Enhanced Artificial Intelligence System for Diagnosing and Predicting Breast Cancer using Deep Learning

Mona Alfifi[1], Mohamad Shady Alrahhal[2], Samir Bataineh[3], Mohammad Mezher[4]

Department of Computer Science[1, 2, 3, 4]

Fahad Bin Sultan University, Tabuk City, King Saudi Arabia[1, 3, 4]

King Abdulaziz University, Jeddah City, King Saudi Arabia[2]

Abstract—**Breast cancer is the leading cause of death among women with cancer. Computer-aided diagnosis is an efficient method for assisting medical experts in early diagnosis, improving the chance of recovery. Employing artificial intelligence (AI) in the medical area is very crucial due to the sensitivity of this field. This means that the low accuracy of the classification methods used for cancer detection is a critical issue. This problem is accentuated when it comes to blurry mammogram images. In this paper, convolutional neural networks (CNNs) are employed to present the traditional convolutional neural network (TCNN) and supported convolutional neural network (SCNN) approaches. The TCNN and SCNN approaches contribute by overcoming the shift and scaling problems included in blurry mammogram images. In addition, the flipped rotation-based approach (FRbA) is proposed to enhance the accuracy of the prediction process (classification of the type of cancerous mass) by taking into account the different directions of the cancerous mass to extract effective features to form the map of the tumour. The proposed approaches are implemented on the MIAS medical dataset using 200 mammogram breast images. Compared to similar approaches based on KNN and RF, the proposed approaches show better performance in terms of accuracy, sensitivity, spasticity, precision, recall, time of performance, and quality of image metrics.**

Keywords—Traditional Convolutional Neural Network (TCNN); Supported Convolutional Neural Network (SCNN); shift; scaling; cancer detection; mammogram; histogram equalization; adaptive median filter

I. INTRODUCTION

The use of artificial intelligence (AI) in the medical sector has many benefits. In light of the danger of spreading COVID-19, machine learning plays a significant role in limiting the number of infected people. The reason is that AI reflects an automation process that can be performed by machines on behalf of users (i.e. doctors, nurses, or medical staff in the medical sector) in regard to diagnosis [1, 2]. Another benefit of AI in the medical sector is the lowering of the cost of treatment and enhancement of the chance of survival [3, 4]. Early breast cancer detection is an excellent example to support this argument in this context. In general, 10 AI applications and advantages can be exploited for the sake of serving humanity, as illustrated by the 2026 estimation in Fig. 1.

Statement of problem. One of the most important causes of the increased death ratio among women society (in contrast to men) is breast cancer. Research has shown that the number of people who died due to breast cancer continued to increase from 2007 to 2018, as shown in Fig. 2.

From a medical point of view, the early detection of breast cancer contributes to saving the lives of patients as well as decreasing the cost of treatment at both the private and governmental medical institution levels. Computer science researchers have employed AI for this purpose, and many approaches have been proposed, such as [7-21]. However, the quality of any proposed approach for breast cancer detection is evaluated based on its accuracy. In other words, due to the sensitivity of diagnosis in the medical field, a high error rate is critical and may lead to death. In the context of AI, the error rate is represented by the false positive (FP) cases that the intelligent machine fails to classify correctly [22, 23]. In real life, it is harmful for the patient to provide a diagnosis result of negative while the diagnosis is actually positive. It is worth mentioning that most of works reviewed in the related work section suffer from this problem. Fig. 3 illustrates the problem being addressed in this work.

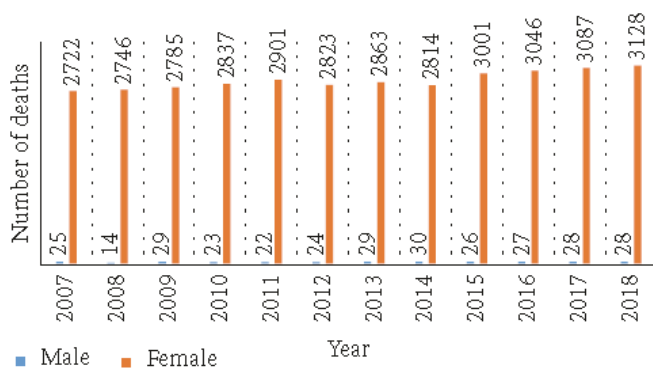Fig. 1.   Benefits of 10 AI Applications in 2026 [5].

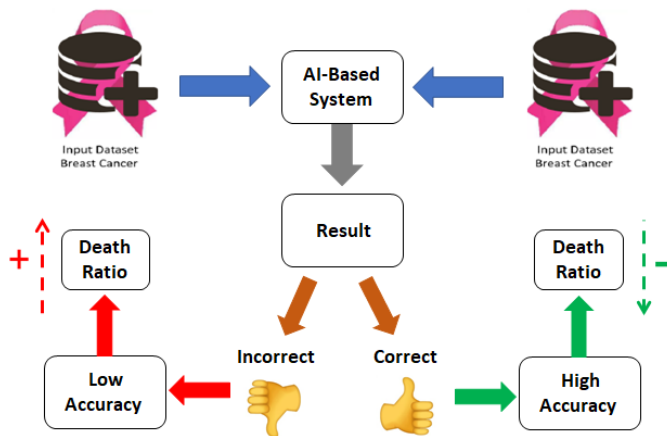Fig. 2. Statistical Survey for Death caused by Breast Cancer [6].

Fig. 3. Problem of AI-based Breast Cancer Detection Systems.

Research questions. On one hand, mammography images are common in hospitals and used for breast cancer detection. On the other hand, image blurring is a recognized issue in the UK [24, 25]. Blurry mammography images are considered the root cause of lower accuracy in AI-based diagnosis systems if they are used to train a classifier. Blurry images are generated because of breast/paddle movement while the exposure is being made [26, 27]. There are also other contributors that negatively impact the clarity of mammography images and are related to inadequate compression, and patient movement together with long exposures may also cause blurring [28]. Such blurring may cause distortion of the generated mammography images in terms of shift and scaling. Therefore, the main research question is how to ensure that AI-based systems generate diagnosis results with high accuracy for noisy mammography images. This in turn leads to the following two questions:

1) How can AI-based systems detect cancer in the breast even if a mammography image has a shift problem?

2) How can AI-based systems control the scaling problem so that it does not affect the accuracy of the detection system?

Using conventional neural networks (CNNs) can contribute to answering the previous questions and lead to better accuracy. That is because we can exploit the architecture of the CNN to control the shifting problem based on conventional layers, while the scaling problem can be absorbed by pooling functions.

In general, the contributions of this works can be listed as follows:

- In responding to the shift problem caused by blurring in mammography images, the architecture of the CNN technique is employed. The padding operation is used in the proposed Traditional convolution neural network (TCNN) to overcome this problem.

- To solve the scaling problem caused by blurring in mammography images, the architecture of the CNN technique is employed. The max function used in the proposed supported convolution neural network (SCNN) to construct the pooling layers is used to overcome this problem.

- To increase the accuracy of the classification process, the flipped rotation-based approach (FRbA) is introduced.

- Extensive experiments are conducted to show the effectiveness of the proposed methods and compare with similar approaches.

The rest of this work is organized as follows. Section II reviews the related work. Section III provides the proposed CNN classification model. In Section IV, the used metrics are presented for evaluation purposes. Section V presents the experiments and discusses the results in light of a comparison with similar approaches. Finally, the paper is concluded in Section VI.

II. RELATED WORK

This section provides a brief background regarding the domain we stand on, followed by a review of approaches that have previously been proposed in the breast cancer detection research field under the AI umbrella.

A. Background

The domain that addresses breast cancer diagnosis, in terms of AI, can be represented by an intersection of four main disciplines, as shown in Fig. 4.
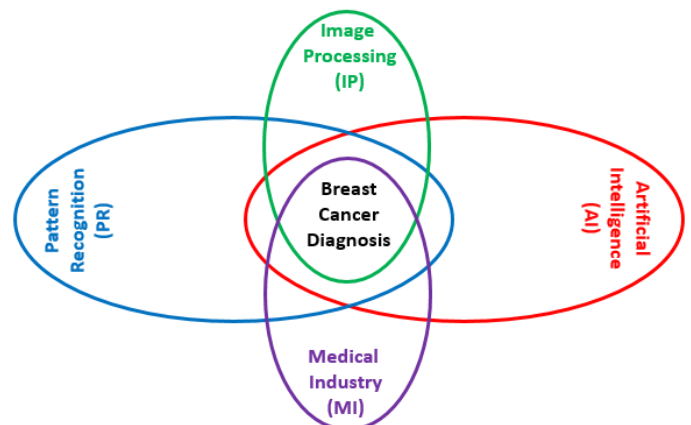

Fig. 4. Intersection of Disciplines that Represent Breast Cancer Diagnosis.

Fig. 4 shows that AI, image processing (IP), medical industry (MI), and pattern recognition (PR) disciplines are involved in the breast cancer diagnosis research field. AI is defined as the science that studies the capability of granting machines the ability to make decisions on behalf of humans [29]. IP is defined as methods of performing some mathematical operations on images based on computer algorithms to obtain improved images or extract various information or features related to the images [30]. PR can be seen as automated learning algorithms that are used to identify patterns such that data are classified into various groups or categories based on the information extracted from the patterns [31]. As for MI, it is a sector within the economic system that provides goods and services to treat patients with curative, preventive, rehabilitative, and palliative care [32].

B. Taxonomy of AI-based Approaches for Breast Cancer Diagnosis

Many efforts have been made to group AI-based systems concerning breast cancer detection, such as [6] and [33-35]. Ref [6] provides details about the description of breast cancer image classification techniques, feature extraction and selection procedures, classification measurement parameterizations, and image classification findings. In addition, it focuses on analysing three main techniques, which are the CNN, random forest (RF), and support vector machine (SVM). The authors of [33] present classification from a medical point of view (i.e. relying on types of cancer) along with four popular architectures used for cancer detection and diagnosis, which are CNN, fully convolutional networks (FCNs), auto-encoders (AEs), and deep belief networks (DBNs). In [34], the authors aim to investigate the state of the art regarding computer-aided diagnosis/detection (CAD) systems for breast cancer by presenting a systematic review. What makes this work recognizable is the good presentation of the databases available for conducting experiments for breast cancer detection. In terms of advantages and disadvantages, [35] summarizes a wide spectrum of the approaches for breast cancer detection. In addition, the performance of the reviewed approaches is investigated and analysed.

In this research, we classify the techniques used for breast cancer detection in the domain of AI based on the three main groups used in this science. Fig. 5 illustrates the supervised, unsupervised, and semi-supervised groups, where each category has its own techniques/algorithms.
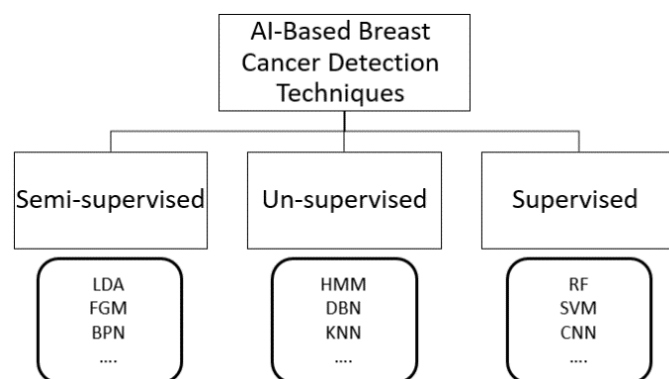

Fig. 5.   Groups of AI-based Systems.

1) Supervised learning-based group: Using some data mining techniques, a comparison study is presented in [7] with the goal of identifying the best technique for breast cancer detection. The authors apply RF, SVM, and Naïve Bayes (NB) on two different data sets, which are INbreast and Breast Cancer Digital Repository (BCDR). The findings show that the RF method is ranked at the top with an accuracy value of 78.3%.

In [8], a super-resolution (SR) approach is proposed that exploits the complementary information provided by multiple images of the same target. In this work, the CAD system consists of four stages: (1) SR computation; (2) extraction of the region of interest; (3) feature extraction; and (4) classification. It is worth mentioning that the authors focus on methods for extracting texture features from a single breast ultrasound (BUS) image for each breast, which mainly suffers from speckle noise and artefacts. As for the dataset used to conduct experiments, the authors used a dataset named Ultrasound Images Private (31 malignant, 28 benign). The proposed CAD system achieves a value of 0.99 in terms of the area under the receiver operating characteristic (ROC) curve (AUC).

The objective of [9] is to evaluate the performance of a system for the detection of masses based on blind feature extraction using independent component analysis (ICA). The system uses different classifiers, such as neural networks (NNs) and SVMs. A conversion process of the image to optical density is performed first as a processing stage. Later, the regions of interest are appropriately resized to facilitate the task of the feature extractor. Finally, the classifiers are applied to detect and classify masses in the mammograms images. The Digital Database for Screening Mammography (DDSM) database is used (5052 images) in this work for training and testing. According to AUC-based evaluation, the proposed method achieves a value of 0.937.

The method presented in [10] aims to help medical experts in the classification of breast cancer lesion through the implementation of Convolutional Neural Network Improvement for Breast Cancer Classification (CNNI-BCC). In their method, the authors 1) decompose the images into patches; (2) have each patch process non-rectangular regions in the image via the masking of certain areas; (3) train on small data sets with a combined rained CNN; and (4) have the presented CNNI-BCC utilize the patch feature relevance to detect and classify the lesion region. The approach is tested using the Mammography Image Analysis Society (MIAS, 221 images) dataset and achieves 90.50% accuracy.

The goal of [11] is to propose a prototype of a medical expert system (based on data mining techniques) that can significantly aid medical experts in detecting breast cancer. The proposed system consists of four major steps: (1) image processing using the discrete wavelet transformation (DWT); (2) extraction of features of the region of interest (ROI); (3) formation of a total of 20 grey level co-occurrence matrix (GLCM) features from the ROI, which are in turn used as inputs for classification algorithms; and (4) application of the actual classification process, where seven different classifiers are used, i.e., SVM, RF, C4.5, K-nearest neighbour (KNN),

naive Bayes (NB), logistic regression (LR), and multilayer perceptron (MLP). In this study, normal breast images and breast images with masses (322 images in total) are manipulated in the training and testing phases. The images are taken from the MIAS database. The SVM and MLP methods achieve accuracies of 74% and 76%, respectively.

A conditional generative adversarial network (cGAN) approach is proposed in [12]. The key idea is to segment the breast tumour within an ROI in mammogram images. The generative network learns to recognize the tumour area and to create the binary mask that outlines it. Experiments are conducted on two datasets, the public INbreast and a private in-house dataset. The proposed segmentation model provides a high dice coefficient and intersection over union (IoU) of 94% and 87%.

2) Unsupervised learning-based Group: Relying on both the wavelet and the hidden Markov model (HMM), a breast cancer detection-based system is provided in [13]. The key idea behind the detection is to identify suspicious areas of breast cancer tumours. Mammographic images are used in this work and processed using a median filter to extract image features. This step is performed to clarify and improve the quality of mammographic images. A combination of MIAS and Paden databases is used to test the proposed method. The detection rate of the proposed method is 96%.

In [14], Al-antari et al. proposed a deep belief network (DBN)-based CAD system to classify breast tissues. Three classes are used in the classification process, which are normal, benign, and malignant. What distinguishes this work is the automatic mass detection algorithm for identifying suspicious masses on mammograms. Then, statistical features are derived from the detected masses. The authors used the DDSM dataset (150 images). The overall accuracies of a DBN are 92.86% and 90.84% for the two ROI techniques, respectively.

Depending on the clustering principle, a microcalcification cluster enhancement method is presented and implemented on digital mammogram images in [15]. The authors handle the dependence problem via the segmentation phase, where a Laplacian of Gaussian filter (LoG) is employed. This leads to optimization of the recognition efficiency. As for the clustering method, fuzzy C-means (FCM) is utilized. The authors used the MIAS dataset to build the detection system. The method is evaluated in terms of accuracy and achieves 95%.

In [16], instead of working in the spatial domain, the authors deal with the frequency domain using a discrete curvelet transform to extract features of the images (image map). The medical images used in this work are mammogram images, and the manipulation passes through three main steps. First, the discrete curvelet transform is applied to the images to compute the four first-order moments. Consequently, two feature sets are obtained: moments from each band and moments from each level. Second, the t-test ranking technique is applied to select the best features from each set. Finally, the KNN classifier is used to distinguish between normal and abnormal breast tissues and classify tumours as malignant or benign. Experiments are performed depending on both the MIAS (252 images) and DDSM (11,553 images) datasets. The

results on the mini-MIAS database show that curvelet moments yield an accuracy of 91.27%.

Focusing on obtaining good mammogram images as an objective, the authors of [17] present the connected component labelling analysis (CCL), which is a multi-level segmentation approach based on artefacts and pectoral muscle removal to ensure good clarity of mammogram images. For the feature extraction and classification stages, they use statistical and Kernel self-optimized fisher discriminant (KSFD) methods, respectively. The MIAS dataset is utilized to implement the proposed approach and conduct the experiments. The KSFD method exhibits an accuracy of 94.46%.

3) Semi-supervised learning-based Group: Ganesan et al. [18] propose a framework for the automated detection of normal, benign, and cancerous mammograms. They rely on using a combination of texture, higher-order spectra (HOS), and discrete wavelet transform (DWT) for feature extraction from digitized mammograms. These features capture the subtle variation in the pixel intensities and contours in the images and serve as significant indicators for the classification process. Decision tree (DT), Fisher, linear discriminant classifier (LDC), nearest mean classifier (NMC), Parzen, and SVM classifiers are used to classify the features. Among the classifiers evaluated, DT performs the best. The accuracy of the proposed method is 91.6% for the DDSM database and 96.8% for SATA.

The goal of [19] is to introduce a classification algorithm based on the fuzzy Gaussian mixture model (FGMM). This is achieved by combining the power of the Gaussian mixture model (GMM) and fuzzy logic system (FLS) to build a CAD system. This system is responsible for classifying the detected regions in mammogram images into malignant or benign categories. The dataset used for this work is a subset of the DDSM (300 images). The results show that the proposed FGMM classifier has achieved an overall accuracy of prediction of 93%.

A tool for the diagnosis of mammography abnormalities is proposed in [20]. The key idea is to depend on semi-supervised classification using a transductive support vector machine (TSVM) with different kernel functions and heterogeneous feature families. The DDSM database (200 images) is used to implement the proposed approach and obtain the results. The TSVM method exhibits an accuracy of 93.1%.

The rough set theory (RST) method is provided in [21]. The RTS is used for many functions. First, it is integrated with back-propagation to handle more uncertain data. Second, RST has also been used to remove pectoral muscles and segmentation. As for artefacts and labels, they are removed using a vertical and horizontal sweeping method. Mammogram images are acquired from the MIAS database. Features are extracted from the segmented mammogram image using GLCM, GLDM, SRDM, NGLCM, and GLRM. Finally, the features are normalized, discretized, and then reduced using RST (as the third and final function of RST). After that, the classification is performed using RNN. The RNN classifier achieves a 99.2% overall classifying accuracy.

III. PROPOSED APPROACH

In this work, we propose two approaches for breast cancer detection, illustrating the impact of image processing on the accuracy of the classification process. The first approach is called the traditional CNN (TCNN)-based approach. The second is called the supported CNN (SCNN)-based approach. Both the TCNN and SCNN approaches follow the same steps shown in Fig. 6.

As shown in Fig. 6, there are 6 steps in the proposed approaches, starting with selection of the dataset and finishing with the output (i.e., the type of breast cancer). It is worth mentioning that the difference between the TCNN and the SCNN is highlighted by step 2 (image pre-processing).

In general, the system that implements the proposed approaches has an input and output. The input is a mammogram image, while the output is one of the three types of cancer (normal, benign, and malignant). Mathematically, x_{image}^{mam} refers to the input image, y_{type} refers to the output of the system (classifier), and $S = \{N.B.M\}$ is the set of the types of cancer. Then, the classifier is modelled as:

$$Classifier\ (x_{image}^{mam}) = y_{type} | type \in S \qquad (1)$$

A. Selecting the Dataset

Due to the sensitivity of the medical area, it is difficult to access real medical images to conduct experiments. The sensitivity of the medical area is reflected by a high level of security and privacy issues [36-43]. In this work, the data are taken from the MIAS [44]. Table I summarizes the information of the MIAS dataset.

B. Image Pre-Processing

The objective of this step is to enhance the quality of the mammogram images. This in turn leads to higher classification accuracy. The reason behind this is that, the clearer edges and details are, the greater the ability to identify the area of the tumour is, and consequently the more accurate diagnosis is.

Formally, let x_{image}^{en-mam} denote the enhanced mammogram image (i.e. the image after the pre-processing step). Then, both the TCNN and SCNN approaches use the same x_{image}^{en-mam} symbol as an output of the pre-processing step. However, the mechanism of generating the enhanced mammogram image differs, as shown in Fig. 7.

In the TCNN approach, the input mammogram image x_{image}^{mam} is processed using the Histogram Equalization ($HisE$) function to generate the enhanced mammogram image according to:

$$HisE\ (x_{image}^{mam}) = x_{image}^{en-mam} \qquad (2)$$

HisE is defined as a technique for adjusting image intensities to enhance contrast [45]. Mathematically, let f be a given image represented as an m_r-by-m_c matrix of integer pixel intensities ranging from 0 to $\alpha - 1$. α is the number of possible intensity values, often 256. Let p denote the normalized histogram of f with a bin for each possible intensity. Thus,

$$p_i = \frac{number\ of\ pixels\ with\ intensity\ i}{total\ number\ of\ pixels} \qquad i = 0.1. \dots . \alpha - 1 \quad (3)$$

The histogram equalized image ($img = x_{image}^{mam}$) will be defined by

$$img_{a.b} = floor((\alpha - 1) \times \sum_{i=0}^{f_{a.b}} p_i) \qquad (4)$$

where $floor()$ rounds down to the nearest integer.

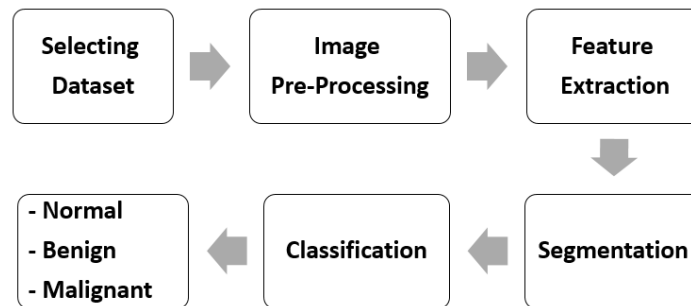Fig. 8 shows the impact of the $HisE$ function on the original breast mammogram image.

Fig. 6. Flowchart of the Proposed Approaches.

TABLE I. MIAS DATASET DESCRIPTION

| NO of Images	322 images							
Type	Normal	Abnormal						
		Benign			Malign			
Abnormal Categories	Circumscribed Masses	Spiculated Masses	Microcalcifications		Ill-defined Masses	Architectural Distortion		Asymmetry
Size	1024×1024 pixels							
Accuracy	8-bit accuracy (grey level)							
Location Information	Centre of a circle surrounding the tumour	Radius	Breast Position		Type of breast tissues			Tumour type
			Left	Right	Fatty	Fatty-glandular	Dense	Benign	Malign
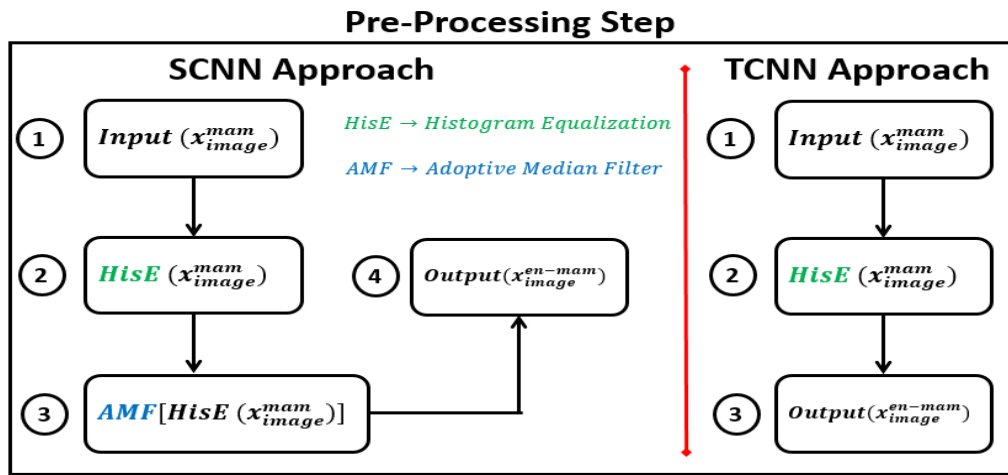
Fig. 7. Details of the Pre-Processing Step for Both the TCNN and SCNN Approaches.
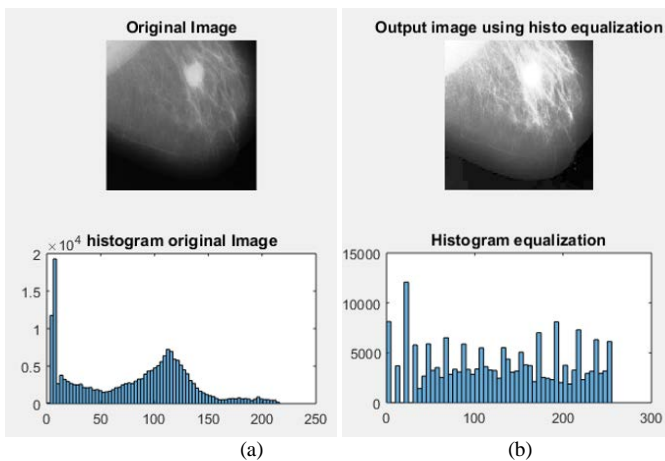

Fig. 8. Breast mammogram image after applying HisE. (a) The original image with the corresponding histogram. (b) The details of the breast are clearer; however, some blurring is generated due to HisE.
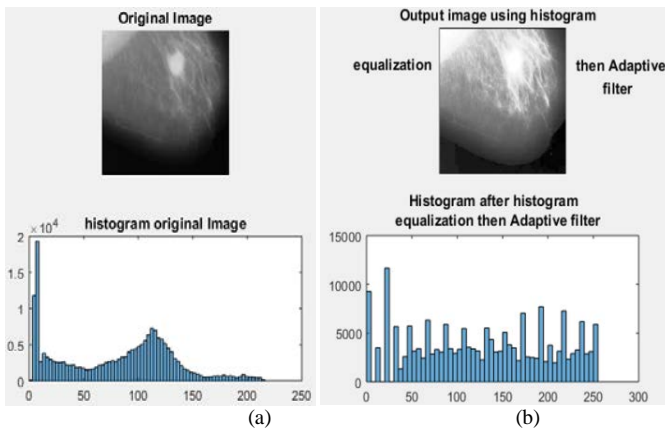

Fig. 9. Breast mammogram image after applying AMF (HisE). (a) The original image with the corresponding histogram. (b) The image after enhancement is performed using AMF.

The AMF algorithm works in two levels, denoted as level alpha (α) and level beta (β), as follows.

Where we take into consideration the following notations arranged in Table II, we define the AMF algorithm.

Algorithm 1: Adoptive Median Filter (AMF)
1: begin
2:    Label Level α:
3:       begin
4:          $\alpha_1 = In_{med} - In_{min}$;
5:          $\alpha_2 = In_{med} - In_{max}$;
6:          if $\alpha_1 > 0$ and $\alpha_2 < 0$ then
7:             begin
8:               go to level β;
9:             end;
10:         *else* increase the w_{xy};
11:         if $w_{xy} \leq w_{max}$ then
12:             begin
13:               repeat level α;
14:             end;
15:         *else* output In_{med};
16:    end;
17:    label Level β:
18:       begin
19:          $\beta_1 = In_{xy} - In_{min}$;
20:          $\beta_2 = In_{xy} - In_{max}$;
21:          if $\beta_1 > 0$ and $\beta_2 < 0$ then
22:             begin
23:               output In_{xy};
24:             end;
25:         else output In_{med};
26:    end;
27: end;

TABLE II. NOTATIONS USED FOR AMF

Notation	Description
w_{xy}	Rectangular window area.
In_{min}	Minimum intensity value in w_{xy}.
In_{med}	Median of intensity values in w_{xy}.
In_{xy}	Intensity value at coordinates (x,y).
In_{max}	Maximum intensity value in w_{xy}.
w_{max}	Maximum allowed size of w_{xy}.

In the SCNN approach, a further enhancement is performed using an adaptive median filter (AMF), as shown in Fig. 9 above. Since the blurring generated by $HisE$ can be considered as a kind of noise or distortion, AMFs contribute by adding additional enhancement for the mammogram input images. That is because they have the following benefits [46]:

1) Removal of salt and-pepper (impulse) noise.

2) Smoothing of other noise (may not be impulsive).

3) Reduction of distortion, such as excessive thinning or thickening of object boundaries.

In the SCNN approach, the input mammogram image $HisE\ (x_{image}^{mam})$ is processed using the AMF function to generate the enhanced mammogram image according to:

$$AMF[HisE\ (x_{image}^{mam})] = x_{image}^{en-mam} \qquad (5)$$

C. Feature Extraction

In general, feature extraction techniques have some advantages as listed below:

1) Accuracy improvements.
2) Overfitting risk reduction.
3) Acceleration of training.
4) Improved data visualization.
5) Increase in explainability of our model.

The basic objective of the feature extraction step is to reduce the number of features in a dataset by creating new features from the existing ones (and then discarding the original features). This new reduced set of features should then be able to summarize most of the information contained in the original set of features [47].

In the CNN architecture, feature extraction is achieved by forming conventional layers. In depth, the conventional layers are generated by scanning the mammogram input image by a certain kernel (filter). Fig. 10 illustrates the architecture of the CNN as well as the general mechanism used to generate conventional layers.

The more features extracted, the greater the ability of the CNN model to define and classify the Region of Interest (RoI), where RoI refers to the cancerous mass. Depending on the previous fact, a flipped rotation-based approach (FRbA) is proposed to generate more useful features. The key idea behind the FRbA is to apply rotations as mathematical operations to give different views of the cancerous mass located within the mammogram input image. Fig. 11 shows the general idea of the FRbA.
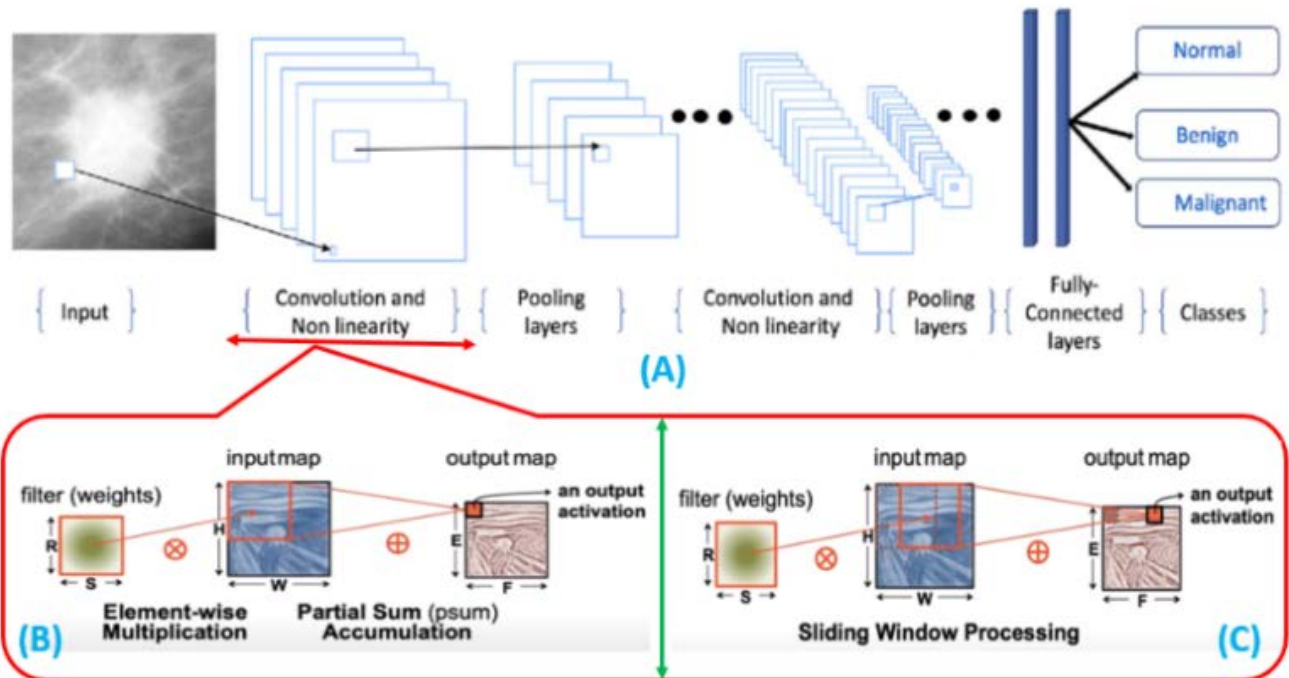

Fig. 10. Architecture of the CNN with Details of Forming Convolution Layers. (A) Represents the General Structure of the CNN model, where many Convolution Layers are Generated. (B) Illustrates the First Step of the Mechanism used to form the Convolution Layers. The Kernel (Filter) is Applied on the Image by Calculating the Sums of Multiplications (i.e. Multiplying the Weights of the Filter by the Intensity Values of the Input Mammogram Image and then Summing the Multiplications). (C) Illustrates the Sliding Window of the Filter to Scan the Entire Image, Covering all of the Pixels of the Image.
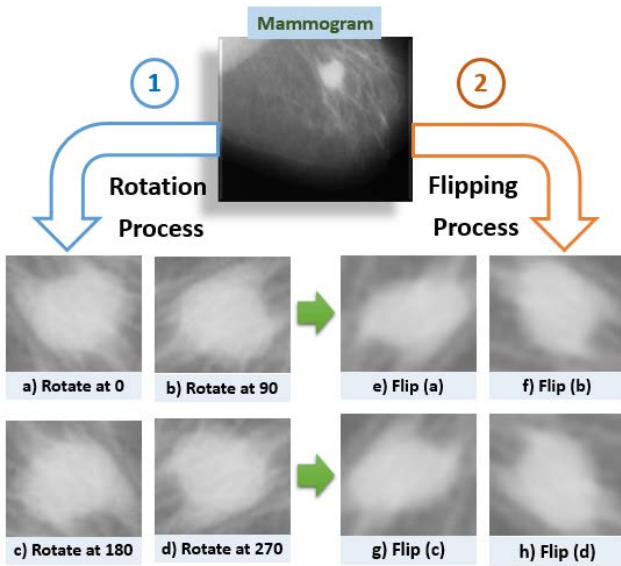
Fig. 11. General view of the FRbA.

Algorithm 2 shows the pseudocode of the FRbA in steps.

Algorithm 2: Flipped Rotation-based Approach (FRbA)

1:   begin

2:    Read Input (x_{image}^{b-mam} , x_{image}^{m-mam} , x_{image}^{n-mam});

3:    Pre-process (x_{image}^{b-mam} , x_{image}^{m-mam} , x_{image}^{n-mam});

4:    (x_{image}^{b-mam} , x_{image}^{m-mam} , x_{image}^{n-mam}) subsample according to transformation ;

5:    *For* (x_{image}^{b-mam}) , *ROTATE* to 90º, 180º, and 270º;

6:    *For* (x_{image}^{m-mam}), *ROTATE* to 90º, 180º, and 270º;

7:    *For* (x_{image}^{n-mam}), *ROTATE* to 90º, 180º,and 270º;

8:    Till all images had been re-aligned;

9:    Perform FLIP on all aligned images;

10:    Write all images in as subsampled images;

11:    Repeat on other training data;

12:   end;

After applying rotation operations, the flipping operation is performed, and the resultant images undergo the filtering process (i.e. forming the convolution layers using kernels). The forming of convolution layers requires passing the filter over all of the pixels of images. The values of the cells of the matrix filter represent the weights of the CNN. The new values that form the convolution layers are computed based on the sums of the multiplication in the CNN. This contributes to solving the shift problem because the first few pixels (that form the boundary) in the input image are manipulated by the filters based on a padding process. Fig. 12 illustrates the idea behind solving the shift problem.

Solving shift problem. In Fig. 12, the region of shifting (RoS) is constructed by the padding process. This in turn gives

space to shift the image in the four directions without any resultant distortion in terms of math. That is because the multiplications by zero do not affect the result of the calculation (i.e. sums of the multiplications). However, the RoS is limited by the size of the filter that is used. Therefore, a distortion is generated in the case exceeding the region (in Fig. 12, no more than 1-pixel shifting). When using the 5×5-size filter, the RoS is 2 pixels. In general, the RoS is defined in terms of the filter's size as:

$$RoS = \frac{sum\ of\ rows\ (filter) - 1}{2} \tag{6}$$

Solving the scaling problem. It is worth mentioning that (C) in Fig. 13 illustrates the key idea behind the solution of the scaling problem. In other words, scaling means that the intensity values of a given image may increase (referred to as scaling up or scaling with a positive "+" value), decrease (referred to as scaling down or scaling with a negative "-" value), or remain constant (which is a special case of scaling that leads to the same image without any change or scaling with a zero "0" value). Because the max function is employed in the pooling process, all values that are lower than or equal to the maximum values of each quarter of the image will be contained automatically in the output. In (C) in Fig. 13, the first quarter has increased intensity values (19 becomes 22, for example). Although this change is performed, this does not affect the output of the pooling process. This in turn leads to the scaling problem being absorbed by the max function.

D. Segmentation

The objective of this step is to identify the edges as well as the features included within the edges (i.e. the pooled features extracted from the previous step). To this end, the Gaussian mixture model (GMM) is used. The GMM is a weighted sum of k component Gaussian densities. Formally, it is represented by:

$$p(x|\delta) \sum_{r=1}^{k} \omega_r\ g(x|\beta_r.\textstyle\sum_r) \tag{7}$$

where x is a D-dimensional continuous-valued data vector (i.e., measurements or features), ω_r, $r =1, 2, ..., k$ are the mixture weights, and $g(x|\beta_r.\sum_r)$, $r =1, 2, ..., k$ are the component Gaussian densities. Each component density is a D-variate Gaussian function of the form:

$$g(x|\beta_r.\textstyle\sum_r) = \frac{1}{(2\pi)^{D/2}|\sum_r|^{1/2}} exp\left\{ -\frac{1}{2}(x - \beta_r)' \textstyle\sum_r^{-1}(x - \beta_r) \right\} \tag{8}$$
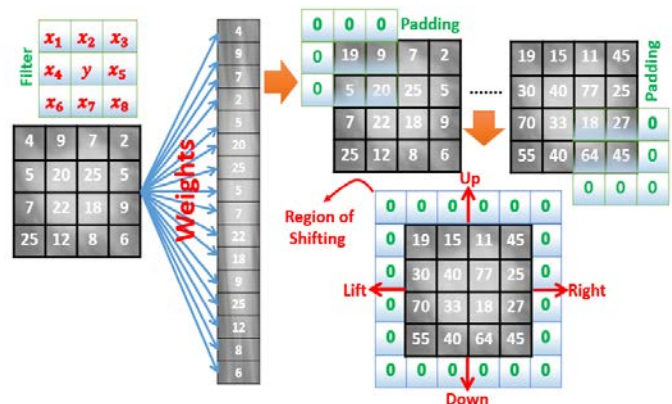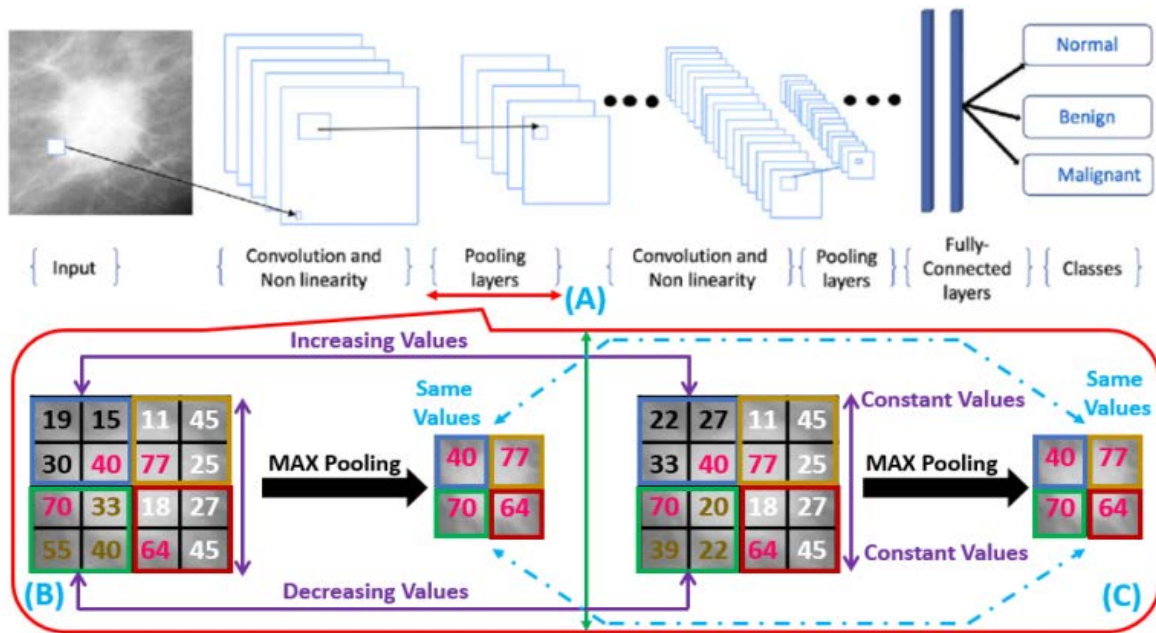

Fig. 12. Solving Shift Problem.

Fig. 13. Details of the Construction of the Pooling Layers. (A) Represents the General Structure of the CNN Model, where many Pooling Layers are Generated. (B) Actual Pooling Process based on the Max Function, where the Highest Values of Intensities are Selected to form Pooling Layers (40, 77, 70, and 64 are the Highest Values that Correspond to the Four Quarters of the Intensity Matrix). (C) Illustrates the Three main Options that may Occur During the Manipulation (Increasing Values of Some Intensities, Decreasing Values of Some Intensities, and Remaining Constant). In Spite of the Three Options Occurring at the Same Time, the Output is the Same (i.e. 40, 77, 70, and 64).

with mean vector β_r and covariance matrix Σ_r. The mixture weights satisfy the constraint that $\sum_{r=1}^{k} \omega_r = 1$. The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices and mixture weights from all component densities [48]. These parameters are collectively represented by the notation:

$$\delta = \{\omega_r\,,\,\beta_r\,,\,\Sigma_r\}\ |r\ = 1.2.\ldots.k \qquad (9)$$

Visually, the output of the GMM for a pre-processed mammogram image is shown in Fig. 14.

E. Classification

This step is the final one in the prediction process. The objective of this step is to show the output of the classifier (T/S CNN). The output of the classifier is one of the three classes, which are normal, benign, and malignant. From a numerical point of view, a mathematical function is required to perform the prediction process. In this context, the softmax function is used for multi-classification. Generally, it is often used in neural networks to map the non-normalized output to a probability distribution over predicted output classes [49].

The softmax function normalizes the outputs of each neuron such that each numerical value will be between 0 and 1. Due to the normalization process, each output is divided into parts that meet the following criteria (the total sum of the outputs is equal to 1). Formally, the softmax function is given by:

$$\partial(VoI)_\tau = \frac{e^{VoI_\tau}}{\sum_{\sigma=1}^{\sigma} e^{VoI_\sigma}} \qquad (10)$$

where VoI defines a vector of the inputs to the output layer. If there are 10 output units, then there are 10 elements in VoI. τ indexes the output units, so that $\tau = 1.2\ldots.\sigma$ [48].

Fig. 15 visually illustrates the three main output classes in this work.

It is obvious that, in Fig. 15, if the ∂ value is 0.11, this means that the input mammogram image has a malignant mass. If the ∂ value is 0.85, the input mammogram image has a benign mass. If the ∂ value is 0.04, this means that the input mammogram image is normal.
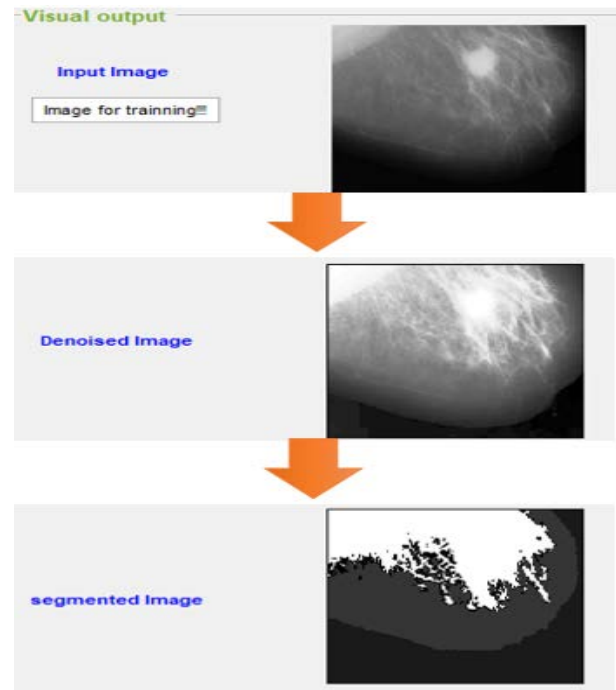

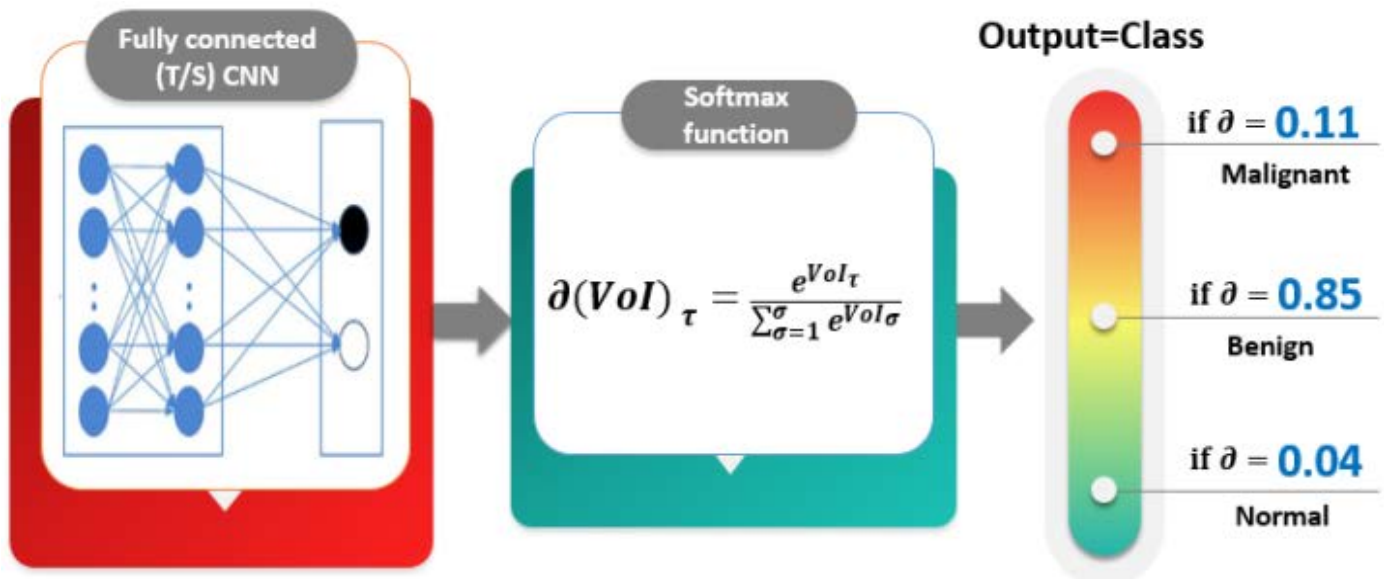Fig. 14. Output of the Segmentation Step (Applying the GMM Filter).

Fig. 15. Output Classes of the Prediction Process in the T/S CNN from a Numerical Perspective.

IV. USED METRICS

Three types of metrics are presented for use in the evaluation process. They are AI-based metrics, performance-based metrics, and quality of image metrics.

A. AI-based Metrics

In general, the confusion matrix (CoMa) is an effective benchmark for analyzing how well a classifier can recognize the images of different classes. The CoMa is formed considering the following terms [50]:

1) True positives (TP): positive images that are correctly labelled by the classifier.

2) True negatives (TN): negative images that are correctly labelled by the classifier.

3) False positives (FP): negative images that are incorrectly labelled as positive.

4) False negatives (FN): positive images that are mislabeled as negative.

Table III shows the CoMa in terms of the TP, FN, FP, and TN.

Relying on the CoMa, the accuracy (Acc), sensitivity (Sen), specificity (Spe), precision (Pre), and recall (Rec) metrics are driven. For a given classifier, the accuracy can be calculated by considering the recognition rate, which is the percentage of the test set images that are correctly classified. The accuracy is defined as:

$$Acc = \frac{(TP+TN)}{number\ of\ all\ records} \quad (11)$$

Accuracy-based evaluation. In this context, a higher accuracy corresponds to a better classifier output. The maximum value of the accuracy metric is 1 (or 100%), which is achieved when the classifier classifies the images correctly without any error in the classification process.

Sensitivity refers to the true positive recognition rate. It is given by:

$$Sen = \frac{TP}{P} \quad (12)$$

Sensitivity-based evaluation. In this context, a higher sensitivity corresponds to a better classifier output. The maximum value of the sensitivity metric is 1 (or 100%), which is achieved when the proportion of true positive cases equals the number of P cases.

For specificity, it refers to the true negative recognition rate. It is given by:

$$Spe = \frac{TN}{N} \quad (13)$$

Specificity-based evaluation. In this context, a higher specificity corresponds to a better classifier output. The maximum value of the specificity metric is 1 (or 100%), which is achieved when the proportion of true negative cases equals the number of N cases.

For precision, it refers to the exactness (what % of tuples that the classifier labelled as positive that are actually positive). It is given by:

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

TABLE III.   CONFUSION MATRIX

Actual class (Predicted class)	Confusion matrix		
	C1	*¬ C1*	*total*
C1	True Positives (TP)	False Negatives (FN)	TP + FN = P
¬ C1	False Positives (FP)	True Negatives (TN)	FP + TN = N

Precision-based evaluation. In this context, a higher precision corresponds to a better classifier output. The maximum value of the precision metric is 1 (or 100%), which is achieved when FP=0.

Recall refers to the completeness (what % of positive tuples did the classifier label as positive?). It is given by:

$$Rec = \frac{TP}{TP+FN} \tag{15}$$

Recall-based evaluation. In this context, a higher recall corresponds to a better classifier output. The maximum value of the recall metric is 1 (or 100%), which is achieved when FN=0.

ROC curves are used to enable the visual comparison of different classification models. These curves indicate the balance between the true and false positive rates, and the area under the ROC curve denotes the accuracy of the classifier [47].

ROC-based evaluation. In this context, a model representing a line closer to the diagonal line (i.e. the closer the area is to 0.5) is a less accurate model.

B. Performance-based Metrics

Time of response (ToR) is used to evaluate approaches in terms of performance. The ToR is calculated based on the total time of the four main steps that are illustrated in Fig. 6 above (i.e. pre-processing, feature extraction, segmentation, and classification steps). ToR is given as:

$$ToR = T_{pre} + T_{fext} + T_{seg} + T_{clas} \tag{16}$$

Where $T_{pre}.T_{fext}.T_{seg}.and\ T_{clas}$ denote the time consumed by the pre-processing, feature extraction, segmentation, and classification steps, respectively.

It is worth mentioning that, the shorter the ToR is, the better the performance of the approach is.

C. Quality of Image Metrics

Here, we employ the mean square error (*MSE*) and peak signal-to-noise ratio (*PSNR*) metrics. These metrics provides numerical values to measure the amount of distortion. The PSNR measures the percentage of the signal to the noise. If its value is high, the quality of the image is good. It is given by:

$$PSNR = 20\ log_{10}\ (\frac{max_i}{\sqrt{MSE}}) \tag{17}$$

Where the MSE is given by:

$$MSE = \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m}\left\|img_{prep}^{after} - img_{prep}^{before}\right\|^2 \tag{18}$$

Where img_{prep}^{after} and img_{prep}^{before} denote the mammogram images after the pre-processing step and before the pre-processing step, respectively. It is worth mentioning that there is an inverse mathematical relationship between the MSE and PSNR metrics.

V. Experimental Results and Evaluations

This section is arranged so that first it first presents the setup and the approaches that are intended to be compared. Then, the actual results and the corresponding discussions are provided.

A. Setup

The proposed AI-based system is implemented on a laptop that has the specifications organized in Table IV.

We selected two approaches to compare with the proposed approaches. They are arranged in Table V.

Since refs [7] and [11] are based on RF and KNN, we present a brief overview of these techniques, as described below.

The RF algorithm, or forest of decision trees, is a classification algorithm that reduces the variance of forecasts from a decision tree alone, thereby improving their performance. For this, it combines many decision trees in a bagging-type approach.

The RF algorithm performs parallel learning on multiple decision trees constructed randomly and trained on different subsets of data. The ideal number of trees, which can go up to several hundred or more, is an important parameter; it is very variable and depends on the problem. Concretely, each tree in the random forest is trained on a random subset of data according to the bagging principle, with a random subset of features according to the principle of random projections. Predictions are then averaged when the data are quantitative or used for a vote on qualitative data, in the case of classification trees. The random forest algorithm is known to be one of the most efficient classifiers. Fig. 16 shows the basic concept of the RF technique.

The KNN algorithm is an unsupervised machine learning algorithm that can be used to solve both classification and regression problems. The KNN algorithm does not require a learning phase; it is just necessary to store the set of learning data.

TABLE IV. SPECIFICATIONS

Item	Details (value)
Operating system	Microsoft Windows 10 Home
System type	x64-based PC
System model	HP Laptop 15-bs0xx
Processor	Intel(R) Core(TM) i5-7200U CPU @ 2.50 GHz
RAM	4 GB
Display chip type	Intel(R) HD Graphics Family

TABLE V. SELECTED APPROACHES

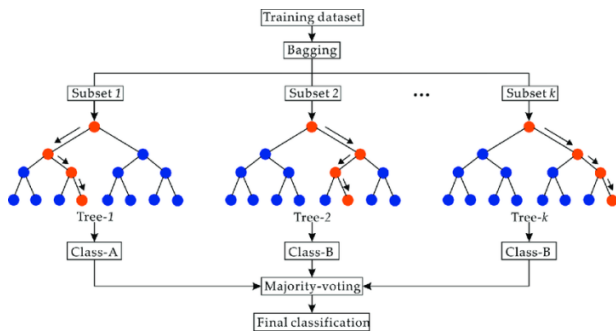| Ref	Selected Approaches		
	Used technique	*Journal*	*Year*
[7]	RF-based	Journal of medical systems	2016
[11]	KNN-based	IEEE	2013

Fig. 16. RF Technique Concept.

Let E be a set containing n labeled data: $E = \{(y_i . x_i)\}$ with $1 < i < n$, where y_i corresponds to the class (label) of the data i and where the vector x_i of dimension p $(x_i = (x_{i1} . x_{i2} . \dots . x_{ip}))$ represents the predictor variables of the data i.

Consider a data u that does not belong to E and that does not have a label (u is only characterized by a vector x_u of dimension p).

Let d be a function that returns the distance between the data um and any data um belonging to E. Let an integer k be less than or equal to n. Below are the general steps of the KNN algorithm:

1) Calculate the distances between the data u and each data belonging to E using the function d.

2) Retain the k data from the data set E closest to u.

3) Attribute to u the class that is most frequent among the k closest data.

Fig. 17 illustrates the concept of the KNN technique.

B. Results and Discussion

1) Results without the effects of shift and scaling problems: In the context of these experiments, 200 mammogram images are used. Since the AI-based metrics are inspired from the CoMa, the values of the cells that form the CoMa are needed. Table VI shows the values of the CoMa elements.

Based on the metrics that are derived from the AI, Table VII presents the values of the metrics along with the comparison with the RF-based and KNN-based approaches.

AI-based metric discussion. The proposed SCNN approach achieves the best values. The reason is that the mammogram images are scanned perfectly by the kernels/filters to extract features and then form the entire image feature map. The pre-processing step that includes HisE followed by AMF contributes to enhancement of the values of all of the metrics. That is because the contrast and the details of the input images are enhanced, which in turn leads to better training (i.e. training on very clear images). As a result, the accuracy level is high. This justification can be used to explain all of the high values of the rest of the metrics. The TCNN is ranked second. The reason is that it uses only the HisE method in the pre-processing step. This leads to lowering the values of the metrics slightly. Compared to the SCNN and TCNN

approaches, the RF performs poorer, but it performs better when compared to the KNN approach. This can be justified by the nature of the RF approach, which takes into consideration all of the possible options that may be generated in the decision trees while processing the images to make the final decision (i.e. classifying the mammogram input image).

Fig. 18 illustrates the ROC-based comparison among the four approaches.
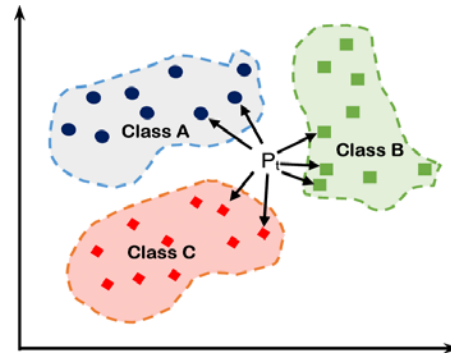
Fig. 17. KNN Technique Concept.

TABLE VI. VALUES OF CoMa ELEMENTS

200 images 131 (P) 69 (N)	Values Of CoMa Elements			
	TP	TN	FP	FN
RF	102	54	29	15
KNN	99	50	32	19
TCNN	119	65	12	4
SCNN	124	66	7	3

TABLE VII. AI-BASED METRIC VALUES

Approach	Metric			
	Acc	Sen	Spe	Pre
KNN	74%	84%	61%	76%
RF	78%	87%	65%	78%
TCNN	92%	97%	84%	91%
SCNN	95%	98%	90%	95%

Fig. 18. ROC Curves.

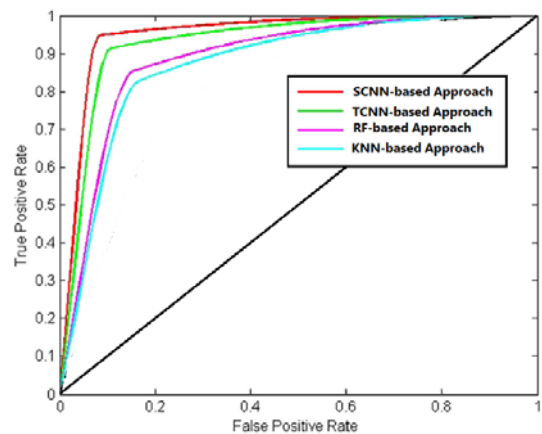ROC-based discussion. The results obtained from the AI-based metrics are supported by the ROC curves. This is reflected through the area under the ROC, where the SCNN approach has the highest area. The TCNN approach comes in next due to the same reasons presented under the discussion of the AI-based metrics. Again, the RF approach performs better than the KNN approach.

Performance-based metric discussion. Fig. 19 illustrates the comparison according to the response time. In terms of time, the SCNN approach comes in at the bottom (i.e. performs the worst), followed by the TCNN approach. The KNN approach comes in at the top, and the RF approach is ranked second. The reason behind the previous order is related to the intensity of the operations used in each step involved in the prediction process (i.e. from the pre-processing step to the classification step). Because the CNN is included in both the proposed TCNN and SCNN approaches, the operations that are involved (constructing the convolution and pooling layers), as well as using extra methods for pre-processing, consume more time. This in turn leads to a long ToR metric value. As for the KNN approach, it directly manipulates the input mammogram images and provides the prediction without any additional delays, which leads to the shortest ToR metric value.

Quality of image-based discussion. From an image processing point of view, the greater the clarity of the image is, the more accurate the output is. This plays a critical role in the medical field due to the sensitivity of diagnosis based on image processing. Since we used HisE and AMF in the pre-processing step, we examine the four options that are generated from them. Table VIII summarizes the values of the MSE and PSNR metrics.

Applying HisE followed by AMF yields the lowest MSE and the highest PSNR, which means that this option is the best for generating the clearest mammogram images (actually, this option is the one used in this work). The reason is that the HisE enhances the details of the mammogram input image. These details undergo another optimization related to surrounding the details within edges, including the area of the cancerous mass. In other words, double enhancement is applied on the mammogram images, leading to a good isolation of the cancerous mass, and this is the root reason behind the higher accuracy of classification as mentioned previously. In contrast, when applying the AMF followed by the HisE, the MSE yields a higher error rate, and the PSNR yields lower clarity of the output processed image. That is because identifying the location of cancerous mass within the mammogram image first leads to the inclusion of details that may be out of interest. Thus, when applying the HisE on the image, the details will affect the accuracy of the classification process because the unwanted details are seen as noise. It is obvious that applying only AMF or HisE generates poor-quality images compared to the other options.

2) Results taking into consideration the effect of shift and scaling problems: In the context of this part of the experiments, we manually add some noise to the mammogram breast images to generate the shift and scaling problems. Then, robustness against the shift and scaling problems is measured based on repeating the same scenario (calculating the values of the AI-based metrics, performance metrics, and quality of image metrics) presented in the first part of the experiments.

Fig. 20 illustrates an output image, as a sample, after adding shift and scaling problems.

The values of the elements of the CoMa are arranged in Table IX.

Table X shows the values of the AI-based metrics.
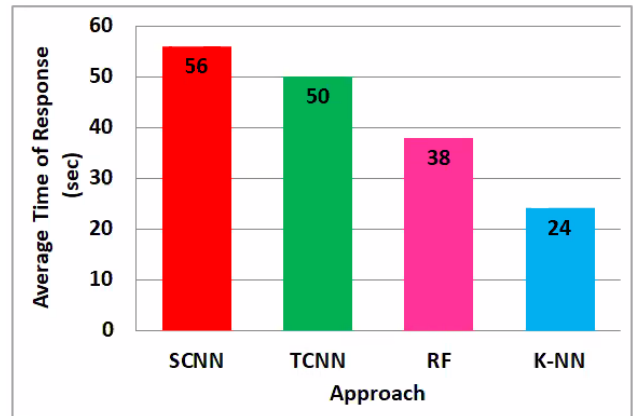
Fig. 19. Performances of Approaches.

TABLE VIII. VALUES OF MSE AND PSNR METRICS

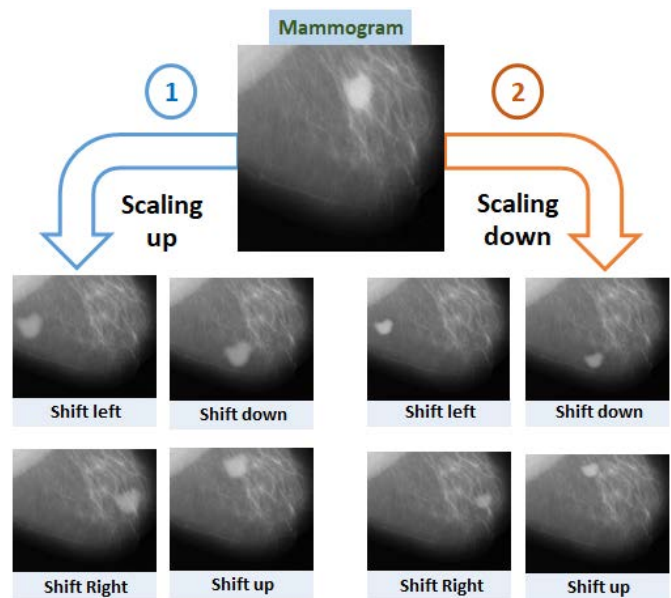| Filter	Metric	
	MSE	*PSNR*
HisE	28.96769	13.493
AMF	32.6703	13.1818
HisE+AMF	12.80	25.98
AMF+HisE	15.72638	20.2516

Fig. 20. Output Mammogram Image after Adding Shift and Scaling Problems.

TABLE IX.   VALUES OF THE CoMa's ELEMENTS AFTER APPLYING SHIFT
AND SCALING PROBLEMS

200 images 131 (P) 69 (N)	Values of CoMa Elements			
	TP	*TN*	*FP*	*FN*
RF	91	49	40	29
KNN	85	45	46	24
TCNN	116	61	15	8
SCNN	122	63	9	6

TABLE X.   VALUES OF AI-BASED METRICS AFTER APPLYING SHIFT AND
SCALING PROBLEMS

Approach	Metric			
	Acc	*Sen*	*Spe*	*Pre*
KNN	65%	78%	49%	65%
RF	70%	76%	55%	69%
TCNN	89%	94%	80%	89%
SCNN	93%	95%	88%	93%

AI-based metrics discussion. In general, the shift and scaling problems have a negative impact on all of the AI-based metrics for all of the approaches. However, the SCNN approach shows the highest resistance against these problems. That is because of the support it offers via a strong pre-processing step, which has the ability to remove noise. On the other hand, the padding operation used to form the convolution layers and the max function used to construct the pooling layers contribute to limiting the negative impact of the shift and scaling problems. In numbers, the accuracy of the SCNN approach decreases from 95% to 93% after applying the shift and scaling problems. The TCNN approach comes in next in terms of resistance against the problems, with its accuracy decreasing from 92% to 89%. Here, the use of only HisE to remove the noise is the reason the accuracy of the TCNN is lower than the accuracy of the SCNN. For the KNN and RF approaches, there is a significant decrease in the accuracy value, from 74% to 65% and from 78% to 70%, respectively.

Fig. 21 shows the performance of the four approaches involved in the comparison.
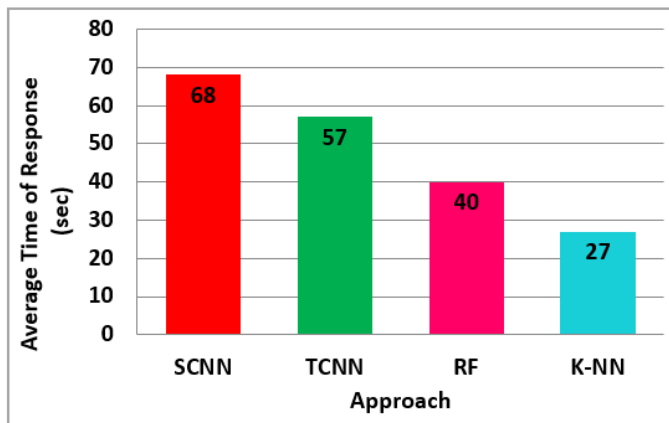

Fig. 21. Performances of the Approaches after Applying the Shift and Scaling Problems.

Performance-based metric discussion. In general, the value of the ToR metric shows poor performance due to the additional processing to remove the noise that is added by the shift and scaling problems. However, the order of the approaches in terms of performance did not change after applying the shift and scaling problems (when compared to the first scenario). It is noticed that the TCNN and SCNN approaches experienced higher rates of increase compared to the KNN and RF approaches. This is normal because the pre-processing step needs more time to generate clear output images.

Table XI summarizes the results of the quality of image metrics after applying shift and scaling problems.

TABLE XI.   VALUES OF QUALITY OF IMAGE METRICS UNDER IMPACT OF
THE SHIFT AND SCALING PROBLEMS

Filter	Metric	
	MSE	*PSNR*
HisE	30.75	12.44
AMF	35.82	12.24
HisE+AMF	14.79	22.30
AMF+HisE	18.54	17.23

Quality of image-based discussion. In general, the quality of the mammogram input images is negatively affected by the shift and scaling problems. This is reflected in the value of the MSE metric (it shows increasing values) and the PSNR metric (it shows decreasing values). However, the approach that depends on the HisE method followed by the AMF method has the best values. That is because of the effectiveness of employing the histogram equalization and the adaptive filter to remove noise generated by the shift and scaling problems. As for the rest of the approaches in Table XI, they suffer from the two problems, which means a lower resistance against the noise. In terms of ranking, the approach that relies on the AMF method followed by the HisE method comes in second, followed by the HisE and AMF approach.

VI. CONCLUSION

Diagnosis in the medical field plays an important role in saving the lives of patients. This is highlighted specially in the cases where early detection of illness is required, such as breast cancer. Employing machine learning to perform diagnosis on behalf of physicians makes the accuracy of diagnosis a critical issue. That is because low accuracy in diagnosis (i.e. mistakes) sometimes leads to death. The accuracy level is also tightly coupled with blurring of the processed images in the medical field. This leads to the following research questions: (1) How can the accuracy of the artificial systems be enhanced? (2) How can the noise that is generated by the shift problem included in the images be manipulated before being processed by machines before training? (3) How can the noise generated by the scaling problem of the images be addressed before entering the training phase? This work presents the TCNN and SCNN approaches, which respond to the blurring issue. The TCNN and SCNN exploit the mechanism of the neural networks used to construct the convolution layers to deal with the shift problem. Specifically, the padding operation is

exploited during the scanning of the input mammogram images by filters. To manipulate the scaling problem, the TCNN and SCNN also exploit the mechanism of the neural networks used to construct the pooling layers. Specifically, this is done by depending on the max function to generate the feature map of the mammogram images. As for enhancing the accuracy level, the FRbA is proposed. The key idea is to generate different instances of the cancerous mass from the input mammogram images to generate more features. Specifically, this is done by relying on both the flipping and rotation mathematical operations that can be applied on the mammogram images. The proposed CNN and SCNN approaches are trained and tested using the MIAS dataset available online, where 200 mammogram breast images are utilized. The proposed CNN and SCNN approaches are evaluated under different confusion matrix-based matrices and compared with similar approaches. The SCNN and TCNN achieved accuracies of 95% and 92%, respectively, while the RF-based and KNN-based approaches exhibited accuracies of 78% and 74%. In terms of applicability, the results of the accuracy related to both the TCNN and the SCNN make them promising approaches to be used in medical sector.

For limitations, due to the high level of both the privacy and the security in medical sector, the implementation of the proposed approaches did not perform on real medical data. In addition, response time did not take into consideration in this work.

In future work, we intend to implement the proposed approaches with a real medical dataset. In addition, enhancing performance in terms of response time will be taken into consideration by implementing the proposed approaches on Hadoop.

REFERENCES

[1] Alhashmi, Shaikha FS, et al. "A Systematic Review of the Factors Affecting the Artificial Intelligence Implementation in the Health Care Sector." Joint European-US Workshop on Applications of Invariance in Computer Vision. Springer, Cham, 2020.

[2] Vuong, Quan-Hoang, et al. "Artificial intelligence vs. natural stupidity: Evaluating AI readiness for the vietnamese medical information system." Journal of clinical medicine 8.2 (2019): 168.

[3] Garbuio, Massimo, and Nidthida Lin. "Artificial intelligence as a growth engine for health care startups: Emerging business models." California Management Review 61.2 (2019): 59-83.

[4] Gruson, Damien, et al. "Data science, artificial intelligence, and machine learning: Opportunities for laboratory medicine and the value of positive regulation." Clinical biochemistry (2019).

[5] Byteant website. Available [online]: https://www.byteant.com/blog/ai-adoption-in-healthcare-10-pros-and-cons/, (Accessed on June 6, 2020).

[6] Nahid, Abdullah-Al, and Yinan Kong. "Involvement of machine learning for breast cancer image classification: a survey." Computational and mathematical methods in medicine 2017 (2017).

[7] Diz, Joana, Goreti Marreiros, and Alberto Freitas. "Applying data mining techniques to improve breast cancer diagnosis." Journal of medical systems 40.9 (2016): 203.

[8] Abdel-Nasser, Mohamed, et al. "Breast tumor classification in ultrasound images using texture analysis and super-resolution methods." Engineering Applications of Artificial Intelligence 59 (2017): 84-92.

[9] García-Manso, Antonio, et al. "Consistent performance measurement of a system to detect masses in mammograms based on blind feature extraction." Biomedical engineering online 12.1 (2013): 2.

[10] Ting, Fung Fung, Yen Jun Tan, and Kok Swee Sim. "Convolutional neural network improvement for breast cancer classification." Expert Systems with Applications 120 (2019): 103-115.

[11] Radovic, Milos, et al. "Application of data mining algorithms for mammogram classification." 13th IEEE International Conference on BioInformatics and BioEngineering. IEEE, 2013.

[12] Singh, Vivek Kumar, et al. "Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network." Expert Systems with Applications 139 (2020): 112855.

[13] Hosaini, Sayedeh Somayeh, and Mehran Emadi. "Breast Cancer Tumor Diagnosis from Mammography Images Using Wavelet Transform and Hidden Markov Model." International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (2015).

[14] Al-antari, Mugahed A., et al. "An automatic computer-aided diagnosis system for breast cancer in digital mammograms via deep belief network." Journal of Medical and Biological Engineering 38.3 (2018): 443-456.

[15] Vivona, Letizia, et al. "Fuzzy technique for microcalcifications clustering in digital mammograms." BMC medical imaging 14.1 (2014): 23.

[16] Dhahbi, Sami, Walid Barhoumi, and Ezzeddine Zagrouba. "Breast cancer diagnosis in digitized mammograms using curvelet moments." Computers in biology and medicine 64 (2015): 79-90.

[17] Li, Jun-Bao. "Mammographic image based breast tissue classification with kernel self-optimized fisher discriminant for breast cancer diagnosis." Journal of medical systems 36.4 (2012): 2235-2244.

[18] Ganesan, Karthikeyan, et al. "Decision support system for breast cancer detection using mammograms." Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine 227.7 (2013): 721-732.

[19] Aminikhanghahi, Samaneh, et al. "A new fuzzy Gaussian mixture model (FGMM) based algorithm for mammography tumor image classification." Multimedia Tools and Applications 76.7 (2017): 10191-10205.

[20] Zemmal, Nawel, Nabiha Azizi, and Mokhtar Sellami. "CAD system for classification of mammographic abnormalities using transductive semi supervised learning algorithm and heterogeneous features." 2015 12th International Symposium on Programming and Systems (ISPS). IEEE, 2015.

[21] Rajakeerthana, K. T., C. Velayutham, and K. Thangavel. "Mammogram image classification using rough neural network." Computational Intelligence, Cyber Security and Computational Models. Springer, New Delhi, 2014. 133-138.

[22] Fondón, Irene, et al. "Automatic classification of tissue malignancy for breast carcinoma diagnosis." Computers in biology and medicine 96 (2018): 41-51.

[23] Yu, Kun-Hsing, Andrew L. Beam, and Isaac S. Kohane. "Artificial intelligence in healthcare." Nature biomedical engineering 2.10 (2018): 719-731.

[24] Ma, W. K., et al. "Extra patient movement during mammographic imaging: an experimental study." The British journal of radiology 87.1044 (2014): 20140241.

[25] Ma, Wang KEi, et al. "Analysis of motion during the breast clamping phase of mammography." The British journal of radiology 89.1059 (2016): 20150715.

[26] Ma, W. K., et al. "What is the minimum amount of simulated breast movement required for visual detection of blurring? An exploratory investigation." The British journal of radiology 88.1052 (2015): 20150126.

[27] Ma, Wang Kei, et al. "A method to investigate image blurring due to mammography machine compression paddle movement." Radiography 21.1 (2015): 36-41.

[28] Seddon, D., K. A. Schofield, and C. A. Waite. "Investigation into possible causes of blurring in mammograms." Breast Cancer Research. Vol. 2. No. 2. BioMed Central, 2000.

[29] Ganor, Boaz. "Artificial or Human: A New Era of Counterterrorism Intelligence?." Studies in Conflict & Terrorism (2019): 1-20.

[30] Nixon, Mark, and Alberto Aguado. Feature extraction and image processing for computer vision. Academic Press, 2019.

[31] Wang, Fei, et al. "Daily pattern prediction based classification modeling approach for day-ahead electricity price forecasting." International Journal of Electrical Power & Energy Systems 105 (2019): 529-540.

[32] Definitions website. [online], available: https://www.definitions.net/definition/health+care+industry . Accesses May25,2020.

[33] Hu, Zilong, et al. "Deep learning for image-based cancer detection and diagnosis− a survey." Pattern Recognition 83 (2018): 134-149.

[34] Yassin, Nisreen IR, et al. "Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review." Computer methods and programs in biomedicine 156 (2018): 25-45.

[35] Cheng, Heng-Da, et al. "Automated breast cancer detection and classification using ultrasound images: A survey." Pattern recognition 43.1 (2010): 299-317.

[36] Alrahhal, Mohamad Shady, Maher Khemakhem, and Kamal Jambi. "Agent-Based System for Efficient kNN Query Processing with Comprehensive Privacy Protection." International Journal Of Advanced Computer Science And ApplicationS 9.1 (2018): 52-66.

[37] Alrahhal, Mohamad Shady, et al. "AES-route server model for location based services in road networks." International Journal Of Advanced Computer Science And Applications 8.8 (2017): 361-368.

[38] Alrahhal, Mohamad Shady, Maher Khemakhem, and Kamal Jambi. "A Survey on Privacy of Location-Based Services: Classification, Inference Attacks, and Challenges." Journal of Theoretical & Applied Information Technology 95.24 (2017).

[39] Alrahhal, Mohamad Shady, Maher Khemekhem, and Kamal Jambi. "Achieving load balancing between privacy protection level and power consumption in location based services." (2018).

[40] Alrahhal, H.; Alrahhal, M.S.; Jamous, R.; Jambi, K. A Symbiotic Relationship Based Leader Approach for Privacy Protection in Location Based Services. ISPRS Int. J. Geo-Inf. 2020, 9, 408.

[41] Al-Rahal, M. Shady, Adnan Abi Sen, and Abdullah Ahmad Basuhil. "High level security based steganoraphy in image and audio files." Journal of theoretical and applied information technology 87.1 (2016): 29.

[42] Alluhaybi, Bandar, et al. "A Survey: Agent-based Software Technology Under the Eyes of Cyber Security, Security Controls, Attacks and Challenges." International Journal of Advanced Computer Science and Applications (IJACSA) 10.8 (2019).

[43] Fouz, Fadi, et al. "Optimizing Communication And Cooling Costs In Hpc Data Center." Journal of Theoretical and Applied Information Technology 85.2 (2016): 112.

[44] Suckling, J., et al. "Mammographic Image Analysis Society (MIAS) database v1. 21 [Dataset], 2015." (1810).

[45] Rao, Boyina Subrahmanyeswara. "Dynamic Histogram Equalization for contrast enhancement for digital images." Applied Soft Computing 89 (2020): 106114.

[46] Saleem, S. Abdul, and T. Abdul Razak. "An effective noise adaptive median filter for removing high density impulse noises in color images." International Journal of Electrical and Computer Engineering 6.2 (2016): 611.

[47] Towards data science website. Online, available: https://towards datascience.com/feature-extraction-techniques-d619b56e31be. (Accessed on, June 3, 2020).

[48] Reynolds, Douglas A. "Gaussian Mixture Models." Encyclopedia of biometrics 741 (2009).

[49] Zeng, Rui, et al. "Quaternion softmax classifier." Electronics letters 50.25 (2014): 1929-1931.

[50] Alrahhal, Mohamad Shady, and Adnan Abi Sen. "Data mining, big data, and artificial intelligence: An overview, challenges, and research questions." (2018).