Article

Automatic Breast Tumor Screening of Mammographic Images with Optimal Convolutional Neural Network

Pi-Yun Chen [1], Xuan-Hao Zhang [1], Jian-Xing Wu [1], Ching-Chou Pai [1,2], Jin-Chyr Hsu [3], Chia-Hung Lin [1,*] and Neng-Sheng Pai [1,*]

1   Department of Electrical Engineering, National Chin-Yi University of Technology,
    Taichung City 41170, Taiwan; chenby@ncut.edu.tw (P.-Y.C.); jefi889c1ckqw2012@gmail.com (X.-H.Z.);
    jian0218@gmail.com (J.-X.W.); actirin1945@gmail.com (C.-C.P.)
2   Division of Cardiovascular Surgery, Show-Chwan Memorial Hospital, Changhua 500, Taiwan
3   The Consultant Physician of Taichung Hospital of the Ministry of Health and Welfare,
    Taichung City 403, Taiwan; jinchyr.hsu@msa.hinet.net
*   Correspondence: eechl53@gmail.com (C.-H.L.); pai@ncut.edu.tw (N.-S.P.)

Abstract: Mammography is a first-line imaging examination approach used for early breast tumor screening. Computational techniques based on deep-learning methods, such as convolutional neural network (CNN), are routinely used as classifiers for rapid automatic breast tumor screening in mammography examination. Classifying multiple feature maps on two-dimensional (2D) digital images, a multilayer CNN has multiple convolutional-pooling layers and fully connected networks, which can increase the screening accuracy and reduce the error rate. However, this multilayer architecture presents some limitations, such as high computational complexity, large-scale training dataset requirements, and poor suitability for real-time clinical applications. Hence, this study designs an optimal multilayer architecture for a CNN-based classifier for automatic breast tumor screening, consisting of three convolutional layers, two pooling layers, a flattening layer, and a classification layer. In the first convolutional layer, the proposed classifier performs the fractional-order convolutional process to enhance the image and remove unwanted noise for obtaining the desired object's edges; in the second and third convolutional-pooling layers, two kernel convolutional and pooling operations are used to ensure the continuous enhancement and sharpening of the feature patterns for further extracting of the desired features at different scales and different levels. Moreover, there is a reduction of the dimensions of the feature patterns. In the classification layer, a multilayer network with an adaptive moment estimation algorithm is used to refine a classifier's network parameters for mammography classification by separating tumor-free feature patterns from tumor feature patterns. Images can be selected from a curated breast imaging subset of a digital database for screening mammography (CBIS-DDSM), and K-fold cross-validations are performed. The experimental results indicate promising performance for automatic breast tumor screening in terms of recall (%), precision (%), accuracy (%), F1 score, and Youden's index.

Keywords: convolutional neural network (CNN); fractional-order cconvolutional operation; adaptive moment estimation algorithm

1. Introduction

As per statistics provided in 2020 by Taiwan's Ministry of Health and Welfare, cancer (malignant tumors) is the primary cause of death among Taiwanese people. In recent years, breast cancer (BC) in females is among the top four cancers (first place) and is one of the diseases that most definitely cannot be ignored. The age at which women possibly develop BC is between 45 and 69 years. As per latest figures on the cause of death from the Ministry of Health and Welfare and cancer registration data from the National Health Agency [1], the standardized incidence and mortality rates of female BC are 69.1 and 12.0 (per 100,000 people), respectively. Each year, more than

10,000 women are diagnosed with BC and more than 2000 women die from it. This is about 31 women being diagnosed with BC every day and six women losing their precious lives because of BC. BC is a malignant tumor that grows from epithelial cells or lobules of the breast, thus resulting in excessive tissue accumulation or hyperplasia. Ductal carcinoma and lobular carcinoma are common types of BC. When proliferating cells mutate or lose control, they can invade or destroy other adjacent tissues and organs, transfer to other organs via the blood or the lymphatic system, and cause breast pain. Symptoms include irregular lumps (partial) on breasts, sunken skin, orange peel skin, redness, ulceration, abnormal secretions, venous vasodilation, and swollen lymph nodes under the arms. For clinical testing, visual inspection or palpation helps identify the existence and location of the tumor (hard mass). However, it is not immediately clear whether the tumor is benign or malignant or whether metastasis is happening. Either additional diagnostic tools or instruments must be used for verification. As per statistics, there is improved chance of tumors on the right side of the breast. The most common location is the upper right corner of the breast, accounting for about 26% of all occurrences. As per the size of the tumor and severity of lymph node (sentinel lymph node) metastasis, BC can be divided into four stages. If the abnormality is quickly detected, the survival rate is higher, and the treatment is more effective. Hence, the early symptoms of BC are traceable, whereas the survival rate of BC, if detected early, can be larger than 90%.

Accordingly, in recent years, artificial intelligence (AI) methods and the collection of data for big data (BD) are playing increasingly important roles for automatic tumor screening, such as liver, lung, and breast cancers. AI methods include models based on machine learning, deep learning, and broad learning [2–9]. For example, in Taiwan, data collection for BD includes medical images of relevant major diseases that major medical centers have been actively collecting both locally and abroad in recent years (since 2018). These medical images comprise fifteen categories, including X-ray images, angiography, magnetic resonance imaging (MRI), and computer tomography (CT). Additionally, a database such as CBIS-DDSM (Curated Breast Imaging Subset of Digital Database for Screening Mammography), is a database containing approximately 2500 enrolled subjects, employed for the studies of the mammographic classification of breast lesions. This classification consists of normal cases, benign cases, malignant cases, and pathology information [10], which provides ground truth validation that makes the DDSM applicable in the development and validation of the decision support systems. The region of interest (ROI) and pathological information can be gathered continuously in the current database. Coupled with the improvement of information communication, digital data processing, AI methods, and machine vision models, software and hardware equipment can process the large amount of digital data [11–13]. However, a decision support system with computer-aided diagnosis and detection algorithms for breast tumor screening in mammography are needed. Currently, regarding AI algorithms, the multilayer machine vision recognition model that is composed of convolutional neural networks (CNNs) is the most commonly used model. It can be used for digital image processing, such as feature enhancement, feature extraction, data simplification, and pattern recognition task [2–9,13–15]. Hence, this study uses breast mammogram images to establish a two-dimensional (2D) fractional order-based CNN classifier that can automatically perform breast tumor screening tasks in clinical applications. We expect that a set of automatic screening tools suitable for clinical usage that takes in mammogram images can be developed to achieve the rapid identification of whether a tumor is malignant or benign. When any breast tumor is suspected, the results of the rapid screening can be used as a basis of reference for subsequent fine needle aspiration cytology/biopsy examinations. Hence, this assistive tool solves the problem of insufficient human resources for manual screening, which can potentially lead to additional problems in the medical diagnosis process. Solving this procedural congestion allows clinicians to focus more on follow-up medical strategies.

The 2D CNNs may comprise several convolutional-pooling layers and a fully connected network in the classification layer, such as back-propagation neural networks and Bayesian networks, which combine the image enhancement, feature extraction, and classification tasks to form an individual scheme [16,17] which achieves promising accuracy for image classification in breast tumor screening. These CNNs are usually greater than 10 convolutional-pooling layers which perform the above mentioned image preprocessing and postprocessing tasks and then increase the identification accuracy. Hence, this multilayer design may gradually replace machine-learning (ML) methods [18,19], which perform image segmentation and feature extraction as an image preprocess for mammograms and breast MRIs and then use the fixed features obtained to train a classifier. Both CNN and ML-based image segmentation [20] can learn the specific features or knowledge representations to automatically identify the boundaries of ROI and then detect the breast lesions. Traditional ML methods have fewer parameters that can easily be optimized by the gradient descent optimization or back-propagation algorithms through training with small-to-medium-sized datasets [21,22]. Through a series of convolutional and pooling processes, the multilayer CNN can enhance and extract the desired object at different scales and different levels from low-level features (extract object's edge) to high-level information (extract object's shape) for detecting nonlinear features, which can increase nonlinearity and obtain feature representation. Then, the pooling process with maximum pooling (MP) is used to reduce the sizes of feature maps for obtaining abstract features. Thus, in contrast to the traditional machine-learning method, CNN-based methods can learn to extract the feature patterns from the raw data and improve the classification accuracy significantly. However, small- or medium-sized datasets are insufficient to train a deep-learning-based CNN. For example, from the existing literature [23–26], such as AlexNet (eight-layer CNN) [25] and ZFnet [26], it can be observed that the deep-learning-based CNN requires several convolutional-pooling layers and fully connected layers for the large-scale image classification (ImageNet image database [27,28]). This CNN can learn to optimize features during the training stage, process large inputs with sparsely connected weights, adapt to different sizes of 2D images, and reduce error rates. Furthermore, this approach demonstrates greater computational efficiency compared with the traditional fully connected multilayer perceptron (MLP) networks. Despite its many advantages, however, a deep-learning-based CNN presents several drawbacks and limitations, such as the number of convolutional-pooling layers' determination, the number of convolutional windows and pooling windows, the sizes of convolutional window assignment (3×3, 5×5, 7×7, 9×9, 11×11), the high computational complexity and large-scale dataset requirement for training the CNN-based classifier, and the poor suitability for real-time applications. Additionally, multi-convolutional-pooling processes with different sizes of convolution masks will result in a very large information loss for feature extraction, and this will result in increased complexity levels. The multilayer CNN must be performed with a graphics processing unit (GPU) to speed up the training and classification tasks by making use of a large amount of training and testing data.

Therefore, to simplify the image processing and classification tasks, this study aimed to design a suitable number of convolutional-pooling layers and a classification layer that is capable of increasing the identification accuracy of image classification, to facilitate automatic breast tumor screening. As observed from Figure 1, we utilized a multilayer classifier, consisting of a fractional-order convolutional layer, two convolutional-pooling layers, a flattening layer, and a multilayer classifier in the classification layer. In the first convolutional layer, a 2D spatial convolutional process with two 3×3 fractional-order convolutional masks was used to perform the enhancement task and to remove unwanted noise from the original mammography image, to distinguish the edges and shapes of the object. In the second and third convolutional layers, sixteen 3×3 kernel convolutional windows were used to subsequently enhance and sharpen the feature patterns twice; hence, the tumor contour could easily be highlighted and distinguished for feature pattern extraction. Consequently, two MP processes were used to reduce the dimensions of the feature

patterns, which conducted network training to avoid failing in overfitting problems [29,30]. In the classification layer, a multilayer classifier with an input layer, two hidden layers, and an output layer is implemented to perform the pattern recognition task, which separates tumor-free feature patterns from tumor feature patterns. To reduce the error rates, an adaptive moment estimation method (ADAM) can compute the adaptive learning rates for updating network parameters by storing an exponentially decaying average of past squared gradients [31,32], which combines two stochastic gradient descent approaches, including adaptive gradients and root mean square propagation. Its optimization algorithm uses randomly selected training data subsets to compute the gradient, instead of using the entire dataset. The momentum term can speed up the gradient descent by converging faster. The ADAM algorithm has a simple implementation, computation efficiency, and fewer memory requirements, and is appropriate for operations with large datasets and parameters for training the multilayer CNN models. A total of 78 subjects is selected from the MIAS (Mammographic Image Analysis Society) Digital Mammogram Database (United Kingdom National Breast Screening Program) for experimental analysis. The clinical information was confirmed and agreed upon by expert radiologists for biomarkers, such as image size, image category, background tissue, class of abnormality, and severity of abnormality [33,34]; the image database included a total of 156 mammography images (including right and left images), including 94 normality cases and 62 abnormalities involving benign and malignant cases. The ROIs were extracted by a 100×100 bounding box, and then the 932 feature patterns were extracted by using the proposed convolutional-pooling processes including 564 abnormalities and 368 tumor-free patterns. By making use of cross-validation, the dataset was randomly divided into two halves: 50% of the dataset was used for training the classifier, and 50% of the dataset was used for evaluating the classifier's performance. Thus, tenfold cross-validation is used to verify the performances of the proposed multilayer deep-learning-based CNN with the proposed convolutional-pooling layers in terms of recall (%), precision (%), accuracy (%), F1 score, and Youden's index [35,36]. Therefore, the optimal architecture of multilayer CNN can be determined, and may potentially be applied to establish a classifier for automatic breast tumor screening in clinical applications.
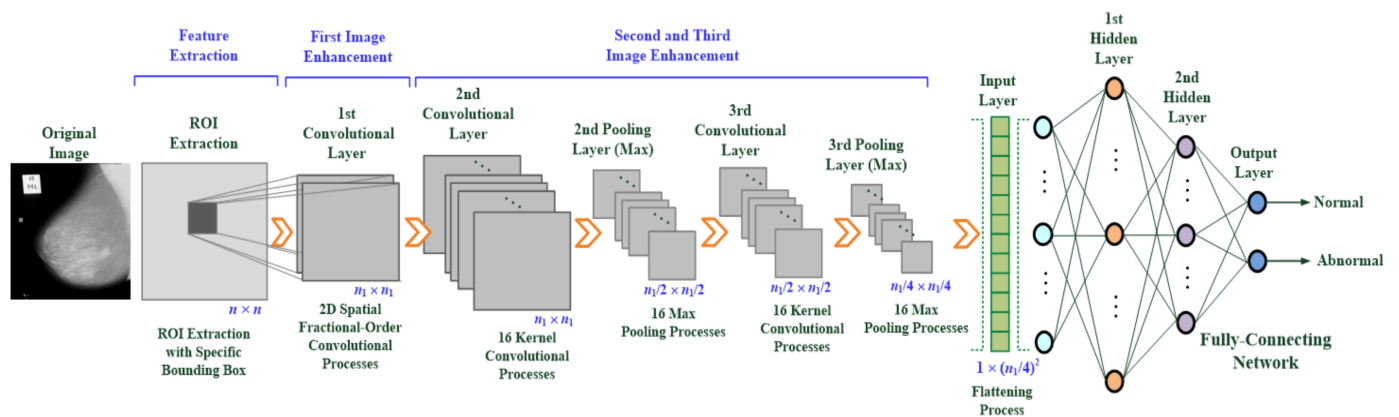

Figure 1. Architecture of the proposed multilayer CNN for automatic breast tumor screening.

The remainder of this study is organized as follows: Section 2 describes the methodology, including the design of the multilayer deep-learning-based CNN, the adaptive moment estimation method, the classifier's performance evaluations, and the computer assistive system. Sections 3 and 4 present the experimental setup, testing of different multilayer CNN models and determination of the suitable CNN architecture, testing of the first convolutional layer for image enhancement, determination of the mask types, feasibility tests, and experiment results for clinical applications, and the conclusions, respectively.

2. Materials and Methods

2.1. Design of the Multilayer Deep-Learning-Based CNN

Multilayer deep-learning-based CNN includes the multiple convolutional layers, pooling layers, and pattern recognition layers (classification layers). It is a multilayer model for image classification that combines multiple functions, such as image feature enhancement, feature extraction, parameter simplification, and pattern recognition. In recent years, deep-learning and broad-learning technologies have been applied for medical image processing, segmentation, and classification [3–9,13–15]. Compared with traditional MLP networks, CNN can process feature enhancement and extraction at the front end of the network; therein lies its advantage in processing 2D medical images. However, image processing generates a considerable number of parameters, increasing the amount of computations and the time required for computing. With pooling processes, the number of features is reduced in the middle of the network to reduce the computational time. Finally, the image classification layer assists with screening tasks using a pattern to separate the normality from abnormalities. This study uses the multilayer CNN model to develop an automatic breast tumor screening based on 2D mammogram images, as shown in Figure 1. This automatic screening process includes: (1) Region of interest (ROI) extraction; (2) feature enhancement and extraction; and (3) rapid screening of breast tumors. Each function is described as follows:

ROI extraction: In this study, we use digital data in the mammogram X-ray image database (161 female patients, 322 images) [33,34] provided by the MIAS. The biomarkers of MIAS database clearly mark the positions and tumor sizes [33]. The statistical results of probability distribution of tumor locations, in accordance with the MIAS database's biomarkers, are shown in Figure 2a. Looking at the distribution probability of tumor locations, the most frequent location of tumors is the right and the upper outer quadrant of the breast. Given a 2D image with 4320 pixels × 2600 pixels, we define the priority for automatically extracting ROI with a specific bounding box based on the distribution probability. Areas with greater probability are first in line for ROI extraction, and the priority order is stored in the work queue. The priority order is shown in Figure 2b. ROI image extraction and tumor detection is performed as per the priority order.
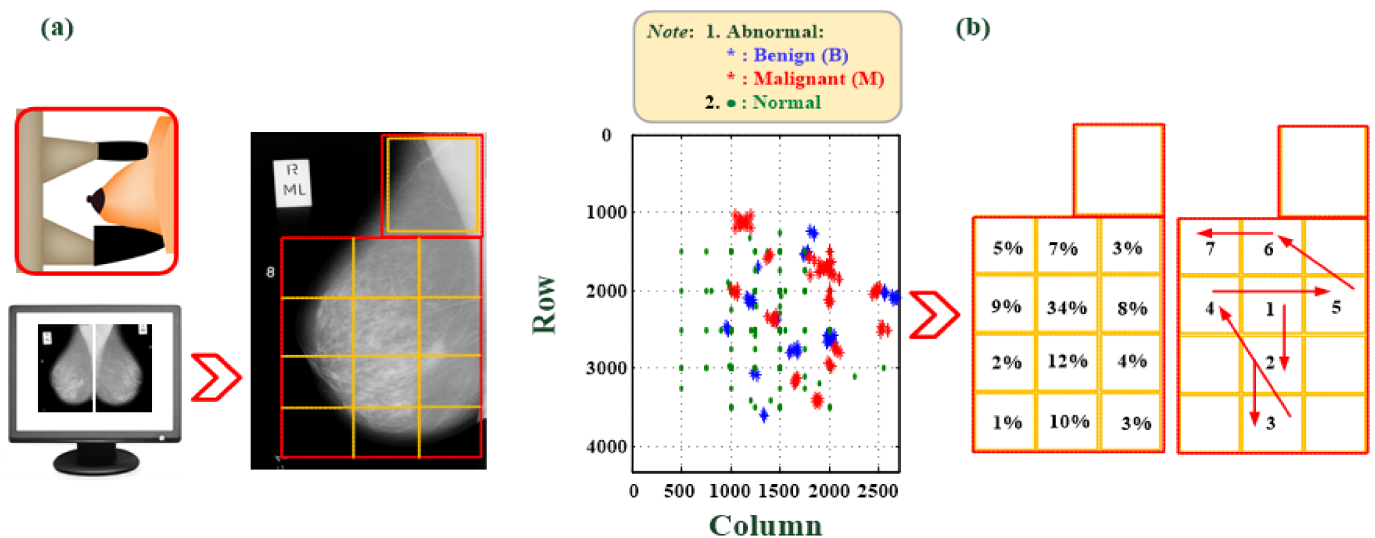
Figure 2. ROI block cutting and priority extraction. (**a**) The statistics of the prevalence of malignant and benign tumors; (**b**) priority of ROI block for automatic ROI extraction.

- Feature enhancement and extraction: A multilayer 2D convolution operation is used to magnify the texture of what might be tumor tissue and edge information (usually two or more layers are used), as shown in Figure 2a. Each layer uses a 3 × 3 sliding window to perform the operation of the convolutional weight. First, a 2D fractional

convolution operation is performed to magnify the tumor characteristics. Then, by combining multilayer convolutional weight calculations, the contour of the tumor is gradually strengthened, noise is removed, and the image is sharpened. These effects help strengthen the target area and retain non-characteristic information. This study applies the 2D spatial fractional-order convolutional processes in the fractional convolutional layer, selects the appropriate fractional order parameters, and performs convolution in the x and y directions, thus yielding a combination of 2D weight values in space, the general formula being [35–38]:

$$C^v I_{xy} = conv(I_{xy}, M(i,j), v)^T \tag{1}$$

$$C_x^v I_{xy} = \sum_{i=-\frac{h-1}{2}}^{\frac{h-1}{2}} \sum_{j=-\frac{h-1}{2}}^{\frac{h-1}{2}} M_x(i,j) I(x+i, y+j) \tag{2}$$

$$C_y^v I_{xy} = \sum_{j=-\frac{h-1}{2}}^{\frac{h-1}{2}} \sum_{i=-\frac{h-1}{2}}^{\frac{h-1}{2}} M_y(j,i) I(x+j, y+i) \tag{3}$$

where $h = 3$ is the dimension of the convolution window, v is a fractional parameter and $v \in (0, 1)$, and $I(x, y) \in [0, 255]$ is the pixel value at point (x, y) in a 2D image. Each fractional-order convolutional mask multiplies each element, $M(i, j)$ or $M(j, i)$, by the corresponding input pixel values, $I(x, y)$, and then obtains an enhanced feature pattern containing spatial features in the x-axis and y-axis directions. These 2D spatial convolutional processes act as two low-pass frequency filters [39] and then remove the high-spatial-frequency components from a breast mammogram. In this study, the image dimension is $n \times n$, $x = 1, 2, 3, \ldots, n$, and $y = 1, 2, 3, \ldots, n$. M_x and M_y are 3×3 convolutional windows that can be written as follows [35–38]:

$$M_x = \begin{bmatrix} 0 & \frac{v^2-v}{2} & 0 \\ 0 & -v & 0 \\ 0 & 1 & 0 \end{bmatrix}, \ M_y = M_x^T = \begin{bmatrix} 0 & 0 & 0 \\ \frac{v^2-v}{2} & -v & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{4}$$

where $v \in (0, 1)$ is the fractional-order parameter. A sliding stride = 1 is selected for spatial domain-based convolution operations in the horizontal and vertical directions. The results of the convolution operation of (1) and (2) are combined and normalized, and the approximate formula is written below:

$$\nabla^v I_{xy} = \begin{bmatrix} C_x^v I_{xy} \\ C_y^v I_{xy} \end{bmatrix}^T, \ \left| \nabla^v I_{xy} \right| \cong \frac{\left| C_x^v I_{xy} \right| + \left| C_y^v I_{xy} \right|}{255} \tag{5}$$

These multilayer convolutional layers are also called the perception layers of the CNN network for feature enhancement and extraction. After feature extraction, the 2×2 sliding window is used to perform maximum pooling (MP), as shown in the general formula (6):

$$MP \Big|_{\frac{n}{2} \times \frac{n}{2}} = \max_{M_{2 \times 2}} (\left| \nabla^v I_{xy} \right| |_{n \times n}) \tag{6}$$

After MP, the number of feature patterns is reduced to 25% of the total number of original feature images. This reduction in the dimensions of the feature patterns can overcome the overfitting problem for training a multilayer classifier.

- Rapid screening of breast tumors: Breast tumors can be identified at the image classification layer, which includes the flattening process (FP) and a multilayer classifier, as seen in Figure 1. The FP can convert a 2D feature matrix into a 1D feature vector, which is then fed as the input vector of the classifier for further pattern recognition.

After two MP treatments, the FP treatment may be written as shown in the general formula (7):

$$X\Big|_{1\times\left(\frac{n}{4}\right)^2} = FP(MP\Big|_{\frac{n}{4}\times\frac{n}{4}}) \tag{7}$$

where X is the 1D feature vector of the multilayer classifier used as input. In this study, the multilayer classifier includes an input layer, two hidden layers (i.e., the first and second hidden layers), and an output layer.

In two hidden layers, the Gaussian error linear unit (GeLU) function [40–42] is used as the activation function in each hidden node. This activation function performs a nonlinear conversion, as shown in Figure 3, which can be expressed as follows:

$$GeLU(x_i) = 0.5x_i(1 + \tanh(\sqrt{\frac{2}{\pi}}(x_i + 0.4472x_i^3))), \ i = 1, 2, 3 \tag{8}$$

where x_i is the 1D feature vector used as input, $i = 1, 2, 3, \ldots, n$, $X = [x_1, x_2, x_3, \ldots, x_n]$. The training of the multilayer classifier uses the back-propagation algorithm to adjust the connecting weighted parameters of the classifier and set the loss function as the convergent condition for terminating the training stage. For multi-class classification, multiple classes of binary cross-entropy functions [7,43–45] are shown in Equation (9):

$$L = -\frac{1}{K}\sum_{j=1}^{m}\sum_{k=1}^{K} t_{j,k}\log_2(y_{j,k}) + (1 - t_{j,k})\log_2(1 - y_{j,k}), \ j = 1, 2, 3 \tag{9}$$

$$Y = GeLU(XW) \tag{10}$$

where $t_{j,k}$ is the target value (desired class), $T = [t_{1,k}, t_{2,k}, t_{3,k}, \ldots, t_{m,k}]$ for multiple classes; $y_{j,k}$ is the outputted prediction value, $Y = [y_{1,k}, y_{2,k}, y_{3,k}, \ldots, y_{m,k}]$; and m is the number of classifications. This study sets $m = 2$, either normal or abnormal, coding as $Y = [1, 0]$ and $Y = [0, 1]$, respectively, $k = 1, 2, 3, \ldots, K$, is the number of training data, and W is the weighted parameter matrix of the classifier with a fully connecting network.
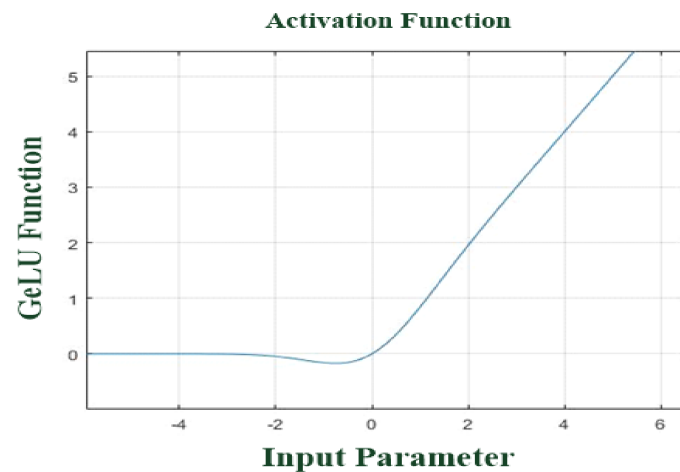
Figure 3. *GeLU* activation function in each hidden node.

2.2. Adaptive Moment Estimation Method

In the classification layer, the network connecting weighted parameters are adjusted by using BPA to minimize the loss function. The smaller the cross-entropy value, the smaller the classification error rate, and the higher the accuracy that can be obtained. The adjustment formula of weight parameters of the classifier uses the adaptive moment estimation (ADAM) optimization method, as follows [31,46]:

$$w(p+1) = w(p) + \eta\frac{\hat{m}(p)}{\sqrt{\hat{v}(p)} + \delta} \tag{11}$$

where $\hat{m}(p) = \frac{m(p)}{1-\beta_1}$ and $\hat{v}(p) = \frac{v(p)}{1-\beta_2}$ are adjustment parameters; η is the learning rate of the classifier; δ is the smoothing value; $\beta_1 = 0.900$ and $\beta_2 = 0.999$ are the attenuation rates of each iteration; $p = 1, 2, 3, \ldots, p_{max}$; and p_{max} is the maximum number of iterations. Each iteration computation adjusts the weighted parameters of the classifier within a limited range with the parameters of Equation (11), as shown in Equations (12) and (13) [31,46]:

$$m(p) = \beta_1 m(p-1) + (1-\beta_2)\frac{\partial L}{\partial w} \tag{12}$$

$$v(p) = \beta_1 v(p-1) + (1-\beta_2)\left(\frac{\partial L}{\partial w}\right)^2 \tag{13}$$

With the above-mentioned formulas, the best parameters can be quickly obtained using matrix operations and the loss function (9) can be minimized.

The proposed multilayer classifier used in this study is a fully connecting network. The number of nodes in the hidden layer in the middle of the network is set as per the graph's data type and complexity. Optimization algorithms are used to adjust the connecting weighted parameters of the classifier to minimize the loss function. The input ROI image size used in this study is 100 pixels × 100 pixels. There is one fractional-order convolutional layer, two convolutional layers, two maximum pooling layers, a flattening layer, and a fully connected multilayer classifier. The relevant information about the proposed multilayer CNN is shown in Table 1.

Table 1. Relevant information about the proposed multilayer CNN.

Layer Function	Manner	Feature Pattern
Input Feature Pattern	ROI Extraction with 100 × 100 Bounding Box	2D, 100 pixels × 100 pixels
1st Convolutional Layer	3 × 3 Fractional-Order Convolutional Window, Stride = 1	2D, 100 pixels × 100 pixels
2nd Convolutional Layer	3 × 3 Kernel Convolution Window, Stride = 1	2D, 100 pixels × 100 pixels
2nd Pooling Layer	2 × 2 Maximum Pooling Layer, Stride = 2	2D, 50 pixels × 50 pixels
3rd Convolutional Layer	3 × 3 Kernel Convolution Window, Stride = 1	2D, 50 pixels × 50 pixels
3rd Pooling Layer	2 × 2 Maximum Pooling Layer, Stride = 2	2D, 25 pixels × 25 pixels
Flattening Layer	Flattening Process	1D, 1 × 625 feature vector
Classification Layer	Multilayer Classifier: 625 input nodes, 168 hidden nodes 64 hidden nodes, 2 output nodes (for normality and abnormality) Algorithm: ADAM Algorithm	1 × 625 Feature Vector Feeding into Multi-Layer Classifier

2.3. Classifier's Performance Evaluations

This study uses the cross-validation ten times to evaluate the performance of the proposed multilayer CNN-based classifier. Each time, the dataset is divided into the two groups of normal and abnormal feature patterns. The dataset is then randomly divided into two halves: 50% training dataset and 50% testing dataset. Repeat the procedure ten folds to confirm the performances of the proposed classifier, such as the evaluation indicators shown in the formulas for precision (%), recall (%), F1 score, and accuracy (%) [35,36,47] in Table 2. For each fold cross-validation, the multilayer CNN-based classifier will produce a confusion matrix comprising four parameters, including TP (True Positive), FP (False Positive), TN (True Negative), and FN (False Negative). These parameters help to determine the indexes for evaluating the performances of the proposed classifier. The precision (%) indicator represents the rate that a TP can be correctly identified among all TPs (positive samples). Generally, a model must be larger than 80% to be recognized as a good classifier. The recall (%) indicator is defined as TP/(TP + FN). The F1 Score ($\in [0,1]$) is the harmonic mean of precision (%) and recall (%) indexes. Its index combines the two in a single evaluation index. The higher the value of the F1 score and the closer it is to 1, the better the classifier is at prediction.

Table 2. Formulas for the evaluation criteria of the proposed classifier, including precision (%), recall (%), accuracy (%), and F1 score.

	Actual		Total	Precision (%)
Predicted	TP	FP	TP + FP	(TP)/(TP + FP)
	FN	TN	FN + TN	
Total	TP + FN	FP + TN	Accuracy (%): (TP + TN)/(TP + FP + TN + FN)	
Recall (%)	(TP)/(TP + FN)			
F1 Score	(2TP)/(2TP + FP + FN)			

2.4. Computer Assistive System for Automatic Breast Tumor Screening

This study uses the LabVIEW 2019 (NI^{TM}) software to develop a computer assistive system for automatic breast tumor screening, integrating: (1) ROI image extraction, (2) feature enhancement and extraction, and (3) breast tumor screening classifier and other functions. Algorithms for functions (1) and (2) are developed using the MATLAB Script tool. The multilayer CNN algorithm and the interface shown in Figure 4 are written by Python software. The interface works as follows:

- Zone ①: Sets the source path of breast mammography images;
- Zone ②: Loads and displays the selected mammography images;
- Zone ③: As per the priority order, extract ROI images and perform automatic tumor screening. In this study, six areas at which tumors are most possibly identified are designated. The CAS automatically prioritizes the ROI cutting feature patterns (100 pixels × 100 pixels), as seen in Figure 2, and then screens those areas. The block marked ③ can show the output of the classifier, the identification result, and the classification information. The red and green circles show the normality and abnormality. The output value of the classifier must be >0.5 to have a high degree of confidence that there is a suspected breast tumor.
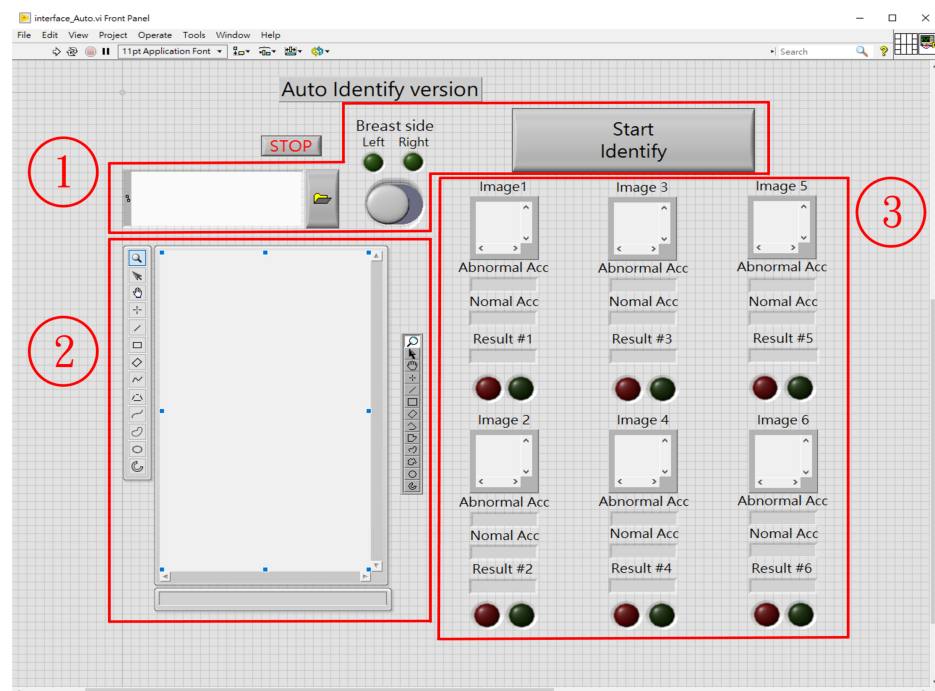

Figure 4. The human machine interface for automatic breast tumor screening: Zone ①: Sets the source path of images, Zone ②: Loads and displays the images, Zone ③: Extract ROI images and perform automatic tumor screening.

The human machine interface designed in this study can be used in clinical applications when switched to manual mode. The user can manually select six ROI blocks and

screenshots and then save these images in a temporary database. When the number of screenshots reaches the set number (i.e., default of six ROIs), the multilayer CNN classifier performs the classification task as per the priority order of the queue and returns the identification results. The clinician then receives messages to confirm the existence of a possible tumor.

2.5. Experimental Setup

In the MIAS database, the most common image size was 4320 pixels × 2600 pixels. Thus, in this study, the dimensions 4320 pixels × 2600 pixels were selected for breast tumor screening [33,34]. The vertical and horizontal resolutions of each image were identical at 600 dpi, with a bit depth of 24 bits. A total of 156 mammography images (78 subjects), including 62 images with malignant (M) or benign (B) tumors and 94 images without tumors, were obtained. Given a specific bounding box measuring 100 pixels × 100 pixels, feature patterns were screenshots from the 156 images. In total, 932 feature patterns, including 564 tumors and 368 tumor-free screenshots, were obtained. In each classifier's training stage, 282 tumor and 184 tumor-free screenshots (50% feature patterns) were randomly selected to train the multilayer CNN classifier. The remaining 50% of the feature patterns were used to evaluate the classifier's performance for each cross-validation. This study used the relevant data, as shown in Table 1, to establish a multilayer CNN-based classifier. We designed a fractional-order convolutional layer, two general convolutional layers, and two MP processing layers for feature enhancement and extraction. The convolutional layer had 16 kernel convolutional windows. In the kernel window, the sliding window moved the number of columns and rows in steps of 1 (stride = 1) at each point of the convolution operation. The padding parameter was set to 1 to maintain the feature pattern after the convolutional operation. During the pooling process, the MP window moved with a stride of 2 (stride = 2) each time. During each feature enhancement, and extraction process, the possible tumor features and contours were gradually enhanced by the convolutional-pooling processes; hence, it can be observed that the multilayer CNN-based classifier can improve the accuracy of pattern recognition based on these enhanced features. Tenfold cross-validation was performed using precision (%), recall (%), F1 score, and accuracy (%) as indicators [35,36,48] to evaluate the prediction performance of the proposed classifier. Figure 5 shows the visualization of the confusion matrix; for example, the classifier used 466 images for rapid screening, and the results show 178 TPs, 13 FPs, 269 TNs, and 6 FNs. The precision (%), recall (%), F1 score, and accuracy (%) can be calculated from the confusion matrix.
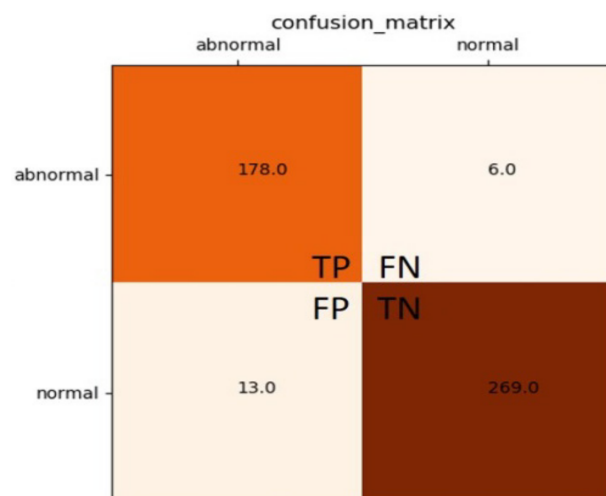

Figure 5. Output confusion matrix of the multilayer CNN-based classifier.

3. Experimental Results

This study compares the training time, accuracy, training curve, and prediction performance of multilayer CNN-based classifiers using different numbers of convolutional layers and pooling layers, different types of convolutional windows, and different sizes of convolutional window dimensions, as seen in Table 3. The comparison items are briefly described as follows:

- The number of convolutional layers and pooling layers: This study increases the number of convolutional layers and pooling layers from 1 to 5 and the sizes of convolution windows from 3×3 to 11×11. The processing windows for the pooling layers are set to 2×2, and the second to fifth convolutional layers have 16 kernel convolution windows to perform feature enhancement and extraction.
- The first convolution windows: This study selects three types of convolutional windows, including fractional-order ($v \in (0, 1)$), Sobel (first order, $v = 1$), and Histeq [35,36,38,47,49,50] windows to pre-enhance the feature pattern.

Table 3. Different convolutional layer models of the multilayer CNN-based classifier (Models #1–#5).

Model	1st Convolution Window	2nd Convolution Window	3rd Convolution Window	4th Convolution Window	5th Convolution Window	Stride	Padding
1	$3 \times 3, 2$	-	-	-	-	1	1
2	$3 \times 3, 2$	$5 \times 5, 16$	-	-	-	1	1
3	$3 \times 3, 2$	$5 \times 5, 16$	$7 \times 7, 16$	-	-	1	1
4	$3 \times 3, 2$	$5 \times 5, 16$	$7 \times 7, 16$	$9 \times 9, 16$	-	1	1
5	$3 \times 3, 2$	$5 \times 5, 16$	$7 \times 7, 16$	$9 \times 9, 16$	$11 \times 11, 16$	1	1

In general, a multilayer CNN may have dozens of layers of convolutional layers. As shown in Table 3, this study designs five different multilayer models of convolutional layers and convolutional window sizes. It compares the training time and accuracy of five models to confirm the feasibility of the CNN model constructed in this study. Moreover, we establish three models for feature enhancement and extraction, as seen in Table 4, by combining different kernel convolutional windows and dimensions (3×3 and 5×5) and comparing the performance of these different models. These tests will help determine the best model for clinical applications in automatic breast tumor screening. In addition, we also use a multi-core personal computer (Intel® Q370, Intel® Core™ i7 8700, DDR4 2400 MHz 8G*3) as a development platform to implement the multilayer CNN-based classifier suggested in this study and use the graphics processing unit (GPU) (NVIDIA® GeForce® RTX™ 2080 Ti, 1755 MHz, 11 GB GDDR6) to speed up the time it takes for digital image processing. The feasibility study was validated as described in detail in the subsequent sections.

Table 4. Comparisons of average training CPU time and average accuracy (%) for five different CNN models.

Model	1	2	3	4	5
Training CPU Time (min)	<30	<240	<7	<10	<180
Average Accuracy (%)	90.99%	90.34%	95.92%	95.28%	95.71%

3.1. Testing of Different Multilayer CNN Models and Determination of the Most Suitable Architecture

As shown in Table 3, this study designs five models comprising 1–5 convolutional layers. The first layer is a fractional-order convolutional layer with two 3×3 convolutional windows; it is used to perform the 2D spatial convolution operations. The second to fifth convolutional layers use 16 kernel convolutional windows with different sizes of convolutional windows ($3 \times 3, 5 \times 5, 7 \times 7, 9 \times 9, 11 \times 11$) [51] and 16 MP windows for feature enhancement and extraction. Finally, a fully connected network with two hidden layers using the adaptive moment estimation optimization method [46,48] adjusts the

connecting weighted parameters with Equations (11)–(13) such that the predetermined classification is obtained. This study uses the same training dataset, specifically, 466 feature patterns (282 tumor and 184 tumor-free screenshots), to train and test the five different CNN models. We randomly generate initial parameters and train each model at least five times, thus recording the required training time and classification accuracy rate to compare the average training CPU time (min) and average accuracy (%) of models, as seen in Table 4. The testing results indicate that the three-layer convolutional layer model (Model #3 shown in Table 3) has an average training CPU time of lower than 7 min and an average accuracy (%) of larger than 95% with 466 untrained feature patterns. While the average accuracy (%) of the four- and five-layer models can reach larger than 95%, these two models require more training CPU time. Therefore, Model #3 is the most suitable model for developing an automatic breast tumor screening classifier in clinical applications.

3.2. Testing of the First Convolutional Layer and Determination of the Window Type

As seen in Table 5, three types of convolution windows in the first convolutional layer, including fractional-order windows, Sobel windows, and Histeq windows, are used perform the 2D spatial convolutions [35,36,47,49–52]. Figure 6a shows the original image and image enhancement results of these three types of convolution windows. Figure 6b shows the pixel grayscale value distribution map after the image is magnified. Compared with the original image grayscale value (0–255) distribution map, the convolution result of the first derivative-based Sobel convolution window [49] has a smoothing effect and is anti-noise but requires a considerable amount of calculations while performing convolutions; moreover, this window type produces a thicker edge contour, which results in lower accuracy in identifying the position of the target object. We can use a second-order-based convolutional window for feature enhancement, but this window is fairly susceptible to noise and thus unsuitable for obtaining the edge contour of the target; this type of window is generally used for binarization applications. The Histeq convolution (histogram regularization) [50] yields a histogram of the number of times each grayscale value appears. This histogram can describe the statistical information of the grayscale values of the image and allows the direct observation of the characteristics of the image, such as its brightness and contrast. It is primarily used for image segmentation and adjustment of grayscale values in the image. As shown in Figure 6b, the non-zero value of the histogram has a wide and uniform distribution, which indicates that the contrast of the image is high. The pixel value of the image may be readjusted to a value between 0 and 255 by using linear, piecewise linear, and nonlinear transformation functions [53]. These transformation functions are primarily used to magnify the contrast of the original image. The overall grayscale value distribution map shifts to the right, and the contrast of the image increases, thereby minimizing the effort required to highlight the outline of the malignant tumor, as shown in Figure 6a. However, this method is susceptible to factors such as illumination, viewing angle, and noise. The Histeq (histogram normalization) function [50] can automatically determine the grayscale transformation function and yield an output image with a uniform histogram. It is primarily used for contrast adjustments over a small range but could amplify background noise.

Table 5. Different convolutional layer models for feature enhancement and extraction (Models #1–#3).

Model	1st Convolutional Window and Window Size	2nd Convolutional Window and Window Size	3rd Convolutional Window and Window Size	Stride/ Padding	Maximum Pooling Window	Stride
1	Fractional Order, 3×3, 2	3×3 or 5×5, 16	3×3 or 5×5, 16	1/1	2×2, 16	2
2	Sobel (First Order), 3×3, 2	3×3 or 5×5, 16	3×3 or 5×5, 16	1/1	2×2, 16	2
3	Histeq, 3×3, 2	3×3 or 5×5, 16	3×3 or 5×5, 16	1/1	2×2, 16	2

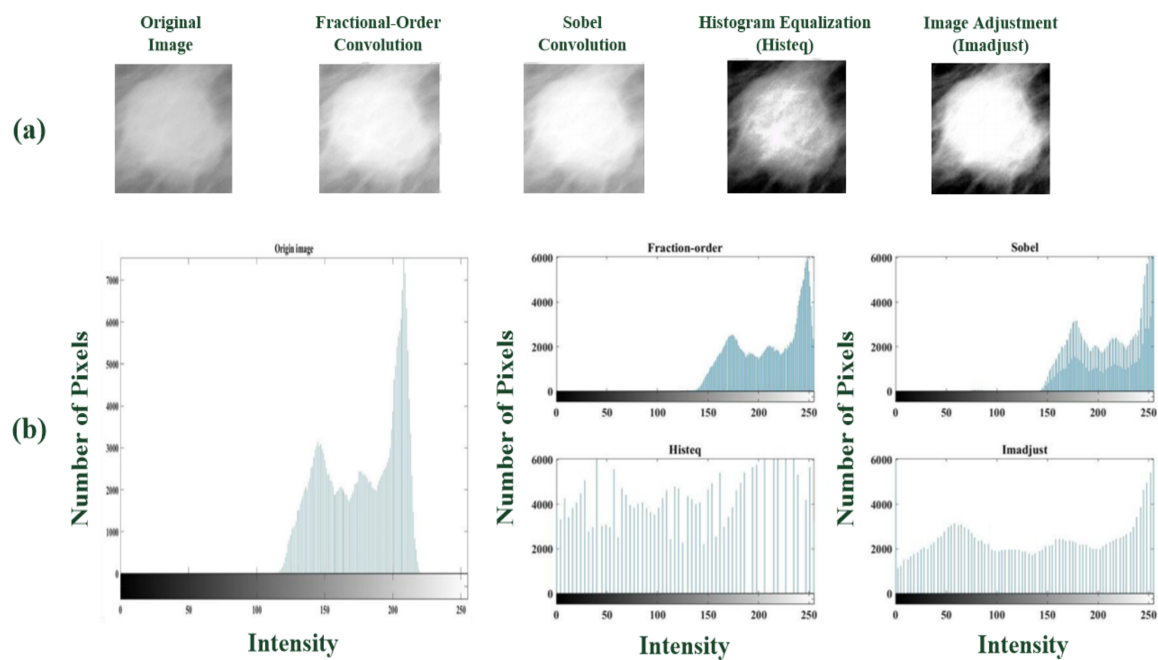**Figure 6.** (**a**) The original image (malignant tumor) and the enhanced image of the three convolutions, (**b**) The original image and the pixel greyscale distribution map after image enhancement.

When fractional-order spatial convolution is conducted in 2D space, the overall grayscale value distribution moves to the right, which increases the contrast of the image and filters out noise. Thus, it shows better performance than the Sobel convolutional operation. Therefore, in the first convolutional layer, this study selects a 3×3 fractional-order convolution window for the first convolutional layer. In addition, the literature [35,36] proposes that setting the fractional-order parameter $v = 0.30$–0.40, which is also used for feature enhancement in X-ray images, could yield promising results. Thus, our study selects the parameter $v = 0.35$. After 2D spatial convolution and normalization operations, the contour of the sharpened target can be obtained by using Equations (1)–(5). Strengthening the target's features, retaining non-characteristic information, and removing noise are helpful for the subsequent second- and third-layer feature extraction operations and further pattern recognition.

3.3. Multilayer CNN-Based Classifier Testing and Validation

This study uses Model #1, as shown in Table 5, which adopts three convolutional layers, and the same completely connected classification layer to develop four models, as shown in Table 6. The convolutional window sizes of the second and third layers are combinations of $(3 \times 3, 3 \times 3)$, $(3 \times 3, 5 \times 5)$, $(5 \times 5, 3 \times 3)$, and $(5 \times 5, 5 \times 5)$ [51]. The image dataset is divided into two groups of equal size. The four models use 466 trained and 466 untrained feature patterns to test and confirm the performance of the classifier. A total of 1000 epochs are set for the training classifier, with the trained and untrained feature patterns. Figure 7 shows (a) the training performance of the classifier and (b) the training history curve of the classification performance validation; in (b), the solid blue line represents the results of the training performance test and the solid orange line indicates the results of classification performance validation. As the number of epoch training increases, the classifier's output accuracy (%) gradually increases. The four classifier models require an average of lower than 240 s (lower than 4 min) CPU time to complete the training and testing tasks, as seen in Table 6. Then, the trained and untrained feature patterns are randomly selected, and the accuracy (%) of the four classifier models is tested by performing 10-fold cross-validation ($K_f = 10$). Table 7 shows the overall cross-validation results. Figure 7a indicates that the accuracy (%) of Models #2 and #3 can be improved over 600 epochs of training. By comparison, the accuracy of Models #1 and #4 can be improved over 200–400 epochs,

after which it converges, and the accuracy (%) of classification approaches the maximum. The training convergence curve of the classifier is shown in Figure 7b. The accuracy (%) of the four models may reach larger than 95%. To shorten the classifier's design cycle and reduce the memory requirements for storing classifier parameters, we recommend using the architectures of Models #1 and #4 to establish and implement the multilayer CNN-based classifiers.

Table 6. Comparisons of the training CPU time for different models of multilayer CNN-based classifier.

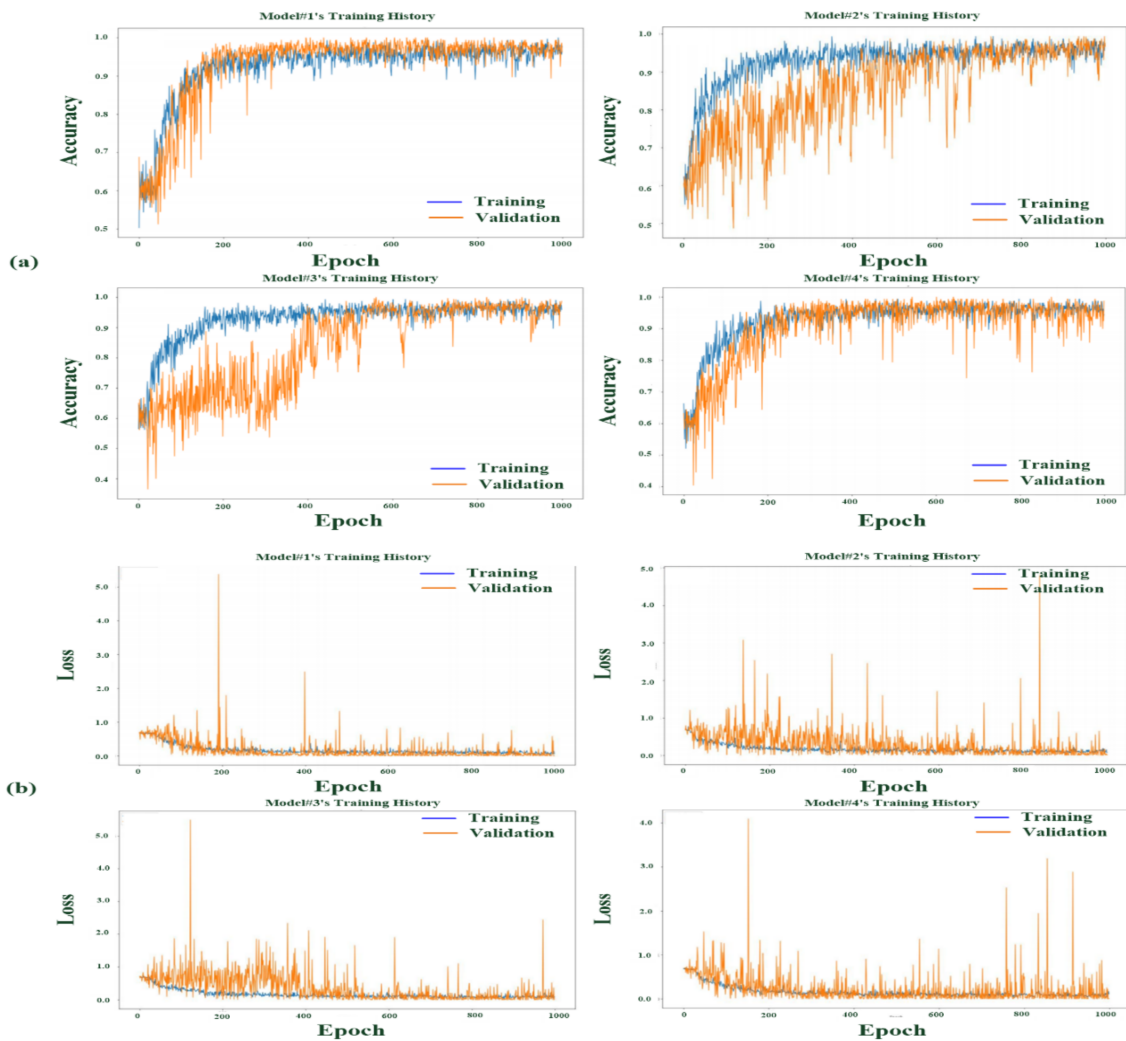| Model	1st Convolutional Layer	2nd Convolutional Layer	2nd Pooling Layer	3rd Convolutional Layer	3rd Pooling Layer	Classification Layer (Fully Connecting Network)	Average Training Time (s)
1		$3 \times 3, 16$		$3 \times 3, 16$			<280
2	$3 \times 3, 2$	$3 \times 3, 16$	$2 \times 2, 16$	$5 \times 5, 16$	$2 \times 2, 16$	625 input nodes, 168 1st hidden nodes, 64 2nd hidden nodes, 2 output nodes	<220
3		$5 \times 5, 16$		$3 \times 3, 16$			<240
4		$5 \times 5, 16$		$5 \times 5, 16$			<330

Figure 7. Training history curves of the multilayer CNN-based classifier. (**a**) Training performance test and classification performance validation as seen classification accuracy versus the training epoch and (**b**) classifier training convergence curve as a loss function versus the training epoch.

Table 7. Cross-validation results for different multilayer CNN models ($K_f = 10$).

Test Fold / Model	1	2	3	4	5	6	7	8	9	10	Average Accuracy
1	96.14	97.43	98.07	97.96	98.93	98.07	96.35	95.60	96.89	98.28	97.37
2	97.42	98.93	98.28	97.64	99.14	97.21	97.85	98.28	99.14	97.42	98.13
3	96.14	97.64	98.50	98.71	99.14	97.85	97.64	95.06	97.21	99.57	97.75
4	98.93	98.71	96.14	97.32	99.36	98.18	91.74	90.34	97.64	90.02	95.84

Considering the experimental results listed in Table 6, the architecture of Model #1 is selected to establish the screening classifier. After training is completed, 466 untrained feature patterns, including 184 abnormal and 282 normal patterns, are randomly selected from the dataset to validate the performance of the classifier. The experimental results of the classifier produce a visual confusion matrix. The testing result of the abnormal pattern yields TP = 178 and FP = 6, while that of the normal pattern yields TN = 269 and FN = 13; these values can be used as variables in Table 2 to compute the four evaluation indices of the classifier. In this study, precision (%) = 96.74%, recall (%) = 93.19%, F1 score = 0.9493, and accuracy (%) = 95.92%. Precision (%) is the standard for predicting TP, and recall (%) is the true accuracy of TP. Both indicators may be greater than 80%. Recall (%) is also called the positive predictive value (PPV), which is the so-called TP in the detection case. The general PPV index is larger than 80%, which means the proposed classifier has promising predictive performance. The F1 score fuses the indicators of precision (%) and recall (%), and F1 score larger than 0.9000 generally indicates a good classification model. Youden's index (YI) is a fusion evaluation index of sensitivity (Sens) and specificity (Spec) [54], which reflects the performance of the classifier for detecting abnormalities. The larger the YI, the better the performance of the classifier for detection and validation and the greater its authenticity. The testing results show YI = 91.01% (Sens = 93.19%, Spec = 97.82%). Given that all evaluation indicators considered in this work exceed 90%, Model #1 indeed has an architecture that supports good classification accuracy and performance, as seen in the tenfold cross-validation ($K_f = 10$) for averages of precision (%), recall (%), accuracy (%), and F1 score in Table 8. Hence, we suggest Model #1 to carry out a multilayer CNN-based classifier for automatic breast tumor screening. In addition, as seen in Table 9, we also set 4, 8, 16, and 32 Kernel convolutional windows and 4, 8, 16, and 32 maximum pooling windows in second and third convolutional-pooling layers, respectively, for establishing four models (Models #1–1 to #1–4). With the tenfold cross-validation, trained feature patterns are randomly selected, the average training CPU time of Models #1–1 and #1–2 is less than Model #1–3 with 16 Kernel convolutional windows and 16 maximum pooling windows. It can be seen that Model #1–4 comprises 32 Kernel convolutional windows and 32 maximum pooling windows will increase the average training CPU time and complex computational processes at each cross-validation. With the tenfold cross-validation, untrained feature patterns are also randomly selected, as seen in Tables 10–13, the proposed architecture of multilayer classifier (Model #1–3) has promising classification accuracy and performance in terms of average precision (%), average recall (%), average accuracy (%), and average F1 score. Additionally, the proposed CNN architecture with different convolutional windows in the first convolutional layer, including fractional-order, Sobel (first-order), and Histeq convolutional windows, is used to test the performance of breast tumor screening model. Through the tenfold cross-validation, the CNN classifier with a fractional-order convolutional window in the first convolutional layer, as Model #1 in Table 14, has better classification accuracy (larger than 95%) than Model #2 (larger than 85%) and Model #3 (larger than 90%).

Table 8. Experimental results of K-fold cross-validation (K_f = 10) for the proposed deep-learning-based CNN.

Test Fold	Precision (%)	Recall (%)	Accuracy (%)	F1 Score
1	95.00	96.48	95.60	0.9574
2	94.38	95.92	95.20	0.9514
3	94.82	91.80	93.80	0.9389
4	95.02	96.50	95.60	0.9575
5	96.51	91.68	95.40	0.9577
6	94.09	94.84	94.40	0.9447
7	92.80	97.61	95.00	0.9515
8	92.77	95.85	94.40	0.9429
9	96.46	95.70	96.00	0.9608
10	95.19	95.54	95.00	0.9536
Average	95.19	95.19	95.04	0.9516

Table 9. Comparisons of the training CPU time for multilayer CNN-based classifiers with different numbers of Kernel convolutional windows and maximum pooling windows in second and third convolutional-pooling layers.

Model	1st Convolutional Layer	2nd Convolutional Layer	2nd Pooling Layer	3rd Convolutional Layer	3rd Pooling Layer	Classification Layer (Fully Connecting Network)	Average Training Time (s)
1–1		$3 \times 3, 4$	$2 \times 2, 4$	$3 \times 3, 4$	$2 \times 2, 4$		<150
1–2		$3 \times 3, 8$	$2 \times 2, 8$	$3 \times 3, 8$	$2 \times 2, 8$	625 input nodes, 168 1st hidden nodes,	<240
1–3	$3 \times 3, 2$	$3 \times 3, 16$	$2 \times 2, 16$	$3 \times 3, 16$	$2 \times 2, 16$	64 2nd hidden nodes, 2 output nodes	<280
1–4		$3 \times 3, 32$	$2 \times 2, 32$	$3 \times 3, 32$	$2 \times 2, 32$		<330

Table 10. Experimental results of K-fold cross-validation (K_f = 10) for Model #1–1 with 4 Kernel convolutional windows and 4 maximum pooling windows in each convolutional-pooling layer.

Test Fold	Precision (%)	Recall (%)	Accuracy (%)	F1 Score
1	85.30	84.70	87.80	0.8640
2	87.10	93.00	91.20	0.9000
3	82.80	93.40	89.00	0.8760
4	86.10	91.50	93.20	0.9200
5	84.00	92.00	89.20	0.8780
6	93.20	84.80	91.08	0.8880
7	91.20	95.70	95.40	0.9480
8	90.10	94.80	93.60	0.9260
9	92.40	86.70	91.40	0.8950
10	97.20	99.10	98.40	0.9810
Average	88.94	91.57	92.30	0.9076

Table 11. Experimental results of K-fold cross-validation (K_f = 10) for Model #1–2 with 8 Kernel convolutional windows and 8 maximum pooling windows in each convolutional-pooling layer.

Test Fold	Precision (%)	Recall (%)	Accuracy (%)	F1 Score
1	96.60	95.70	96.80	0.9620
2	94.80	95.60	96.00	0.9530
3	96.10	93.00	95.40	0.9450
4	88.30	96.20	93.00	0.9210
5	90.10	95.30	93.60	0.9260
6	92.10	93.40	93.80	0.9270

Table 11. *Cont.*

Test Fold	Precision (%)	Recall (%)	Accuracy (%)	F1 Score
7	93.40	93.80	94.70	0.9360
8	92.30	93.00	93.80	0.9270
9	94.10	97.60	96.40	0.9580
10	90.30	96.70	94.20	0.9340
Average	92.81	95.03	94.77	0.9389

Table 12. Experimental results of K-fold cross-validation (K_f = 10) for Model #1–3 with 16 Kernel convolutional windows and 16 maximum pooling windows in each convolutional-pooling layer.

Test Fold	Precision (%)	Recall (%)	Accuracy (%)	F1 Score
1	96.70	96.20	97.00	0.9640
2	97.60	95.30	96.60	0.9570
3	95.10	93.40	95.40	0.9420
4	97.10	94.00	96.20	0.9540
5	97.20	97.10	97.60	0.9680
6	93.00	94.00	94.40	0.9340
7	95.60	92.40	95.00	0.9400
8	98.00	97.20	98.10	0.9760
9	96.50	95.70	96.00	0.9610
10	95.20	95.50	95.00	0.9540
Average	96.30	95.04	95.93	0.9553

Table 13. Experimental results of K-fold cross-validation (K_f = 10) for Model #1–4 with 32 Kernel convolutional windows and 32 maximum pooling windows in each convolutional-pooling layer.

Test Fold	Precision (%)	Recall (%)	Accuracy (%)	F1 Score
1	96.70	96.20	97.00	0.9640
2	99.00	96.20	99.00	0.9760
3	90.80	93.40	93.20	0.9210
4	95.60	93.40	95.40	0.9230
5	99.50	97.60	98.80	0.9860
6	95.30	95.70	96.20	0.9550
7	66.00	95.60	77.20	0.7800
8	96.70	98.60	98.00	0.9770
9	97.40	91.50	95.40	0.9440
10	99.00	94.80	97.40	0.9690
Average	93.60	95.30	94.60	0.9395

Table 14. Comparisons of average accuracy (%) for the CNN-based classifier with different convolutional windows in first to third convolutional layers.

Model	First Convolutional Window	Second and Third Convolutional Window	Second and Third Pooling Window	Classification Layer	Average Accuracy
1	Fractional-Order Convolutional Window (2)	Kernel convolution Window (16)	Maximum Pooling Window (16)	One Input Layer, Two Hidden Layers, One Output Layer	>95%
2	Sobel (First-Order) Convolutional Window (2)	Kernel convolution Window (16)	Maximum Pooling Window (16)		>85%
3	Histeq Convolutional Window (2)	Kernel convolution Window (16)	Maximum Pooling Window (16)		>90%

Figure 8 shows the left breast mammogram image of a patient (File Name #mdb31-5ll [33,34]). In this case study, the right breast mammogram image (File Name #mdb316rl) is normal (background tissue: Dense-glandular (D)), the left breast has a

benign tumor (B), the center coordinates of the tumor are (1900, 317), the background tissue is D, and the circumscribed masses are labeled (CIRC) [33,34]. In this study, using the automatic screening system developed [55], as per the pre-selected priority order of screening (①→⑥), the sequence of ROI blocks is shown in Figure 8, and the automatic screening results show four TP (①,②,③, and⑥) and two TNs (④and⑤). In this case, the large tumor spans four ROI blocks①–③and⑥. Therefore, the screening results show TP for identifying a possible breast tumor, the reliability of the classifier output judged to be abnormal is larger than 0.50, and the abnormality is flagged by a red message. The screening system can be switched to manual mode. Similar to the automatic screening results, the four ROI blocks①–③and⑥can be manually circled, screenshotted, and stored in the queue in the order of manual screenshots. The classifier automatically performs image recognition in sequence, and the corresponding recognition results and messages are returned so that the clinician can confirm the possible tumor conditions.
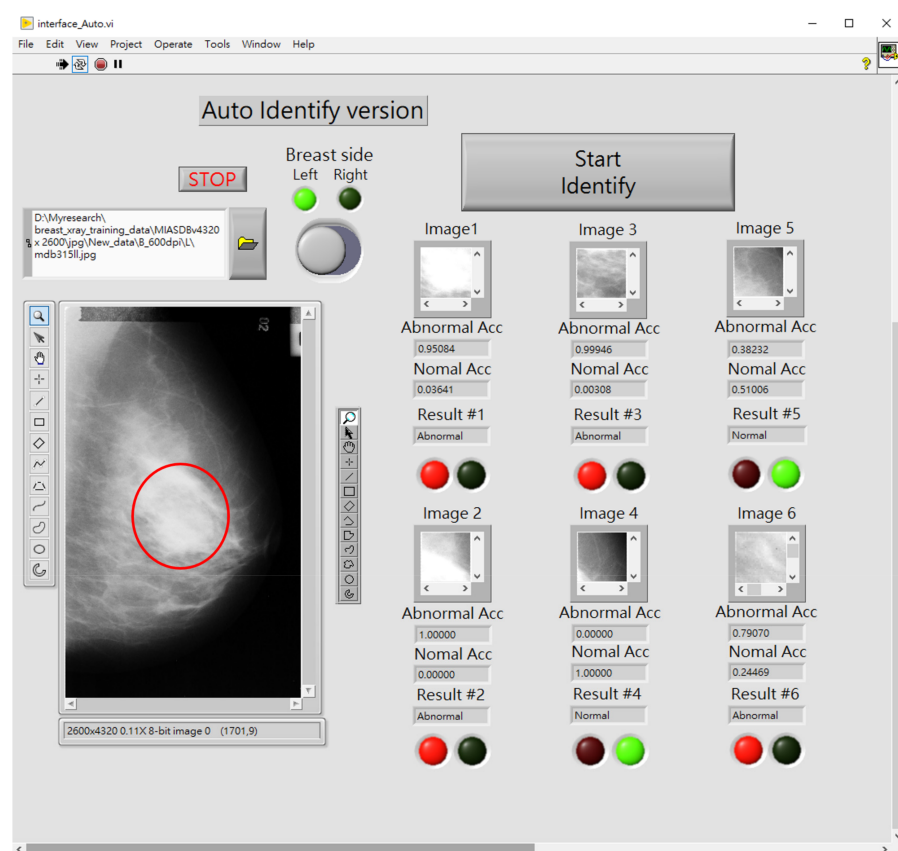
Figure 8. Human–machine interface of the computer assistive system and its automatic screening results.

3.4. Discussion

This study designs a mammography classification method incorporating a multilayer CNN-based classifier for automatic breast tumor screening in clinical applications. The proposed classifier algorithm is implemented in the LabVIEW 2019 (NI^{TM}) software, MATLAB Script tools, and open-source Tensorflow platform (Version 1.9.0) [28] and integrated into a computer assistive system with the automatic and manual feature extraction and breast tumor screening modes. The fractional-order convolutional layer and two convolutional-pooling layers allow the image enhancement and sharpening of the possible tumor edges, contours, and shapes via one fractional-order and two kernel convolutional processes in the feature patterns. Through a series of convolution and pooling processes at different scales and different dimensions, the classifier can obtain nonlinearity feature representation from low-level features to high-level information [29]. Then, with the specific bounding

boxes (automatic or manual mode) for ROI extraction, enhanced feature patterns can then be distinguished for further breast tumor screening by the multilayer classifier in the classification layer. A gradient-descent optimization method, namely, the ADAM algorithm, is used in the back-propagation process to adjust the network weighted parameters in the classification layer. With K-fold (K_f = 10) cross-validation, the 466 randomly selected untrained feature patterns for each test fold, the proposed multilayer CNN-based classifier, has high recall (%), precision (%), accuracy (%), and F1 scores for screening abnormalities in both right and left breasts. Experimental results show that the proposed multilayer CNN model offers image enhancement, feature extraction, automatic screening capability, and higher average accuracy (larger than 95%) for separating the normal condition from the possible tumor classes. It has been observed from previous literature [3–7,10,56] that multilayer CNNs comprised several convolutional-pooling layers and a fully connecting network to establish a classifier for automatic breast tumor screening, and could also be applied for CT, MRI, chest X-ray, and ultrasound image processes, such as image classification and segmentation in clinical applications [19,23,28,35,36,51,55]. The combination of a cascade of deep learning and a fully connecting networks is also carried out by a multilayer CNN-based classifier, and a decision scheme [56]. For the screened suspicious region on mammograms, the cascade of the deep-learning method had 98% sensitivity and 90% specificity on the SuReMapp (Suspicious Region Detection on Mammogram from PP) dataset [56], and 94% sensitivity and 91% specificity on the mini-MIAS dataset [56]. This CNN-based multilayer classifier could extract multi-scale feature patterns, and increase the depth and width feature patterns by using multi-convolutional-pooling processes, which had an overall increase in accuracy. However, excessive multi-convolutional processes would completely lead to a loss of the internal data about the position and the orientation of the desired object, and an excessive multi-pooling processing would lose valuable information relating to the spatial relationships between features; thus, many processes were required to perform with GPU hardware for complex computational processes. Hence, the proposed optimal multilayer CNN architecture contained 2D spatial information in the fractional-order convolutional layer (with two fractional-order convolutional windows), and continuously enhanced the features with two-round convolutional-pooling processes (with 16 Kernel convolutional windows and 16 maximum pooling windows), which could extract the desired features at different scales and different levels. Thus, in comparison with the other deep-learning methods, the proposed multilayer classifier exhibited promising results for the desired medical diagnostic purpose. Hence, we have some advantages for the proposed CNN-based classifier, as follows:

- The ROI extraction, image enhancement, and feature classification tasks are integrated into one learning model;
- The fractional-order convolutional process with fractional-order parameter, v = 0.30–0.40, is used to extract the tumor edges in the first convolutional layer; subsequently, two kernel convolution processes are used to extract the tumor shapes;
- The ADAM algorithm is easy to implement and operate with large datasets and parameter adjustment;
- The proposed CNN-based classifier has better classification accuracy than the CNN architecture with Sobel and Histeq convolutional windows in the first convolutional layer.

4. Conclusions

The proposed CNN architecture had better learning ability for complex feature patterns in massive-sized training datasets, and also had more promising classifier performance than traditional CNN-based classifiers and a cascade of deep-learning-based classifiers. Through experimental test and validation, we suggested optimal architecture for a simplified and established multilayer CNN-based classifier, which consisted of a fractional-order convolutional layer, two Kernel convolutional-pooling layers, and a classification layer. Hence, this optimal CNN-based classifier could replace manual screening for tasks requiring specific expertise and experience for medical image examination, which could

also raise its indication in clinical applications with CBIS-DDSM and SuReMapp dataset for the proposed training classifier. Additionally, in real-world applications, clinical mammography with biomarkers are continuously obtained, the new feature patterns can be extracted and added to the current database to further train the CNN-based classifier, which can keep its intended medical purpose and can also be used as a computer-aided decision-making tool and software in a medical device tool.

Abbreviations

CNN	Convolutional Neural Network
CBIS-DDSM	Curated Breast Imaging Subset of a Digital Database for Screening Mammography
BC	Breast Cancer
AI	Artificial Intelligence
BD	Big Data
MRI	Magnetic Resonance Imaging
CT	Computer Tomography
ROI	Region of Interest
2D	Two-Dimensional
1D	One-Dimensional
ML	Machine Learning
MP	Maximum-Pooling
FP	Flattening Process
MLP	Multilayer Perceptron
GPU	Graphics Processing Unit
ADAM	Adaptive Moment Estimation Method
GeLU	GeLU
MIAS	Mammographic Image Analysis Society
SuReMapp	Suspicious Region Detection on Mammogram from PP
TP	True Positive
FP	False Positive
TN	True Negative
FN	False Negative
PPV	Positive Predictive Value
YI	Youden's Index
Sens	Sensitivity
Spec	Specificity
B	Benign
M	Malignant

References

1. Ministry Health and Welfare, Taiwan. 2020 Cause of Death Statistics. 2021. Available online: https://dep.mohw.gov.tw/dos/lp-1800-113.html (accessed on 1 January 2022).
2. Tsui, P.-H.; Liao, Y.-Y.; Chang, C.-C.; Kuo, W.-H.; Chang, K.-J.; Yeh, C.-K. Classification of benign and malignant breast tumors by 2-D analysis based on contour description and scatterer characterization. *IEEE Trans. Med. Imaging* **2010**, *29*, 513–522. [CrossRef] [PubMed]
3. Kallenberg, M.; Petersen, K.; Nielsen, M.; Ng, A.Y.; Diao, P.; Igel, C.; Vachon, C.M.; Holland, K.; Winkel, R.R.; Karssemeijer, N.; et al. Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring. *IEEE Trans. Med. Imaging* **2016**, *35*, 1322–1331. [CrossRef] [PubMed]
4. Samala, R.K.; Chan, H.; Hadjiiski, L.; Helvie, M.A.; Richter, C.D.; Cha, K.H. Breast cancer diagnosis in digital breast tomosynthesis: Effects of training sample size on multi-stage transfer learning using deep neuralnets. *IEEE Trans. Med. Imaging* **2019**, *38*, 686–696. [CrossRef] [PubMed]
5. Valkonen, M.; Isola, J.; Ylinen, O.; Muhonen, V.; Saxlin, A.; Tolonen, T.; Nykter, M.; Ruusuvuori, P. Cytokeratin-supervised deep learning for automatic recognition of epithelial cells in breast cancers stained for ER, PR, and Ki-67. *IEEE Trans. Med. Imaging* **2020**, *39*, 534–542. [CrossRef]
6. Lee, S.; Kim, H.; Higuchi, H.; Ishikawa, M. Classification of metastatic breast cancer cell using deep learning approach. In Proceedings of the 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Jeju Island, Korea, 13–16 April 2021; pp. 425–428.
7. Chougrad, H.; Zouaki, H.; Alheyane, O. Deep convolutional neural networks for breast cancer screening. *Comput. Methods Programs Biomed.* **2018**, *157*, 19–30. [CrossRef]
8. Jia, G.; Lam, H.-K.; Althoefer, K. Variable weight algorithm for convolutional neural networks and its applications to classification of seizure phases and types. *Pattern Recognit.* **2021**, *121*, 108226. [CrossRef]
9. Li, X.; Zhai, M.; Sun, J. DDCNNC: Dilated and depthwise separable convolutional neural network for diagnosis COVID-19 via chest X-ray images. *Int. J. Cogn. Comput. Eng.* **2021**, *2*, 71–82. [CrossRef]
10. University of South Florida. *DDSM: Digital Database for Screening Mammography, Version 1 (Updated 2017/09/14)*; University of South Florida: Tampa, FL, USA. Available online: http://www.eng.usf.edu/cvprg/Mammography/Database.html (accessed on 1 January 2022).
11. McGarthy, N.; Dahlan, A.; Gook, T.S.; Hare, N.O.; Ryan, M.; John, B.S.; Lawlor, A.; Gurran, K.M. Enterprise imaging and big data: A review from a medical physics perspective. *Phys. Med.* **2021**, *83*, 206–220. [CrossRef]
12. Yaffe, M.J. Emergence of big data and its potential and current limitations in medical imaging. *Semin. Nucl. Med.* **2019**, *49*, 94–104. [CrossRef]
13. The European Federation of Organisations for Medical Physics (EFOMP). White Paper: Big data and deep learning in medical imaging and in relation to medical physics profession. *Phys. Med.* **2018**, *56*, 90–93. [CrossRef]
14. Diaz, O.; Kushibar, K.; Osuala, R.; Linardos, A.; Garrucho, L.; Igual, L.; Radeva, P.; Prior, F.; Gkontra, P.; Lekadir, K. Data preparation for artificial intelligence in medical imaging: A comprehensive guide to open-access platforms and tools. *Phys. Med.* **2021**, *83*, 25–37. [CrossRef] [PubMed]
15. Qiu, Y.; Lu, J. A visualization algorithm for medical big data based on deep learning. *Measurement* **2021**, *183*, 109808. [CrossRef]
16. Saranya, N.; Priya, S.K. Deep convolutional neural network feed-Forward and back propagation (DCNN-FBP) algorithm for predicting heart disease using internet of things. *Int. J. Eng. Adv. Technol.* **2021**, *11*, 283–287.
17. Zhang, J.; Qu, S. Optimization of backpropagation neural network under the adaptive genetic algorithm. *Complexity* **2021**, *2021*, 1718234. [CrossRef]
18. Sadad, T.; Munir, A.; Saba, T.; Hussain, A. Fuzzy C-means and region growing based classification of tumor from mammograms using hybrid texture feature. *J. Comput. Sci.* **2018**, *29*, 34–45. [CrossRef]
19. Comelli, A.; Bruno, A.; Di Vittorio, M.L.; Ienzi, F.; Legalla, R.; Vitabile, S.; Ardizzone, E. Automatic multi-seed detection for MR breast image segmentation. *Int. Conf. Image Anal. Process.* **2017**, *10484*, 706–717.
20. Lindquist, E.M.; Gosnell, J.M.; Khan, S.K.; Byl, J.L.; Zhou, W.; Jiang, J.; Vettukattilb, J.J. 3D printing in cardiology: A review of applications and roles for advanced cardiac imaging. *Ann. 3D Print. Med.* **2021**, *4*, 100034. [CrossRef]
21. Drozdzal, M.; Chartrand, G.; Vorontsov, E.; Shakeri, M.; Di Jorio, L.; Tang, A.; Romero, A.; Bengio, Y.; Pal, C.; Kadoury, S. Learning normalized inputs for iterative estimation in medical image segmentation. *Med. Image Anal.* **2018**, *44*, 1–13. [CrossRef]
22. Racz, A.; Bajusz, D.; Heberger, K. Multi-level compaeison of machine learning classifier and thrir performance metrics. *Molecules* **2019**, *24*, 2811. [CrossRef]
23. Allen, J.; Liu, H.; Iqbal, S.; Zheng, D.; Stansby, G. Deep learning-based photoplethysmography classification for peripheral arterial disease detection: A proof-of-concept study. *Physiol. Meas.* **2021**, *42*, 054002. [CrossRef]
24. Panwar, M.; Gautam, A.; Dutt, R.; Acharyya, A. CardioNet: Deep learning framework for prediction of CVD risk factors. In Proceedings of the 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Seville, Spain, 12–14 October 2020; pp. 1–5.
25. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classi-fication with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]

26. Zeiler, M.D.; Fergus, R. Visualizing and understanding convolutional networks. In *Lecture Notes in Computer Science*; 8689 LNCS; Springer: Cham, Switzerland, 2014; pp. 818–833.

27. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Li, F.-F. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.

28. Li, Y.-C.; Shen, T.-Y.; Chen, C.-C.; Chang, W.-T.; Lee, P.-Y.; Huang, C.-C. Automatic detection of atherosclerotic plaque and calcification from intravascular ultrasound Images by using deep convolutional neural networks. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2021**, *68*, 1762–1772. [CrossRef] [PubMed]

29. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural network. *Pattern Recognit.* **2018**, *77*, 354–377. [CrossRef]

30. Kiranyaz, S.; Avci, O.; Abdeljaber, O.; Ince, T.; Gabbouj, M.; Inman, D.J. 1D convolutional neural networks and applications: A survey. *Mech. Syst. Signal Process.* **2021**, *151*, 107s398. [CrossRef]

31. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference for Learning Representations, San Diego, CA, USA, 7–9 May 2015.

32. Ma, J.; Yarats, D. Quasi-hyperbolic momentum and Adam for deep learning. In Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019. [CrossRef]

33. Pilot European Image Processing Archive. The Mini-MIAS Database of Mammograms. 2012. Available online: http://peipa.essex.ac.uk/pix/mias/ (accessed on 1 January 2022).

34. Mammographic Image Analysis Society (MIAS). Database v1.21. 2019. Available online: https://www.repository.cam.ac.uk/handle/1810/250394 (accessed on 1 January 2022).

35. Wu, J.-X.; Chen, P.-Y.; Li, C.-M.; Kuo, Y.-C.; Pai, N.-S.; Lin, C.-H. Multilayer fractional-order machine vision classifier for rapid typical lung diseases screening on digital chest X-ray images. *IEEE Access* **2020**, *8*, 105886–105902. [CrossRef]

36. Lin, C.-H.; Wu, J.-X.; Li, C.-M.; Chen, P.-Y.; Pai, N.-S.; Kuo, Y.-C. Enhancement of chest X-ray images to improve screening accuracy rate using iterated function system and multilayer fractional-order machine learning classifier. *IEEE Photonics J.* **2020**, *12*, 1–19. [CrossRef]

37. Pu, Y.-F.; Zhou, J.-L.; Yuan, X. Fractional differential mask: A fractional differential-based approach for multiscale texture enhancement. *IEEE Trans. Image Process.* **2010**, *19*, 491–511.

38. Zhang, Y.; Pu, Y.-F.; Zhou, J.-L. Construction of fractional differential masks based on Riemann-Liouville definition. *J. Comput. Inf. Syst.* **2010**, *6*, 3191–3199.

39. Thanh, D.N.H.; Kalavathi, P.; Thanh, L.T.; Prasath, V.B.S. Chest X-ray image denoising using Nesterov optimization method with total variation regularization. *Procedia Comput. Sci.* **2020**, *171*, 1961–1969. [CrossRef]

40. Ba, L.J.; Frey, B. Adaptive dropout for training deep neural networks. *Adv. Neural Inf. Process. Syst.* **2013**, *26*, 1–9.

41. Clevert, D.; Unterthiner, T.; Hochreite, S. Fast and accurate deep network learning by exponential linear units (ELUs). In Proceedings of the 4th International Conference on Learning Representations, San Juan, Puerto Rico, 2–4 May 2016.

42. Hendrycks, D.; Gimpel, K. Gaussian Error Linear Units (GELUs). *arXiv* **2016**, arXiv:1606.08415.

43. Boer, P.; Kroese, D.P.; Mannor, S.; Rubinstein, R.Y. A tutorial on the cross-entropy method. *Ann. Oper. Res.* **2005**, *134*, 19–67. [CrossRef]

44. Rubinstein, R.Y.; Kroese, D.P. *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning*; Information Science and Statistics; Springer: New York, NY, USA, 2004; pp. 1–47.

45. Ho, Y.; Wookey, S. The real-world-weight cross- entropy loss function: Modeling the costs of mislabeling. *IEEE Access* **2019**, *8*, 4806–4813. [CrossRef]

46. Chen, X.; Liu, S.; Sun, R.; Hong, M. On the convergence of a class of ADAM-type algorithms for non-convex optimization. *arXiv* **2018**, arXiv:1808.02941.

47. Syntax: Edge, 1994–2021 Years. Available online: https://www.mathworks.com/help/images/ref/edge.html (accessed on 1 January 2022).

48. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, *21*, 6. [CrossRef]

49. Syntax: Histeq, 1994–2021 Years. Available online: https://www.mathworks.com/help/images/ref/histeq.html (accessed on 1 January 2022).

50. Syntax: Imadjust, 1994–2021 Years. Available online: https://www.mathworks.com/help/images/ref/imadjust.html (accessed on 1 January 2022).

51. Wu, J.-X.; Liu, H.-C.; Chen, P.-Y.; Lin, C.-H.; Chou, Y.-H.; Shung, K.K. Enhancement of ARFI-VTI elastography images in order to preliminary rapid screening of benign and malignant breast tumors using multilayer fractional-order machine vision classifier. *IEEE Access* **2020**, *8*, 164222–164237. [CrossRef]

52. Valenzuela, G.; Laimes, R.; Chavez, I.; Salazar, C.; Bellido, E.G.; Tirado, I.; Pinto, J.; Guerrero, J.; Lavarello, R.J. In vivo diagnosis of metastasis in cervical lymph nodes using backscatter coefficient. In Proceedings of the 2018 IEEE International Ultrasonics Symposium (IUS), Kobe, Japan, 22–25 October 2018.

53. Chansong, D.; Supratid, S. Impacts of Kernel size on different resized images in object recognition based on convolutional neural network. In Proceedings of the 2021 9th International Electrical Engineering Congress (iEECON), Pattaya, Thailand, 10–12 March 2021.

54. Sidek, K.A.; Khalil, I.; Jelinek, H.F. ECG biometric with abnormal cardiac conditions in remote monitoring system. *IEEE Trans. Syst. Man Cybern. Syst.* **2014**, *44*, 1498–1509. [CrossRef]

55. Zhang, X.-H. A Convolutional Neural Network Assisted Fast Tumor Screening System Based on Fractional-Order Image Enhancement: The Case of Breast X-ray Medical Imaging. Master's Thesis, Department of Electrical Engineering, National Chin-Yi University of Technology, Taichung City, Taiwan, July 2021.

56. Bruno, A.; Ardizzone, E.; Vitabile, S.; Midiri, M. A novel solution based on scale invariant feature transform descriptors and deep learning for the detection of suspicious regions in mammogram images. *J. Med. Signals Sens.* **2020**, *10*, 158–173.