detect breast cancer. The accuracy values of these two methods are 92.50% and 88.10%, which are less than the accuracy of the proposed approach. Similarly, the 2D-DWT and curvelet-based methods in [14], [19], and [16] have achieved accuracy values less than the proposed 2D-EWT with the FBPs-based DL method. The hybrid features of thermogram images with different ML-based methods in [17] and [18] have produced accuracy values of 95.80% and 97.18%, respectively. The proposed MSADIDL method has produced higher classification performance than

TABLE VII
COMPARISON OF THE RESULTS OF THE MSADIDL MODEL WITH DIFFERENT TRANSFER LEARNING MODELS FOR BREAST CANCER DETECTION USING HOLD-OUT VALIDATION

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Kappa (%)	Trainable parameters
ResNet101 [39]	78.16 \pm 0.39	81.49 \pm 0.59	78.11 \pm 0.39	77.54 \pm 0.40	56.28 \pm 0.78	8,194
VGG19 [39]	75.54 \pm 0.16	78.78 \pm 0.53	75.48 \pm 0.15	74.80 \pm 0.16	51.02 \pm 0.32	2,050
EfficientNetV2B3 [39]	73.63 \pm 1.10	74.79 \pm 1.42	73.60 \pm 1.10	73.31 \pm 1.05	47.24 \pm 2.21	6,146
DenseNet121 [39]	72.59 \pm 0.49	74.30 \pm 0.54	72.54 \pm 0.49	72.07 \pm 0.52	45.13 \pm 0.98	4,098
MobileNetV2 [39]	76.65 \pm 0.48	80.07 \pm 0.51	76.60 \pm 0.48	75.94 \pm 0.55	53.25 \pm 0.96	5,122
Xception [39]	71.93 \pm 0.70	75.51 \pm 1.14	71.87 \pm 0.70	70.88 \pm 0.85	43.80 \pm 1.41	8,194
InceptionV3 [39]	70.88 \pm 1.08	72.61 \pm 1.58	70.84 \pm 1.08	70.30 \pm 0.98	41.71 \pm 2.16	8,194
Proposed MSADIDL	99.54 \pm 0.39	99.54 \pm 0.39	99.54 \pm 0.39	99.54 \pm 0.39	99.08 \pm 0.78	12,2038

TABLE VIII
COMPARISON OF THE RESULTS OF MSADIDL MODEL WITH ALL TRANSFER LEARNING MODELS FOR BREAST CANCER DETECTION USING FIVEFOLD CV

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Kappa (%)
ResNet101[39]	77.18 \pm 0.57	80.87 \pm 0.36	77.12 \pm 0.57	76.44 \pm 0.71	54.30 \pm 1.14
VGG19[39]	75.54 \pm 0.60	79.51 \pm 0.98	75.48 \pm 0.60	74.66 \pm 0.58	51.02 \pm 1.21
EfficientNetV2B3[39]	76.52 \pm 0.73	80.15 \pm 0.83	76.46 \pm 0.73	76.46 \pm 0.73	52.99 \pm 1.47
DenseNet121[39]	72.59 \pm 0.16	75.47 \pm 0.77	72.53 \pm 0.15	71.76 \pm 0.23	45.11 \pm 0.32
MobileNetV2[39]	77.37 \pm 0.35	80.66 \pm 0.37	77.32 \pm 0.35	76.73 \pm 0.39	54.70 \pm 0.71
Xception[39]	72.52 \pm 0.70	77.28 \pm 0.96	72.45 \pm 0.70	71.23 \pm 0.82	44.97 \pm 1.40
InceptionV3[39]	73.11 \pm 1.20	76.66 \pm 1.92	73.05 \pm 1.20	72.16 \pm 1.22	46.16 \pm 2.41
Proposed MSADIDL	99.54 \pm 0.49	99.54 \pm 0.48	99.54 \pm 0.49	99.54 \pm 0.49	99.08 \pm 0.98

TABLE IX
COMPARISON WITH EXISTING METHODS FOR BREAST CANCER DETECTION USING THERMOGRAM IMAGES

Authors	Method employed	Accuracy (%)
Milosevic et al. [13]	GLCM features extracted from thermogram images and KNN classifier	92.50
Pramanik et al. [14]	2D DWT domain statistical features from thermogram images and MLP classifier	90.48
Sathish et al. [19]	2D DWT domain statistical features followed by RSFS-based feature selection and SVM classifier	91.00
Dey et al. [21]	Transfer learning model applied on original, and edge defected thermogram images followed by softmax layer	98.80
Gomez et al. [22]	Deep CNN with tree parent estimator based hyper-parameter tuning	92.00
Nasser et al. [17]	Different texture analysis based extraction of features from thermogram images and MLP classifier	95.80
Francis et al. [16]	Curvelet transform domain statistical and texture features from thermogram images and SVM classifier	90.91
Resmini et al. [18]	GLCM, local ternary pattern, DWT, Fractal and Hurst coefficient based features evaluated from thermogram images followed by feature selection and SVM classifier	97.18
Acharya et al. [15]	GLCM based features extracted from thermogram images and SVM classifier	88.10
Proposed work	MSADIDL model	99.54

hybrid feature-based techniques to detect breast cancer using thermogram images. The deep CNN directly applied in the thermogram images has achieved the accuracy values of 92%. Similarly, the deep transfer learning used over the thermogram images and its edge detected images have obtained an accuracy value of 98.80% [21], which is less than the accuracy value of the proposed MSADIDL model to detect breast cancer. In this work, we have also considered the direct thermogram image as input to the fine-tuned VGG16 model to detect breast cancer. The fine-tuned VGG16 model has obtained less accuracy than the proposed MSADIDL model to detect breast cancer. In addition, all transfer learning models have produced accuracy values of less than 80% for the automated detection of breast cancer. The multiscale decomposition using 2DEWT with FBPs helps to obtain the local information or modes of thermogram images. Each mode captures the thermogram image's segregated information in a particular frequency range. Due to this advantage, the MSADIDL model has obtained higher accuracy in detecting breast cancer using thermogram images.

E. Computational Complexity of MSADIDL Model

For one mode of the thermogram image, the DFEB consists of three convolution layers, three max-pooling layers, three dense layers, one BN layer, one dropout layer, one GAP, one GRU, and one attention layer, respectively. For l th convolution layer, the computational complexity is computed as $O(m_1^l m_2^l k_1^l k_2^l I^l F^l)$ [24]. The parameters, m_1^l and m_2^l , are called the size of the input image or feature map to the convolution layer. Similarly, k_1^l and k_2^l are interpreted as the size of the kernel used in the l th convolution layer. Moreover, the parameters I^l and F^l are denoted as the number of input and output subimages or feature maps of l th convolution layer for one mode. For the pooling layer, the computational complexity is calculated as $O(P_1^l P_2^l)$ [24], where P_1^l and P_2^l are the size of the input image or feature map to the l th pooling layer. For the l th dense layer, the computational complexity is calculated as $O(n_1^l n_2^l \log(g^l))$, where n_1^l and n_2^l are the size of the weight matrix in the l th dense layer. The parameter g^l is the number of neurons in the l th dense

layer. For the GRU layer, the computational complexity is evaluated using the following mathematical expression as $O(N_s N_h^2 + N_s N_h N_i)$ [46]. where N_i and N_h are the input and hidden state dimensions, respectively. The parameter, N_s , is interpreted as the length of the sequence or feature vector in the input to the GRU layer. The computational complexity of the BN layer is given as $O(B)$. where B is the size of the input to the BN layer. Similarly, for the dropout layer, the computational complexity is given by $O(N_d)$. where N_d is the input size for the dropout layer. The complexity of the attention layer is given as $O(a_1^2 a_2 + a_2^2 a_1)$. where $a_1 \times a_2$ is the size of the input to the attention layer. The total computational complexity of one mode of DFEB (CC_{DFEB}) can be given by $CC_{DFEB} = O(\sum_{l=1}^3 m_1^l m_2^l k_1^l k_2^l I^l F^l) + O(\sum_{l=1}^4 P_1^l P_2^l) + O(n_1^l n_2^l \log(g^l)) + O(B) + O(N_d) + O(a_1^2 a_2 + a_2^2 a_1)$. Thus, the overall computational complexity of the MSADIDL model is evaluated as the sum of computational complexity values of all seven mode DFEBs, followed by using all seven modes. The overall computational complexity of the MSADIDL model is $\sum_{i=1}^7 CC_{DFEB}^i + O(\sum_{l=1}^3 n_1^l n_2^l \log(g^l)) + O(N_d)$. where $O(\sum_{l=1}^3 n_1^l n_2^l \log(g^l))$ and $O(N_d)$ are the computational complexity values of dense layers and dropout layer in MSADIDL model after DFEB of each mode. The following points are the key findings of the proposed approach.

- 1) The subbands or modes of thermogram images are computed using 2D-EWT with FBPs.
- 2) An interpretable DL architecture that considers different scale information of thermogram images is used for breast cancer detection.
- 3) The suggested MSADIDL approach has outperformed all transfer learning-based methods to detect breast cancer using thermogram images.
- 4) Compared to different existing methods, the MSADIDL approach has demonstrated higher accuracy for breast cancer detection.

For all the transfer learning networks, the pretrained weights are obtained from the models trained using images from the imangenet. Imagenet is a database of millions of training images of different categories (mainly objects and scenes) [47]. It has been observed that the imangenet-based source images' probability distributions vary significantly from the thermogram images. Due to this reason, all transfer learning models produce overall accuracy values less than 80% for both fivefold CV and hold-out validation cases to detect breast cancer. It can be more challenging if the pretrained weights are taken from the models that have been trained using another modality, such as mammogram images, and use these weights to detect breast cancer with the input as thermogram images for all transfer learning models. In the future, different modality-based transfer learning models can be implemented to detect breast cancer. In this study, we have used all 1520 thermogram images. The robustness of the proposed MSADIDL approach can be verified using more thermogram images. The suggested approach can detect other types of cancers using images from different medical imaging modalities. Automated detection of different types of breast cancer is challenging in clinical oncology and medical imaging [48], [49]. In the

future, novel DL-based methods can be developed for the automated detection of various types of breast cancer, such as lobular carcinoma in situ, ductal carcinoma in situ, etc., from thermogram images. The classification performance of the proposed MSADIDL model has been reduced using ultrasound images with an imbalanced dataset case. Hence, there is a scope to develop novel DL-based architectures with higher classification performance in detecting breast cancer using ultrasound images.

V. CONCLUSION

The MSADIDL-based approach has been presented in this article to detect breast cancer disease using thermogram images. The EWT with FBPs has been utilized to evaluate subbands from the thermogram image. The DFEB has been employed to extract learnable subband domain features. All local learnable features are merged and given to various dense layers to obtain reduced feature vectors for detecting breast cancer pathology using thermogram images. The proposed approach is interpretable and produced an accuracy value of more than 99.50% for both fivefold CV and hold-out validation cases to detect breast cancer. The suggested multi-scale domain approach demonstrated higher overall accuracy than different transfer learning-based techniques to detect breast cancer disease using thermogram images. The proposed MSADIDL approach has considered all frequency scale information and produced higher accuracy than the various existing thermogram image-based methods to detect breast cancer.

REFERENCES

- [1] J. G. Elmore, K. Armstrong, C. D. Lehman, and S. Fletcher, "Screening for breast cancer," *J. Amer. Med. Assoc.*, vol. 293, no. 10, pp. 1245–1256, 2005.
- [2] J. E. Bower, "Behavioral symptoms in breast cancer patients and survivors: Fatigue, insomnia, depression, and cognitive disturbance," *J. Clin. Oncol., Off. J. Amer. Soc. Clin. Oncol.*, vol. 26, no. 5, p. 768, 2008.
- [3] V. Levenson, "Biomarkers for early detection of breast cancer: What, when, and where?" *Biochim. Biophysica Acta, Gen. Subjects*, vol. 1770, no. 6, pp. 847–856, Jun. 2007.
- [4] H. G. Kaplan, J. A. Malmgren, M. K. Atwood, and G. S. Calip, "Effect of treatment and mammography detection on breast cancer survival over time: 1990–2007: Improved breast ca survival over time," *Cancer*, vol. 121, no. 15, pp. 2553–2561, Aug. 2015.
- [5] R. Sood et al., "Ultrasound for breast cancer detection globally: A systematic review and meta-analysis," *J. Global Oncol.*, vol. 5, no. 5, pp. 1–17, Dec. 2019.
- [6] C. K. Kuhl, "Abbreviated breast MRI for screening women with dense breast: The EA1141 trial," *Brit. J. Radiol.*, vol. 91, no. 1090, Oct. 2018, Art. no. 20170441.
- [7] L. R. Greene and D. Wilkinson, "The role of general nuclear medicine in breast cancer," *J. Med. Radiat. Sci.*, vol. 62, no. 1, pp. 54–65, Mar. 2015.
- [8] D. Singh and A. K. Singh, "Role of image thermography in early breast cancer detection-past, present and future," *Comput. Methods Programs Biomed.*, vol. 183, Jan. 2020, Art. no. 105074.
- [9] A. Mencattini, M. Salmeri, R. Lojacono, M. Frigerio, and F. Caselli, "Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 7, pp. 1422–1430, Jul. 2008.
- [10] S. H. Heywang-Köbrunner, A. Hacker, and S. Sedlacek, "Advantages and disadvantages of mammography screening," *Breast Care*, vol. 6, no. 3, pp. 199–207, 2011.

- [11] U. Raghavendra, A. Gudigar, T. N. Rao, E. J. Ciaccio, E. Y. K. Ng, and U. R. Acharya, "Computer-aided diagnosis for the identification of breast cancer using thermogram images: A comprehensive review," *Infr. Phys. Technol.*, vol. 102, Nov. 2019, Art. no. 103041.
- [12] B. Yousefi, H. M. Sharifipour, and X. P. V. Maldague, "A diagnostic biomarker for breast cancer screening via Hilbert embedded deep low-rank matrix approximation," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.
- [13] M. Milosevic, D. Jankovic, and A. Peulic, "Thermography based breast cancer detection using texture features and minimum variance quantization," *EXCLI J.*, vol. 13, p. 1204, Nov. 2014.
- [14] S. Pramanik, D. Bhattacharjee, and M. Nasipuri, "Wavelet based thermogram analysis for breast cancer detection," in *Proc. Int. Symp. Adv. Comput. Commun. (ISACC)*, Sep. 2015, pp. 205–212.
- [15] U. R. Acharya, E. Y. K. Ng, J.-H. Tan, and S. V. Sree, "Thermography based breast cancer detection using texture features and support vector machine," *J. Med. Syst.*, vol. 36, no. 3, pp. 1503–1510, Jun. 2012.
- [16] S. V. Francis, M. Sasikala, and S. Saranya, "Detection of breast abnormality from thermograms using curvelet transform based feature extraction," *J. Med. Syst.*, vol. 38, no. 4, pp. 1–9, Apr. 2014.
- [17] M. Abdel-Nasser, A. Moreno, and D. Puig, "Breast cancer detection in thermal infrared images using representation learning and texture analysis methods," *Electronics*, vol. 8, no. 1, p. 100, Jan. 2019.
- [18] R. Resmini, L. Silva, A. S. Araujo, P. Medeiros, D. Muchaluat-Saade, and A. Conci, "Combining genetic algorithms and SVM for breast cancer diagnosis using infrared thermography," *Sensors*, vol. 21, no. 14, p. 4802, Jul. 2021.
- [19] D. Sathish, S. Kamath, K. Prasad, and R. Kadavigere, "Role of normalization of breast thermogram images and automatic classification of breast cancer," *Vis. Comput.*, vol. 35, no. 1, pp. 57–70, Jan. 2019.
- [20] R. Roslidar et al., "A review on recent progress in thermal imaging and deep learning approaches for breast cancer detection," *IEEE Access*, vol. 8, pp. 116176–116194, 2020.
- [21] S. Dey, R. Roychoudhury, S. Malakar, and R. Sarkar, "Screening of breast cancer from thermogram images by edge detection aided deep transfer learning model," *Multimedia Tools Appl.*, vol. 81, no. 7, pp. 9331–9349, Mar. 2022.
- [22] J. Zuluaga-Gomez, Z. Al Masry, K. Benaggoune, S. Meraghni, and N. Zerhouni, "A CNN-based methodology for breast cancer diagnosis using thermal images," *Comput. Methods Biomed. Eng., Imag. Vis.*, vol. 9, no. 2, pp. 131–145, Mar. 2021.
- [23] J. Gilles, G. Tran, and S. Osher, "2D empirical transforms. Wavelets, ridgelets, and curvelets revisited," *SIAM J. Imag. Sci.*, vol. 7, no. 1, pp. 157–186, Jan. 2014.
- [24] N. Muralidharan, S. Gupta, M. R. Prusty, and R. K. Tripathy, "Detection of COVID19 from X-ray images using multiscale deep convolutional neural network," *Appl. Soft Comput.*, vol. 119, Apr. 2022, Art. no. 108610.
- [25] R. Panda, S. Jain, R. Tripathy, and U. R. Acharya, "Detection of shockable ventricular cardiac arrhythmias from ECG signals using FFREWT filter-bank and deep convolutional neural network," *Comput. Biol. Med.*, vol. 124, Sep. 2020, Art. no. 103939.
- [26] L. F. Silva et al., "A new database for breast research with infrared image," *J. Med. Imag. Health Informat.*, vol. 4, no. 1, pp. 92–100, Mar. 2014.
- [27] R. B. Pachori, *Time-Frequency Analysis Techniques and Their Applications*. Boca Raton, FL, USA: CRC Press, 2023.
- [28] N. Bjorck, C. P. Gomes, B. Selman, and K. Q. Weinberger, "Understanding batch normalization," *Adv. neural Inf. Process. Syst.*, vol. 31, 2018.
- [29] Z. Li, S. Wang, R. Fan, G. Cao, Y. Zhang, and T. Guo, "Teeth category classification via seven-layer deep convolutional neural network with max pooling and global average pooling," *Int. J. Imag. Syst. Technol.*, vol. 29, no. 4, pp. 577–583, Dec. 2019.
- [30] J. Chen, H. Jing, Y. Chang, and Q. Liu, "Gated recurrent unit based recurrent neural network for remaining useful life prediction of nonlinear deterioration process," *Rel. Eng. Syst. Saf.*, vol. 185, pp. 372–382, May 2019.
- [31] A. Varshney, S. K. Ghosh, S. Padhy, R. K. Tripathy, and U. R. Acharya, "Automated classification of mental arithmetic tasks using recurrent neural network and entropy features obtained from multi-channel EEG signals," *Electronics*, vol. 10, no. 9, p. 1079, May 2021.
- [32] M.-T. Luong, H. Pham, and C. D. Manning, "Effective approaches to attention-based neural machine translation," 2015, *arXiv:1508.04025*.
- [33] C. Corneauan, M. Madadi, S. Escalera, and A. Martinez, "Explainable early stopping for action unit recognition," in *Proc. 15th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, Nov. 2020, pp. 693–699.
- [34] J. Karhade, S. Dash, S. K. Ghosh, D. K. Dash, and R. K. Tripathy, "Time–frequency-domain deep learning framework for the automated detection of heart valve disorders using PCG signals," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.
- [35] S. K. Ghosh, R. N. Ponnalagu, R. K. Tripathy, G. Panda, and R. B. Pachori, "Automated heart sound activity detection from PCG signal using time–frequency-domain deep neural network," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–10, 2022.
- [36] A. Khamparia et al., "Diagnosis of breast cancer based on modern mammography using hybrid transfer learning," *Multidimensional Syst. Signal Process.*, vol. 32, no. 2, pp. 747–765, Apr. 2021.
- [37] G. Ayana, K. Dese, and S.-W. Choe, "Transfer learning in breast cancer diagnoses via ultrasound imaging," *Cancers*, vol. 13, no. 4, p. 738, Feb. 2021.
- [38] N. Aidossov et al., "Evaluation of integrated CNN, transfer learning, and BN with thermography for breast cancer detection," *Appl. Sci.*, vol. 13, no. 1, p. 600, Jan. 2023.
- [39] *Keras Transfer Learning Models*. Accessed: Jun. 15, 2022. [Online]. Available: <https://keras.io/api/applications/>
- [40] D. Sarkar, R. Bali, and T. Ghosh, *Hands-On Transfer Learning With Python: Implement Advanced Deep Learning and Neural Network Models Using TensorFlow and Keras*. Birmingham, U.K.: Packt Publishing, 2018.
- [41] S. Bhaskarpandit, A. Gade, S. Dash, D. K. Dash, R. K. Tripathy, and R. B. Pachori, "Detection of myocardial infarction from 12-lead ECG trace images using eigendomain deep representation learning," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–12, 2023.
- [42] D. Zhang and D. Zhang, "Wavelet transform," in *Fundamentals of Image Data Mining: Analysis, Features, Classification and Retrieval*. Cham, Switzerland: Springer, 2019, pp. 35–44.
- [43] J.-M. Deng, H.-Z. Yue, Z.-Z. Zhuo, H.-G. Yan, D. Liu, and H.-Y. Li, "A stationary wavelet transform based approach to registration of planning CT and setup cone beam-CT images in radiotherapy," *J. Med. Syst.*, vol. 38, no. 5, pp. 1–9, May 2014.
- [44] Y. Gu et al., "Deep learning based on ultrasound images assists breast lesion diagnosis in China: A multicenter diagnostic study," *Insights Imag.*, vol. 13, no. 1, p. 124, Dec. 2022.
- [45] Y. Li, H. Gu, H. Wang, P. Qin, and J. Wang, "BUSnet: A deep learning model of breast tumor lesion detection for ultrasound images," *Frontiers Oncol.*, vol. 12, Mar. 2022, Art. no. 848271.
- [46] M. Rotman and L. Wolf, "Shuffling recurrent neural networks," in *Proc. AAAI Conf. Artif. Intell.*, 2021, vol. 35, no. 11, pp. 9428–9435.
- [47] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
- [48] W. A. Berg and J. W. Leung, *Diagnostic Imaging: Breast E-Book*. Amsterdam, The Netherlands: Elsevier, 2019.
- [49] K. Atrey, B. K. Singh, N. K. Bodhey, and R. Bilas Pachori, "Mammography and ultrasound based dual modality classification of breast cancer using a hybrid deep learning approach," *Biomed. Signal Process. Control*, vol. 86, Sep. 2023, Art. no. 104919.