A deep learning model using data augmentation for detection of architectural distortion in whole and patches of images

Olaide N. Oyelade *, Absalom E. Ezugwu

School of Mathematics, Statistics, and Computer Science, University of KwaZulu-Natal, King Edward Avenue, Pietermaritzburg Campus, Pietermaritzburg, KwaZulu-Natal, 3201, South Africa

ARTICLE INFO

ABSTRACT

Breast cancer is now widely known to be the second most lethal disease among women. Computer-aided detection (CAD) systems, deep learning (DL) in particular, have continued to provide significant computational solution in early detection and diagnosis of this disease. Research efforts are advancing novel approaches to improve the performance of DL-based models. Techniques such as data augmentation, varying depth of model, image quality enhancement, and choice of classifier have been proposed to improve performance in the characterization of abnormalities in mammograms. However, no significant progress has been made in applying deep learning techniques to the detection of architectural distortion – a form of abnormalities in breast images. In this research, we propose a novel convolution neural network (CNN) model for the detection of architectural distortion by enhancing its performance using data augmentation technique. We also investigate the performance of the proposed model on different operations of image augmentation. Furthermore, the new model was adapted to detect images presenting the right and left breast presented in MLO and CC views. Similarly, we investigate the performance of our model under the fixed-size region of interests (ROIs) and multi-size whole images inputs. Our method was trained on 5136 ROIs from MIAS, 410 whole images from INbreast, 322 whole images from MIAS, and 55,890 ROIs from DDSM + CBS databases. Performance evaluation of the proposed model in comparison with other state-of-the-art techniques revealed that the model achieved 93.75 % accuracy. This study has, therefore, strengthened the need to leverage data augmentation techniques to enhance the detection of architectural distortion, thereby reducing the rate of advanced cases of breast cancer.

1. Introduction

Cancer is the uncontrolled growth and spread of cells. Breast cancer is a type of cancer that has risen to be the second cause of death among women. A fundamental characteristic of all forms of cancer is that the earlier they are detected and attended to, the easier they are able to be cured. This is because the growth rate of the affected cells can be exponential [1]. The Cancer Health Center (CHC) noted that most cases of cancer are detected and diagnosed after a tumor can be felt or when other symptoms have developed [2]. Breast cancer has the second highest mortality rate in women next to lung cancer and is the most common type of cancer in 140 countries of a total of 182 evaluated nations [3]. The US prediction on breast cancer towards 2019 revealed that about 268,600 new cases of invasive breast cancer would be diagnosed, 62,930 new cases of carcinoma in situ will be diagnosed, and 41,760 women will die from breast cancer [4]. Although these figures may

appear to mirror what is obtainable in most developed economies, research has also shown that almost 50 % of breast cancer cases and 58 % of deaths occur in less developed countries. The low survival rates can be explained mainly by the lack of early detection of the disease with over 33 % and 81 % of the population in ages 30–49 years, and 30–59 years, respectively [5–7]. Recently, the American Cancer Society estimated that in 2019 [78], 30 % (14,460) of women diagnosed of ductal carcinoma in situ (DCIS), a type of breast cancer, were in the age gap of 60–69; 28 % (74,820) of those in that same age group progress to have the invasive cancer (an advanced stage of breast cancer); and 24 % (9920) eventually died. In a separate report, US alone was reported to have 1,806,590 new cancer cases and 606,520 cancer deaths were projected for the year 2020 [79], while [80] showed that the estimated rate of new cases is 128.5 per 100,000 and death rate is 20.1 per 100,000 women per year.

The challenge of early detection of breast cancer has promoted

---

research in computer-aided detection (CAD) systems. The high impact of CAD has increased the number of cancers detected by 19.5 %, increased early-stage malignancies detected from 73 % to 78 %, achieved a recall rate of 7.7 %, and detected 50 % of the cases of architectural distortion missed by human experts [8]. Although different approaches like rule-based systems, logic and prolog are some traditional CAD systems adopted for this purpose, however, some of these systems have not been successful at limiting the number of false diagnoses. This has, therefore, advanced research in the area of deep learning, particularly convolutional neural network (CNN) in drastically limiting false diagnosis, false positive rates and increased early detection. CNN and deep learning techniques generally are a type of CAD systems which are based on recognition of patterns while processing images to extract certain set(s) of features. We have directed the focus of this research at the use of deep learning methods for mammograms. This will help us detect if it is of malignant or benign or normal case based on the digital image. We exploit the presence of architectural distortion (AD) in mammography to carry out this differential task. This is necessitated due to the fact that 12–45 % of cancers missed in mammographic screening are AD [3]. Note that malignant cells are cancerous cells which start from abnormal cell growth and might spread rapidly or invade nearby tissue while benign cases are considered as noncancerous, these can be easily removed from the body [9].

Mammography is a type of medical imaging used for screening and diagnosis of breast cancer. It is the most useful and common tool employed by the radiologist when looking out for speculated masses, microcalcifications, bilateral asymmetry and architectural distortion [3, 10]. This implies that making findings from it largely depends on experts – the radiologists. However, screening carried out using this tool is highly characterized with false positive results, over-diagnosis of insufficient lesions, different interpretations of screening results, unreliable/low accuracy of detection and diagnosis, increasing need for carrying out additional examinations, and limitation of the radiologist who makes the findings, thereby leading to patient anxiety [11,12]. These notwithstanding, mammography has made significant contributions towards the early detection of breast cancer, even the detection of something as subtle as architectural distortion [10,13,14]. On the other hand, the technique of biopsy is another method applied to confirm the presence of breast cancer [81]. The study in [82] attempted to apply the combination of K-means, fuzzy C-means, competitive learning neural networks and Gaussian mixture models as clustering techniques in detection of breast cancer based on biopsy images. Architectural distortion is the alteration of the architecture of a normal breast. The presence of architectural distortion does not necessarily mean that it must present with mass or calcification. But this distortion in breast architecture usually manifests in mammographic finding as speculations radiating from a point, focal retraction, and straightening at the edges of the parenchyma [15,16]. Architectural distortion, due to its subtle nature and low prevalence, is, however, frequently discovered in retrospective assessments of false-negative mammography and may represent the earliest manifestation of breast cancer.

Deep learning models have demonstrated great results by improving the state-of-the-art in identifying subtle abnormalities like architectural distortion [69,70]. The application of deep learning technique has proven that using classical methods such as Grey Level Co-occurrence Matrix (GLCM) [83], multi-resolution wavelet [84], Gabor Filter [85–89], pattern recognition [90] for feature selection and classification in characterization of abnormalities in breast images are ineffective and deficient [91]. Hence the adoption of the use of deep learning models which are based on the deep arrangement of layers which are able to extract features in images (and other forms of data) with multiple levels of abstraction [17]. CNN, which is a type of deep learning model, is used in detecting architectural distortion in digital mammography, in addition to localization of regions of interest (ROIs), classification of findings, image retrieval, and risk assessment. The layers which often stack up to build a CNN model are convolutional layer, pooling layer and fully

connected (fc) layer [18]. Variation of hyperparameters (depth of model for example) in CNN has produced different CNN architectures namely: CiFarNet [19], AlexNet [20], GoogLeNet or Inception v1 [21], Inception v3 [22], Inception v4 [23], Xception [24], ResNeXt-50 [25], ResNet [26], VGG [27] and LeNet [28]. These novel architectures have attracted a wide range of acceptance and application among researchers largely due to their outstanding performances even in the detection of architectural distortion [29–32].

Generally speaking, studies have supported the need for advanced tools and techniques for accurate diagnosis and classification of breast cancer since the segmentation and classification phases are challenging [73,74], hence the motivation for the adoption and adaptation of deep learning techniques for these tasks. In addition to that, the major challenge of using CNN for detection of abnormality in mammography is limited data which usually lead to high false positive rates, a poorly trained CNN model that does not generalize and results in overfitting [33–35]. Secondly, another related challenge which may lead to high false positive rate is associated with low contrast in the datasets of mammogram images available for use in CNN models [36]. This calls for the furtherance of research to advance CNN beyond the limitations of radiologists and traditional CAD systems which are duped into having sensitivity ranging from 62 to 87 % and specificity from 75 to 91 % for human experts [37], and sensitivity of 50 % with the number of false positives per image equal to 1.0 for traditional CADs [35]. Thirdly, there are few research efforts geared towards the detection of architectural distortions in digital mammograms, and these are often missed or difficult to detect when screening mammograms [32,43]. The fourth research opportunity aimed at improving deep learning models for detection of ADs remains variation in sizes of ROIs. Although the first problem has been partly addressed by using transfer learning (TL) and data augmentation techniques, these solutions are yet to sufficiently harness for a significant reduction of false positive rates of detection [38] and improved detection of AD. Whereas some literature has argued in favor of the use of both the ground truth-size-based ROIs and fixed-sized ROIs, others have attempted to adapt CNN models to whole mammography images. These input size variations are still a hot topic of research.

The effectiveness of mammography as a tool for screening and diagnosing breast cancer is still relevant [39]. Furthermore, it is also well observed that deep learning models are capable of yielding unprecedented performance on some tasks, given sufficient data [40,35]. In addition to this, the standard and synthetic approaches for data augmentation have produced significant performance in curtailing overfitting. Although techniques like dropout, batch normalization, batch re-normalization or layer normalization have also contributed immensely to tackling overfitting of deep learning models, we still argue that the use of the synthetic approach for data augmentation is not yet sufficiently exploited.

To address the above challenges, we propose an improved computational solution to aid the accurate detection of breast cancer abnormality architectural distortion. This study employed the use of data augmentation [41] technique in generating new synthetic dataset to achieve a CNN model that effectively generalizes. This research investigated the effect of combining both fixed-sized and variable-sized images into the proposed CNN model. The model was adapted to be able to compare images from the left and right breasts and also the craniocaudal (CC) and mediolateral-oblique (MLO) view of each breast. Also, we tweaked the model to increase the performance of our proposed CNN architecture by generating augmented data to increase training data.

The main objective of this study is to design a CNN-based model for the detection of architectural distortion, and apply data augmentation technique to improve the accuracy in reducing false positive classification of architectural distortion. The novelty of our work lies in the proposed CNN model designed, implemented and trained from scratch, the combination of datasets resulting in a good number of experimentations, and the outstanding classification performance for architectural

distortion. Specifically, the technical contributions and novelty of the proposed model are hereby highlighted below:

- Design of a novel CNN-based model for detection of architectural distortion. The proposed model is able to accept multi-size inputs and images of different views.
- Application of a data augmentation approach based on the traditional method of flipping and rotating of images. We experimentally chose the best values for the required parameters (e.g. angle of rotation of image).
- Enhancement of the proposed CNN model to investigate the likelihood of obtaining performance improvement when whole images are served as input.

It is noteworthy to mention that the new deep learning model presented in this study was able to achieve a reduction of false positive rates due to the detection of architectural distortion in images in digital mammography. Similarly, we applied the proposed model and the augmentation technique to a good number of benchmarked datasets to validate the robustness of the model. Further, to validate the efficiency and practicality of the proposed model, it was applied to different augmentation operations. This was done under different experimentation to investigate what combination of operations presents the optimal performance of the proposed model. Finally, a performance evaluation analysis reported in the later section of this paper shows that the new CNN-based model achieved superior accuracy when compared to similar existing state-of-the-art CNN models.

The remaining part of this paper is organized in the following order: Section 2 is focused on reviewing related works in data augmentation techniques, image cropping approaches for the extraction of ROIs, and use of deep learning models in detecting abnormalities in mammography. Section 3 presents the technical contributions of this paper. In Section 4, we report the experimentation carried out based on the proposed model. Furthermore, Section 5 presents the results and discussion of the experimentation. Finally, we conclude this paper in Section 6.

2. Related works

This section presents a review of some related works that used data augmented techniques for training deep learning models in detecting abnormalities from digital mammography and other related areas. Characterization of defects in mammograms is usually categorized into four, namely: malignant mass detection, calcification detection, architectural distortion, and asymmetry of the breast. Our review revealed that computer aided detection systems (CADs) using deep learning techniques have achieved outstanding performances not only in breast cancer but also brain, stomach, gastrointestinal, and lung cancers. This claim is supported by some very interesting studies [71,72,75–77]; we, however, constrained our review to those of breast cancer with the hope of learning from their approach. Although this paper is focused on the characterization of the architectural distortion abnormalities in the breast, we shall not restrict our review to that aspect alone, but instead briefly present the outcome of studies in others.

In [42], the authors approached their task of detection of architectural distortion and speculated masses using Gabor filters and planes. Experimentation was tested on Mini-MIAS and DDSM, and they applied SVM and MLP classifiers for classification. Results showed that they achieved 90 % of sensitivity, 86 % specificity in distinguishing AD from the healthy breast tissue and 93 % sensitivity and 88 % specificity in classifying speculated mass; also, SVM classifiers achieved 96 % sensitivity with 9.6 false positives per image in detection of speculated mass and 97 % sensitivity with 6.6 false positives per image while detecting architectural distortion. In related work [43], also demonstrated the methods for the detection of architectural distortion in prior mammograms of interval-cancer cases based on analysis of the orientation of breast tissue patterns in mammograms. They used the oriented

structures in a given mammogram which are analyzed using Gabor filters and phase portraits to detect node-like sites of radiating or intersecting tissue patterns and pattern classification via quadratic discriminant analysis. Results obtained achieved a sensitivity of 80 % at about five false positives per patient.

Others have leveraged the benefits of R-CNN as in [44] who introduced the detection of architectural distortion using a supervised pre-trained region-based network (R-CNN). Experimentation was based on DDSM dataset, and results showed that they obtained over 80 % sensitivity and specificity, and yielded 0.46 false-positives per image at 83 % true-positive rate. Similarly [45], demonstrated a novel network which combined two learning branches with region-level classification and region ranking in weakly and semi-supervised settings. Their results for weakly supervised learning showed an improvement of 4% in AUC, 10–17 % in partial AUC and 8–15 % in specificity, and 0.85 sensitivity. On the other hand [46], GlimpseNet autonomously extracts multiple regions of interest, classifies them, and then pools them to obtain a diagnosis for the full image. They obtained a result that gained 4.1 %.

Recently, there has being a surge in the use of basic CNN models in the characterization of architectural distortion from mammograms. [47] proposed a framework using a combination of deep Convolutional Neural Network (CNN) models. The model is an eight layer deep learning network that involves three pairs of convolution-max-pooling layers for automatic feature extraction and a multiple layer perceptron (MLP) classifier for feature categorization to process ROIs. The network contained 20, 10, and 5 feature maps of convolution layers. The MLP classifier is composed of one hidden layer and one logistic regression layer. Results of their experimentation achieved an AUC of 0.696 ± 0.044, 0.802 ± 0.037, 0.836 ± 0.036, and 0.822 ± 0.035 for fold 1–4 testing datasets, respectively, while the overall AUC of the entire dataset is 0.790 ± 0.019. Similarly [48], also proposed an in-depth feature-based framework combining intensity information for breast mass classification task. In related work, Bakkouri and Afdel [49] suggested a novel discriminative objective for supervised feature deep learning approach focused on the classification of tumors in mammograms as malignant or benign, using SoftMax layer as a classifier. The proposed network was enhanced with a scaling process based on Gaussian pyramids for obtaining regions of interest with normalized size. The DDSM and BCDR dataset were used in addition to data augmentation (geometric transformation) technique. The result of their experiments showed that they obtained an accuracy of 97.28 %.

Another deep learning model was used by [50], which was a novel supervised deep learning-based framework for region classification into semantically coherent tissues. Their work improvised data for training by training the CNN in an overlapping patch-wise manner and adapting the convolutional neural network (CNN) to learn discriminative features automatically. The experimental result showed that they obtained an average dice coefficients of 0.71. In [51] a multi-task transfer learning DCNN was proposed to translate knowledge from non-medical images to medical diagnostic tasks through supervised multi-task transfer learning, digitized screen-film mammograms (SFMs) and digital mammograms (DMs) were used to train the DCNN, which was then tested on SFMs. Experimentation was done with Institutional Review Board (IRB) approval, SFMs and DMs were collected from our patient files, and additional SFMs were obtained from the Digital Database for Screening Mammography. The data set consisted of 2242 views with 2454 masses (1057 malignant, 1397 benign).

Mammogram-based CNN based models include [52] which exploited the efficiency of pre-trained convolutional neural networks (CNNs) in a combination of pre-existing handcrafted features. These features were combined with low-level to mid-level features using a pre-trained CNN. The dynamic contrast enhanced-MRI [690 cases], full-field digital mammography [245 cases], and ultrasound datasets [1125 cases] were used for experimentation. The research obtained the following results: DCE-MRI [AUC = 0.89, se = 0.01], FFDM [AUC = 0.86, se = 0.01)], and ultrasound [AUC = 0.90, se = 0.01)]). Another use of a CNN model for

classification of breast mass lesions and aided with end-to-end learning process was proposed by [53].

In [54], the authors presented a novel classification technique for a large dataset of mammograms using deep learning: convolutional neural network-discrete wavelet (CNN-DW) and convolutional neural network-curvelet transform (CNN-CT). An augmented dataset is generated by using mammogram patches, in addition to filtering the data using contrast limited adaptive histogram equalization (CLAHE) while using SoftMax layer and support vector machine (SVM) layer as a classifier. Results showed that CNN-DW and CNN-CT had achieved an accuracy rate of 81.83 % and 83.74 %, respectively. In related work, the authors in [55] also explored the possibility of combining the technique of transfer learning with GoogLeNet and AlexNet pre-trained on a large-scale visual database. Results of their research demonstrated that GoogLeNet reached an AUC of 0.88 outperforming AlexNet, which stood at AUC of 0.8.

Finally, here are some other related works which adopted other similar techniques: [56] applied Convolution Neural Network on the mammogram images to enhance the accuracy rate of CAD. Performance of the different classifiers was measured on receiver operating characteristic. Experimentation results showed that the model attained an accuracy of 73 %, with 71.5 % sensitivity and 73.5 % specificity for dense tissue, and accuracy of 79.23 %, 73.25 % sensitivity and 74.5 % specificity was achieved for fatty tissue. Similarly [37], presented two novel techniques - genetic search of image enhancement methods with CLAHE and DCNN - to address inherent challenges in the application of machine learning to the domain of mammography. The research also utilized dual deep convolutional neural networks at different scales for classification of full mammogram images and derivative patches combined with a random forest gating. The result obtained showed a specificity of 0.91 and a specificity of 0.80.

The study proposed in [37] was based on wavelet convolution neural network for the detection of speculated findings in low-contrast noisy mammograms, such as architectural distortions and speculated masses. The dataset used for experimentation consisted of CBIS-DDSM and reached an accuracy of over 85 % for architectural distortions and - 88 % for speculated masses. In [31], the authors proposed a detection scheme composed of two separate channels, each of them being dedicated to the detection of one of the target radiological signs for detection of masses and architectural distortions in DBT datasets. Lastly [32], employed the use of texture models using support vector machine (SVM) classifier for texture classification of architectural distortion. The databases used were IRMA version of a digital database for screening mammogram (DDSM) and Mammographic Image Analysis Society (MIAS). Results pertain to an accuracy of 92.94 % obtained in case of DDSM database for fixed-size ROIs, and for MIAS database, an accuracy of 95.34 %. Similar to the work in [44], the authors in [65] adapted Fast R-CNN to detect and classify malignant or benign lesions on a mammogram using the INbreast dataset.

In [66], the authors designed a novel CNN architecture that can accurately detect breast cancer on screening mammograms using an "end-to-end" training procedure. Their model was able to detect cancer from incomplete annotated datasets. They required the image annotation for the initial stage of the training and then allowed their model to fall back to image-level labels in the second stage. Using the CBIS-DDSM dataset, their proposed model attained an AUC of 0.88, and four-model averaging improved the AUC to 0.91 (sensitivity: 86.1 %, specificity: 80.1 %). The authors also demonstrated that a whole image classifier trained using the proposed end-to-end approach on the dataset can be transferred to another dataset (INbreast) [66]. Meanwhile, in [67], the authors approached their study using the Shallow-Deep Convolutional Neural Network (SD-CNN) which relies on a shallow CNN to derive "virtual" recombined images from low-energy (LE) images; features are then extracted from the LE images using a CNN. The study revealed an accuracy of 0.90 using 49 contrasts enhanced digital mammography cases privately acquired. Similarly, they applied the proposed SD-CNN

Table 1

A summary of some recent state-of-the-art CNN-based studies aimed at characterizing abnormalities in breast images.

Authors and Reference	Year	Data Sources (Number of images)	Technique/ classifier	Classes	Classification accuracy/ performance
Jadoon et al. [63]	2017	IRMA (2796 images)	CNN-DW and CNN-CT with augmented data set /SVM and Softmax classifiers	2 and 3	81.83 % and 83.74 % accuracy
Qiu et al., [47]	2017	Private (560 images)	CNN/MLP/ Logistic regression classifier	2	AUC = 0.790
Ribli et al., [65]	2018	INbreast (115 FFDM cases images) Private dataset (847vFFDM images), and DDSM (2620 images)	Faster R-CNN	3	AUC = 0.95
Shen et al., [66]	2019	CBIS-DDSM (2478) and INbreast (115)	CNN/MLP/ Softmax	5	AUC = 0.98, sensitivity: 86.7 %, specificity: 96.1 %

on 69 digital mammography cases collected from the hospital located at Zhejiang University, China and obtained an accuracy of 0.95. The application of transfer learning to the process of detection of breast cancer from breast density scoring was demonstrated in [68]. The authors proposed a CNN model which uses Gabor filters and local binary pattern (LBP)] and gradient-based features [histogram of oriented gradients (HOG) as well as speeded-up robust features (SURF). After that, the adapted transfer learning approach with ImageNet trained weights to obtain an AUC of 87.3 %.

An overall perspective view of the state-of-the-art applying CNN models to the task of characterization of abnormalities in breast images revealed that most of the architectures adopted either SoftMax or SVM for classification. The architectural evolution of CNN ranges among architectural innovation (variation of depth and width of the model, number of channels, spatial or temporal exploitation, input size, and multi-layered architecture), skillfully selecting parameter optimization approach, use of different activation and loss functions, application of regularization, regulatory units (dropouts and batch normalization), and data augmentation. Performance enhancement of CNN model, therefore, largely depends on how carefully/skillfully researchers can arrange these components. Most data augmentation approaches used in classification model adopt that of traditional data augmentation techniques. This augmented data set is generated by using patches. The data processing technique widely used in similar studies involves data cleaning, feature analysis, data normalization and sampling, the addition of filtering, by contrast, limited adaptive histogram equalization (CLAHE), and data ordering. Table 1 summarizes all the pros and cons of the studies reviewed. Meanwhile, we observed that some studies are already taking advantage of the 3D tomosynthesis of breast images over the 2D mammograms widely used [64].

3. The proposed architectural distortion detection model and algorithm

In this section, we present the data preprocessing and cropping techniques we adopted and also show the design of the proposed CNN model. Meanwhile, we first attempt to give an overview of the complete
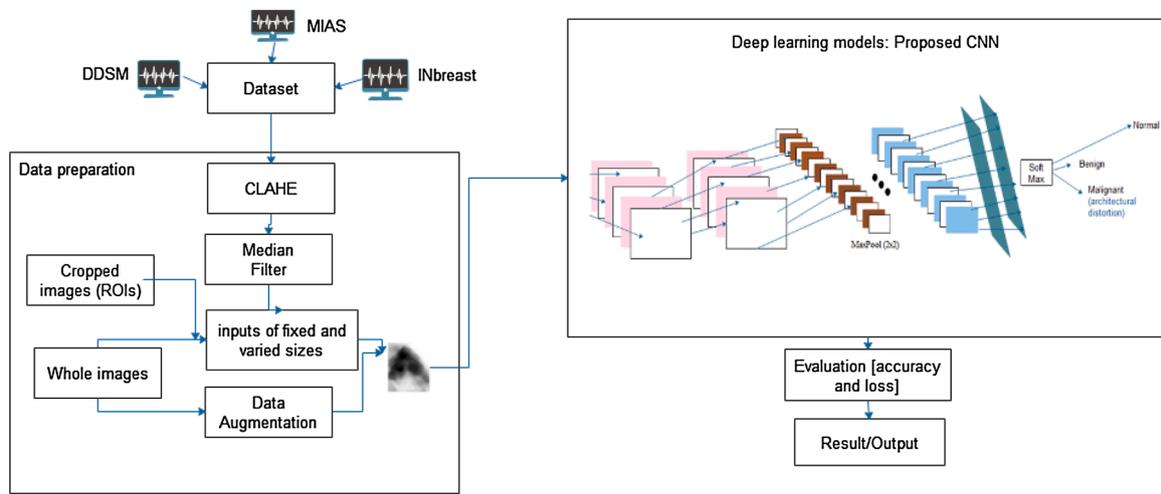
Fig. 1. Block diagram of the proposed CADx system which consists of image/data preprocessing task, traditional data augmentation, and the CNN model (CAPITAL I IN INPUTS).

Table 2
Description of popular benchmarked datasets used for experimentation.

Database	No. of Patients	No. of Images	Cases of abnormalities	Description
MIAS	161	322 (MLO view of images)	All forms of abnormalities (32 show architectural distortion)	Digitised to 50-micron pixel edge, and reduced to 200-micron pixel edge and padded/clipped so that all the images are 1024×1024. Images include radiologist's truth-markings. The database has some associated patient information (like age at time of study) and image information (like spatial resolution). Images are marked with ground truth information about the locations and types of suspicious regions.
DDSM	2620	10,480 (MLO and CC view of images)	All forms of abnormalities (about 137 show architectural distortion)	
INbreast	115	410	All kinds of abnormalities (architectural distortion approximately 8%)	Image matrix is 3328×4048 or 2560×3328 pixels.
Total		11,212		

Table 3
Description of datasets used for experimentation.

	Dataset	Total no. of samples/ ROIs	No. of samples/ ROIs with abnormalities
Benchmark Dataset	DDSM + CBIS-DDSM	The dataset contains 55,890 of which 14 % are positive and the remaining 86 % negative	7824
	MIAS	5136 ROIs	536
	MIAS whole images	322 whole images	115
	INbreast	410 whole images	349

c Data/image generation using the traditional data augmentation technique.

d Apply some forms of standard data augmentation technique on a batch of the acquired datasets.

e Split all the datasets into training, validation and testing sets.

f Apply our proposed deep learning model (in sub-section 3.5) on the datasets.

g Evaluate the performance of the two learning models

In Fig. 1, the illustration of the overview of the approach used in this paper is presented, as outlined above.

3.2. Mammography dataset

In CADx-based CNN models which use digital mammograms for input, most of the databases that are publicly available and widely used are the Mammographic Image Analysis Society (MIAS) database [57] and the Digital Database for Screening Mammography (DDSM) [10]. Similar datasets are the INbreast database, Breast Cancer Digital Repository (BCDR), and Image Retrieval in Medical Applications (IRMA). In this paper, we combine the datasets in the DDSM, MIAS and INbreast databases for experimental purposes. The INbreast database has 115 cases that resulted in 410 images (90 cases with both breasts - 4 images per-person affected, and another 25 cases of mastectomy - 2 images per-person) which consist of abnormalities such as architectural distortion [58]. Meanwhile, MIAS database has 332 images collected from 161 different cases, and DDSM has its images extracted from 2620 cases, each case having two images, resulting in 10,480 images. Table 2 presents the details of the datasets used in this research. We decided to acquire FFDM and FSM images to enhance the detection rate and to
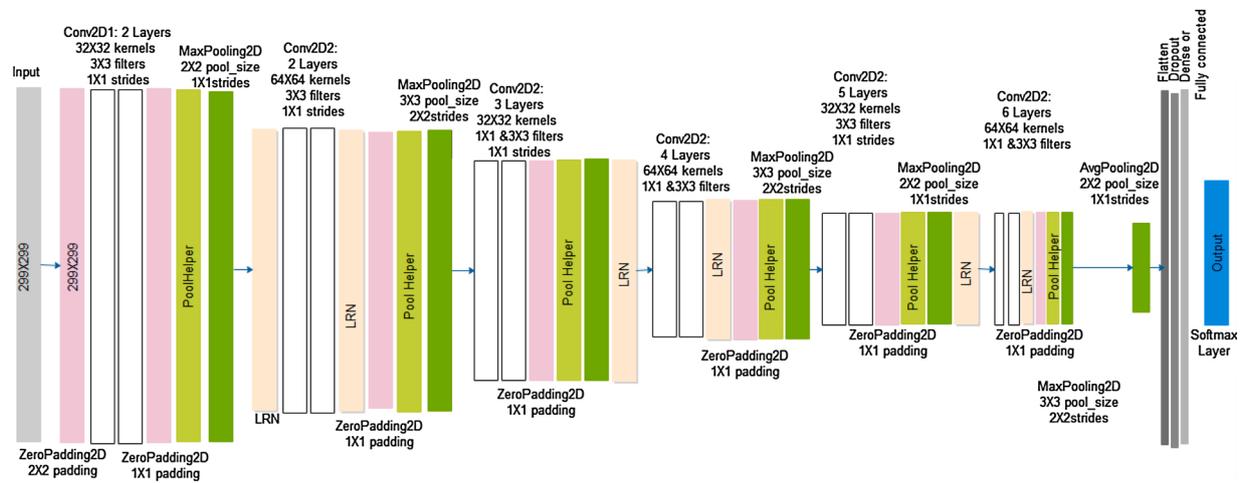
learning model proposed in this study.

3.1. Overview of methodology

The following are the procedures which outline the overview of our approach:

a Images are collected from selected databases (MIAS, INbreast and DDSM).

b Image preprocessing/data preparation and cleanings using CLAHE. To enhance the internal structures, Contrast-Limited Adaptive Histogram Equalization (CLAHE) was applied, followed by a mean filter of kernel 3×3.

Fig. 2. The proposed CNN architecture showing the input, blocks of convolution, max-pooling, fully connected layers and a three-class classifier: Layered-based architecture of the proposed CNN model.

lower the number of false detections [14].

In Table 3, we give more detailed information about the datasets used for the experimentation. We concentrate on the details of the number of ROIs extracted both manually and using an automated method. Also, we give information about the statistics of the occurrence of the four different forms of abnormalities observable in mammograms. Meanwhile, our dataset information will be categorized into those obtained from benchmarked databases and the synthesized datasets. The DDSM and CBIS-DDSM datasets have already been preprocessed and converted to 299×299 images by extracting the ROIs. The MIAS datasets were also preprocessed into its ROI forms of sizes 299×299. We experimented with the INbreast dataset [58] using the whole images

of sizes 2560×3328, and whole MIAS images of sizes 1024×1024. We first preprocessed the images by converting them from the DICOM formats into PNG files which yielded 420 images. Table 3 summarizes the description of samples/ROIs from each of the databases and even the synthesized one. Our MIAS and DDSM + CBS ROI datasets were preprocessed into Numpy files and TFrecords. The Numpy files are much more convenient to use in training than CSVfiles. The whole images (MIAS and INbreast) are in PNG forms.

3.3. Image preprocessing

Image preprocessing stage of this research implies the removal of

Table 4
A detailed representation of each layer of the proposed CNN architecture showing the following: input, output, filter size and number, activation, weights, and number of parameters. Note that convo represents a short form of convolutional layer.

Layer	Input (w x h)	Output (w x h x k)	Filter Size (n x m)	No. filters (k)	Memory	Weights (n*m)*k	Activation
Input layer	220×220	$220 \times 220 \times 1$	–	–	$48,400 \sim 48K$	–	–
Zero Padding	Padding = (3,3)						
Convo1_1	220×220	$224 \times 224 \times 32$	3×3	32	$1,548,800 \sim 1.5M$	320	Relu
Convo1_2	224×224	$224 \times 224 \times 32$	3×3	32	$1,605,632 \sim 1.6M$	9248	Relu
Zero Padding	Padding = (1,1)						
Max Pooling	Pool size = (2,2), Strides = (1,1)						
Convo2_1	223×223	$223 \times 223 \times 64$	1×1	64	$3,182,656 \sim 3.1M$	2112	Relu
Convolutional layer 2_2	223×223	$223 \times 223 \times 64$	3×3	64	$3,182,656 \sim 3.1M$	36,864	Relu
Zero Padding	Padding = (1,1)						
Max Pooling	Pool size = (3,3), Strides = (2,2)						
Convolutional layer 3_1	111×111	$111 \times 111 \times 128$	3×3	128	$1,577,088 \sim 1.5M$	73,856	Relu
Convolutional layer 3_2	111×111	$111 \times 111 \times 128$	3×3	128	$1,577,088 \sim 1.5M$	147,584	Relu
Zero Padding	Padding = (1,1)						
Max Pooling	Pool size = (2,2), Strides = (1,1)						
Convolutional layer 4_1	110×110	$110 \times 110 \times 256$	1×1	256	$6,364,600 \sim 6.3M$	33,024	Relu
Convolutional layer 4_2	110×110	$110 \times 110 \times 256$	3×3	256	$6,364,600 \sim 6.3M$	590,080	Relu
Zero Padding	Padding = (1,1)						
Max Pooling	Pool size = (3,3), Strides = (2,2)						
Convolutional layer 5_1	54×54	$52 \times 52 \times 512$	3×3	512	$1,492,992 \sim 1.4M$	1,180,160	Relu
Convolutional layer 5_2	52×52	$52 \times 52 \times 512$	3×3	512	$1,492,992 \sim 1.4M$	2,359,808	Relu
Zero Padding	Padding = (1,1)						
Max Pooling	Pool size = (2,2), Strides = (1,1)						
Convolutional layer 6_1	51×51	$51 \times 51 \times 1024$	1×1	1024	$2,663,424 \sim 2.6M$	525,312	Relu
Convolutional layer 6_2	51×51	$51 \times 51 \times 1024$	3×3	1024	$2,663,424 \sim 2.6M$	9,438,208	Relu
Zero Padding	Padding = (1,1)						
Max Pooling	Pool size = (3,3), Strides = (2,2), output = 25×25						
Avg. Pooling	Pool size = (2,2), Strides = (1,1), output = 24×24						
Flatten	576 single column						
Dropout layer	Rate = 0.5						
Dense layer	Number_of_classes = 24, L2(0.0002)						
SoftMax (output layer	24						

```
Input:  dataset ≔ MIAS, DDSM, or INbreast datasets

Output: AD Classified ROIs/whole images

1  Start
2      preprocessedDataset ←applyCLAHE(dataset) for image preprocessing
3      benchmarkDatas← applyMedianFilter(preprocessedDataset)
4      synthesizedDatasets ←  DataSynthesizemodel(benchmarkDatas)
5      allDatasets ←  synthesizedDatasets+ benchmarkDatasets
6
7      TrainingSet ← extractTrainingSet(allDataset)
8      ValidationSet ← extractValidationSet(allDataset)
9      TestingSet ← extractTestingSet(allDataset)
10     proposedCNN(TrainingSet, ValidationSet, TestingSet )
11     model ← ComputeCNN(cnn, parameters, hyperparameters)
12     display classification result
13 end
```

noise, breast image contrast enhancement, and image breast segmentation to remove background area, labels, artefacts, and pectoral muscle. This paper uses a variant of adaptive histogram equalization (AHE), called contrast limited adaptive histogram equalization (CLAHE), to improve the contrast in images. In addition, we use a median filter for denoising, and un-sharp mask to smoothen the images.

3.4. Cropping of images

We obtained cropped images containing ROIs of size 299×299 pixels. The ROIs used by our method were of fixed size (ROIs of size 299×299 pixels were selected both from MIAS and DDSM). Meanwhile, we investigated the possibility of feeding in a multi-sized image into our model. To accomplish this task, we used images of size 1024×1024 pixels for MIAS and of size 2560×3328 pixels for INbreast, although the result was not encouraging.

3.5. The proposed deep learning architecture

In CNN, the design of the network architecture largely depends on appropriate choice of model parameters/hyperparameters and conditions like the requirement of the model, the size of the dataset, depth of the architecture, more layers (extract more features, increases accuracy when sufficient training datasets exists), and considerable size of training data. In addition, Minavathi et al. [60] observed that, with deep learning models, using step decay rate while reducing learning rate by some percentage after a set number of training epochs increases the performance of characterization of abnormalities in mammograms.

We first served ROIs images into the CNN model, and after that experimented with whole mammograms. Our first attempt was aimed at reducing the computational time required for extracting features and also to eliminate loss resulting from the down-sampling whole mammogram. We chose a 3×3 filter size because smaller filters collect as much local information as possible; bigger filters represent more global, high-level information. Our decision on this was reinforced by the common knowledge of the subtle nature in detecting architectural distortion from digital mammograms.

The architecture described in Fig. 2 above assumes the form of Conv-Conv-Pool-Conv-Conv-Pool with a number of filters modeled as 32(3, relu)–32(3, relu)-2(2)-64(3, relu)–64(3, relu) and so on. The main model is captured in Fig. 2.

We provide a detailed representation of the proposed CNN model shown in Fig. 2 by summarizing all important parameters in Table 4 below. The proposed architecture applied the ***strides*** of 1; ***bias*** of 1 and the value ***same*** for the ***padding*** parameters across convolutional layers. Also, the proposed architecture used kernel regularizer (L2) with a value of $2*10{-}4$ across convolutional layers. We computed the number of parameters for the convolutional layers using Eq. 1:

$$convoParams = (((n * m) * stride + 1) * filters) \tag{1}$$

where n*m = kernel size. Similarly, we compute the output of each layer using Eq. 2:

$$layerOutput = \left(\left(\frac{(w - n + 2P)}{s} + 1 \right) * \left(\frac{(h - m + 2P)}{s} + 1 \right) \right) * k \tag{2}$$

where w and h stand for the width and height of the input sample image respectively, s stands for the strides, P denotes the zero padding value, and k stands for number of filters. The computation of the weights for each layer is also achieved using Eq. 3:

$$weights = (w * h) * c + b \tag{3}$$

where c denotes the number of image depth or channels or depth, and b is bias. The output of our pooling layers is derived using Eq. 4 which multiplies to 1 (depth) because of the gray-scale images:

$$poolLayerOutput = \left(\left(\frac{(w - n)}{s} + 1 \right) * \left(\frac{(h - m)}{s} + 1 \right) \right) * 1 \tag{4}$$

The proposed CNN model consists of 12 convolutional layers alongside pooling and fully connected layers in their respective positions. The motivation for the selection of this architecture is first motivated from the Google's architecture, named GoogleNet. We adopted the concept of Conv-Conv-Pool-Drop from their architecture. The architecture of the model follows the form of Conv-Conv-Pool-Drop-Conv-Conv-BatchNorm-Pool-Drop-Dense(relu)–BatchNorm-Drop, with a number of filter modeled as 32(3, relu)–32(3, relu)-2(2)-64(3, relu)–64(3, relu) and we applied a normalization technique (layer), namely local response normalization (LRN), after each block of convolutional layer.

The optimization algorithm applied to the proposed model was the Adam optimizer, and the learning rate was 0.001. The proposed CNN model benefits from some deep learning regularization techniques which have demonstrated capacity to combat overfitting issue. Overfitting is the situation when a model learns the training data excellently but falls short of generalizing well when some other data is exposed to it. Regularization techniques such as L2 and L1, dropout, data augmentation, and early stopping have been widely reported to enhance the performance of deep learning models [92,93]. This study, therefore, experimented with some of the techniques to ensure an optimal performance of the proposed deep learning (CNN) model. Hence, we do not just hope to improve performance but also to enable our model to generalize well. A model failing to generalize well will show validation error increases while the training error steadily decreases. In this study, we applied our work to the most common regularization technique L2 which is also referred to as "weight decay". We aimed at applying this weight regularization technique to reduce overfitting. L2 values range between 0 and 0.1 with examples as 0.1, 0.001, 0.0001, and are in

Table 5
Data augmentation parameters.

Parameters	Description	Training dataset	Validation dataset	Testing dataset
zca_whitening	apply ZCA whitening	False	False	False
rotation_range	randomly rotate images in the range (degrees, 0–180)	20	20	10
width_shift_range	randomly shift images horizontally (fraction of total width)	0.2	0.	0.
shear_range	set range for random shear	0.15	0.	0.
channel_shift_range	set range for random channel shifts	0	0	0
height_shift_range	randomly shift images vertically (fraction of total height)	0.2	0.	0.
zca_epsilon	epsilon for ZCA whitening	1e-06	1e-06	1e-06
horizontal_flip	randomly flip images	True	True	True
vertical_flip	randomly flip images	True	True	True
Rescale	set rescaling factor (applied before any other transformation)	None	None	None
fill_mode	set mode for filling points outside the input boundaries	Nearest	nearest	nearest
zoom_range	fraction of images reserved for validation (strictly between 0 and 1)	0.15	0.	0.
data_format	# image data format, either "channels_first" or "channels_last"	channels_first	channels_first	channels_first
CLACHE parameters				
adaptive_equalization		True	True	True
contrast_stretching		True	True	True
histogram_equalization		True	True	True

logarithmic scale. We therefore hope to reduce our model's training error by applying this technique.

3.6. Algorithm of the combined model for data augmentation and classification

Algorithm 1 combines the procedures outlined in Fig. 1. The algorithm accepts one out of the datasets in Table 2; it then performs all necessary preprocessing tasks on the datasets. Some of these functions are: data preprocessing (as described in Section 3.3), generating additional images using data augmentation (as described in Sections 3.1 and 4.1 A), splitting of the dataset into three sets (training, evaluation and trusting), and finally applying the proposed CNN model on the dataset before classification is done using SoftMax classifier (multiclass classification applied here).

The complexity of Algorithm 1 evaluated to $O(n)$. This was derived from line 2 is $(O(n))$, line 3 is $(O(n))$, line 4 is $(O(n))$, lines 5, 7, 8, 9, and 12 each have 1, and lines 10 and 11 are approximated to have complexity of $O(n)$. Following the rules of summing (addition rule) each of them, we arrive at $O(n)$ for Algorithm 1.

4. Computational experiments

4.1. Experimentation setup

Experimentation of the proposed deep learning model was implemented using Keras with Google's Tensorflow as backend. We trained our model on Floyd server [61] using the following configurations: a High-Performance GPU with Tesla V100, 16 GB Memory, 61 GB RAM, and 100 GB SSD; also, we deployed our model for training on a Standard GPU with Tesla K80, 12 GB Memory, 61 GB RAM, and 100 GB SSD. Meanwhile, we attempted to observe the variations of performance when our model is trained using a CPU only-based system, hence, we likewise trained our model on a system with a Standard CPU, Intel Xeon 2 Cores, having 8 GB RAM 100 GB SSD; and also on another configuration of High-Performance CPU, Intel Xeon with 8 Cores, 32 GB RAM and 100 GB SSD.

4.2. Image augmentation and training configuration

A brief description of how images (particularly the highly insufficient images with architectural distortion) are augmented based on the current proposed model adaptation and the parameters settings for training the new model are elaborated in this section.

Table 6
Total samples/ROIs generated in each category of the dataset.

	Dataset	Total no. of samples/ROIs synthesized
Synthesized Dataset	DDSM + CBIS-DDSM	179,447 ROIs
	MIAS (ROIs)	49,724 ROIs
	MIAS whole image	7914 whole images
	INbreast	1688 whole images

4.2.1. Data augmentation

We adopted the standard or what is known as the traditional data augmentation procedure on all the categories of datasets used in the study. Table 5 summarizes the choice of features we applied to the original images to generate or synthesise new data for the training process. The augmentation parameters listed in Table 5 were applied to each of the datasets used in training the proposed model. The outcome of the application of the augmentation described in Table 5 is listed in Table 6.

In Fig. 3, we illustrate samples of images in their varicose categorization based on the dataset and their usage (training, validation and testing purposes). We generated augmented images for the category of our dataset using the parameters listed in Table 5. The outcome of our augmentation operation is shown in Fig. 5. Base on the data augmentation operation applied as detailed in Table 5.

4.2.1.1. MIAS dataset.
In Fig. 3a, we show some samples (ROIs) of the real image from the MIAS dataset of sizes 299 × 299 which we used for training, testing and validating our model.

4.2.1.2. DDSM + CBS dataset.
Samples of real images collected from the DDSM + CBS dataset for training and validation purposes were also applied to our model.

4.2.1.3. MIAS whole image (1024 × 1024) dataset.
We experimented our model with whole mammogram images to observe its performance. The MIAS whole images of sizes 1024 × 1024 were used, and their corresponding augmented versions generated. These are shown in Fig. 3b.

4.2.1.4. INbreast whole image (2560 × 3328) dataset.
This dataset was first converted to PNG formats from its original DICOM formats, then we produced augmented version before serving into our model as input.
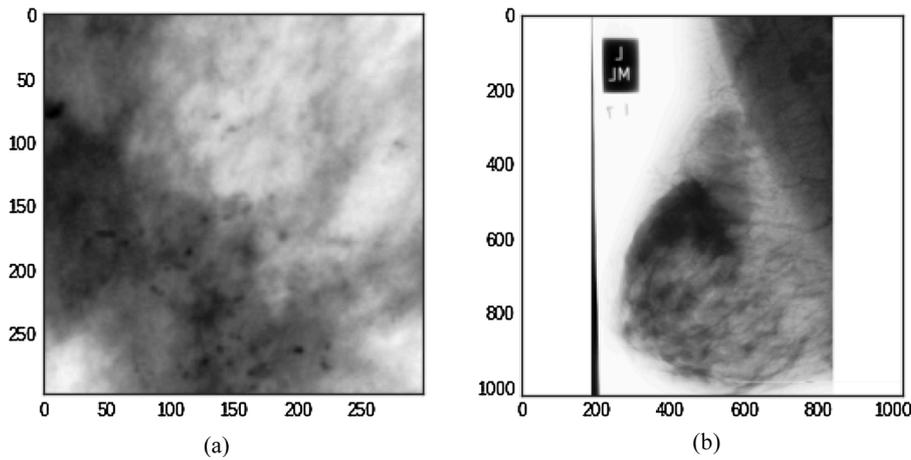
Fig. 3. (a): Samples of MIAS 299 × 299 real images used for training, training, and testing. (b): A sample of real MIAS(1024 × 1024) image with architectural distortion, and two samples of augmented MIAS(1024 × 1024) image.

Table 7
List of parameters and hyperparameters used during training.

Parameters & Hyperparameters		
Used under real and synthesized datasets	Kernels	Kernel size:3 × 3 number of kernels: 64 Activation function: ReLU Stride = 1 Padding = 1
	Weights	Number of weights, Activation function = SoftMax
	Callbacks	learning rate: 0.001
	Early stopping [monitor='val_loss', patience = 2]	beta_1 = 0.9,
	Checkpointing [monitor='val_loss', mode='auto', period = 1]	beta_2 = 0.999,
		epsilon = 1e-8
	Optimerzer: Adam	loss = categorical_crossentropy

4.2.2. Training parameters

Table 7 summarizes the parameters and hyperparameters used in training our model on the different datasets presented earlier.

5. Results and discussion

In this section, we present the results of our experiment and also compare the performance of the proposed CNN model with similar existing models in the literature.

5.1. MIAS dataset (299 × 299)

The loss value and accuracy obtained while training our model with the MIAS dataset revealed that the model performs well in detecting architectural distortion in breast patches presented to it. Training the proposed model without applying augmentation technique resulted in an accuracy of 84.30 % as the loss continued to drop progressively during training. Fig. 4 present the pattern of change in accuracy and loss during training and evaluation.

The results in Fig. 4 above show that as the loss value decreases, our model achieves an incremental inaccuracy. Applying the proposed model to data augmentation technique operations (horizontal and vertical flip, shift range, shear range and zoom range) during training and evaluation as shown in Fig. 5 produced a better result compared to the earlier approach. The result confirms the argument that performance improvement is attainable when applying data augmentation to the deep learning model. The result shows that we obtained an accuracy of 93.75 % at the loss value of 0.29.

Although our model performed well when applied to the MIAS datasets, however a further improvement of the current model using image augmentation technique yielded some performance gain. This assertion is drawn from results illustrated in Fig. 5. Furthermore, to investigate the performance of the proposed model on different data augmentation operations, we varied the process by adding and removing operations. This allowed the study to present an optimal combination of operations necessary for applying the augmentation method to CNN models. This is informed from the understanding that our proposed model tends to learn features that are unaffected by their positioning in the input, hence the liberty to transform images into
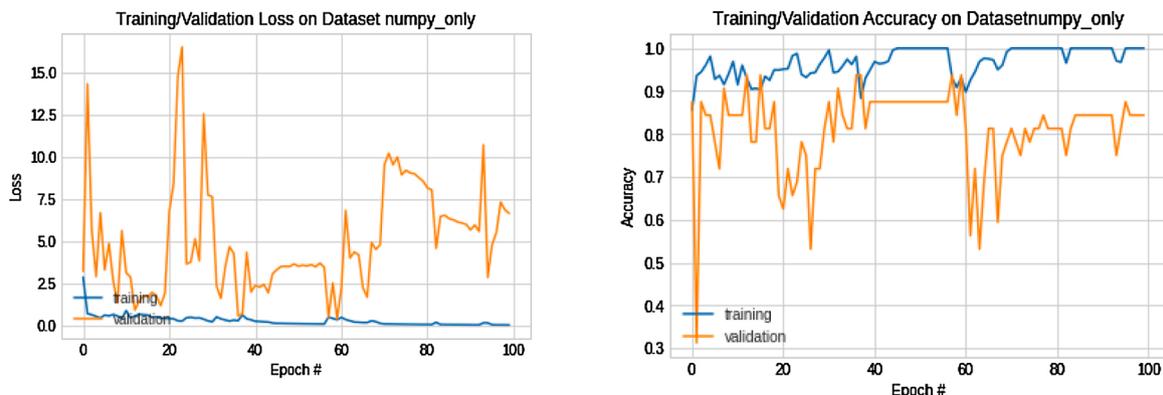

Fig. 4. Pattern of change in loss function and accuracy of training and evaluation MIAS dataset.
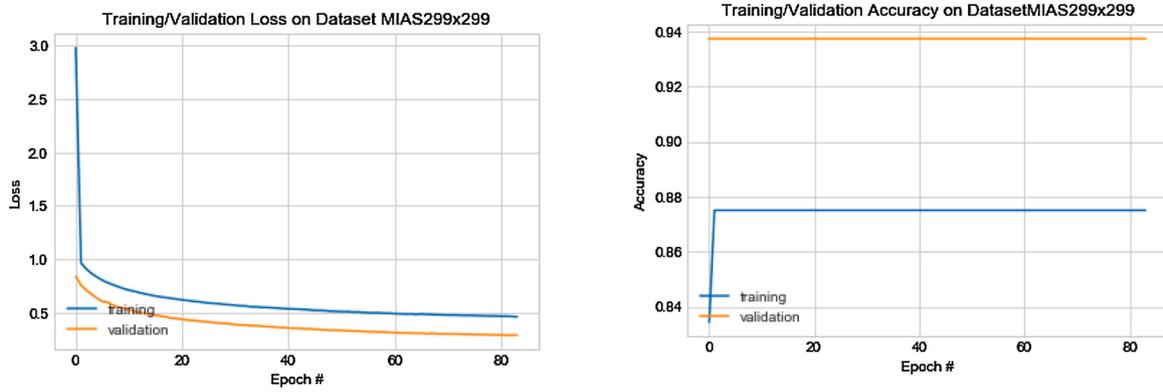
Fig. 5. Pattern of change in loss function and accuracy of training and evaluation MIAS dataset with data augmentation.
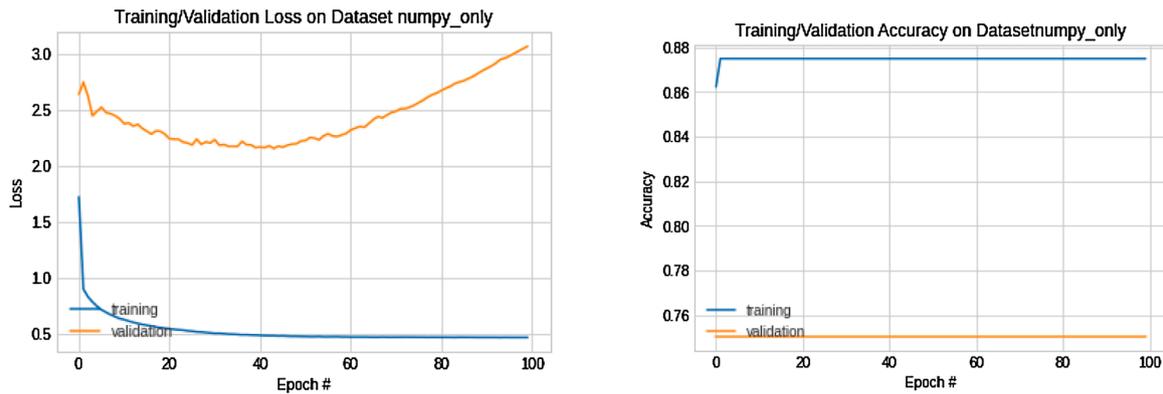

Fig. 6. Accuracy and loss of training and evaluating MIAS dataset with data augmentation defined in item (i).
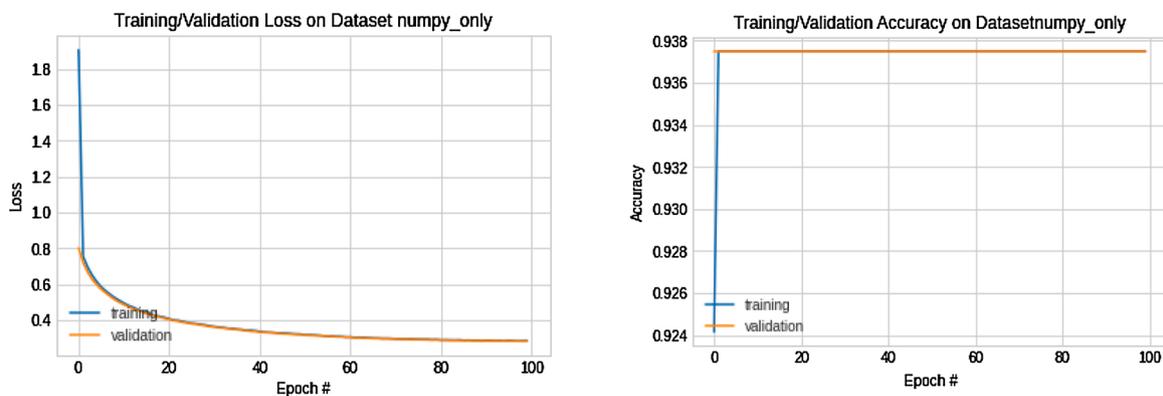

Fig. 7. Accuracy and loss of training and evaluating MIAS dataset with data augmentation defined in item (ii).

different forms using augmentation technique. To do so, images were randomly selected from each batch and were transformed as follows:

i Vertical and horizontal flip, rotation at 90 degrees, width shift range of 50 %, and results are shown in Fig. 6.
ii Vertical flip, rotation at 180 degrees, height shift range of 50 %, application of ZCA whitening, validation split 0.25, the zoom range of 50 %, width shift range, and results are shown in Fig. 7.
iii Horizontal flip; rotation at 270 degrees; both width and height shift range of 50 %, application of ZCA whitening, validation split of 0.15, and results are shown in Fig. 8.

The performances of the three different scenarios of augmentation parameters defined in items (i)-(iii) above show that the proposed model

performed better under the case described by item (ii). The case described in option (ii) demonstrates a unique combination of augmentation operations on the input dataset.

5.2. DDSM + CBS Dataset (299 × 299)

The proposed model also yielded an excellent performance using the DDSM dataset. It was observed that the accuracy and loss value of the model under this dataset improved with data augmentation just like the MIAS dataset. Although the performance of the model was lower than what was obtained with the MIAS datasets, it was, however, discovered that the model was effective. Results obtained show an accuracy of 87.38 % and loss values of 0.58 in training phase while the accuracy of 86.80 % with loss value of 0.60 was achieved during evaluation. These
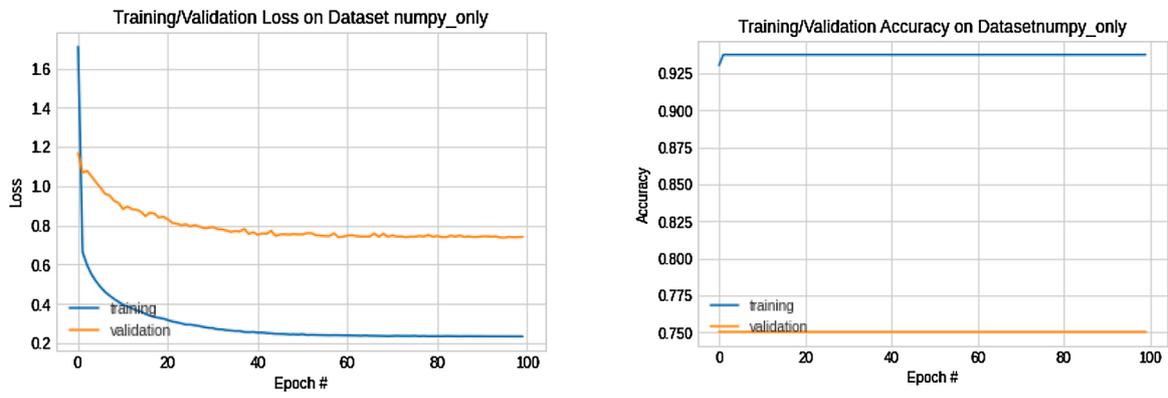
Fig. 8. Accuracy and loss of training and evaluating MIAS dataset with data augmentation defined in item (iii).
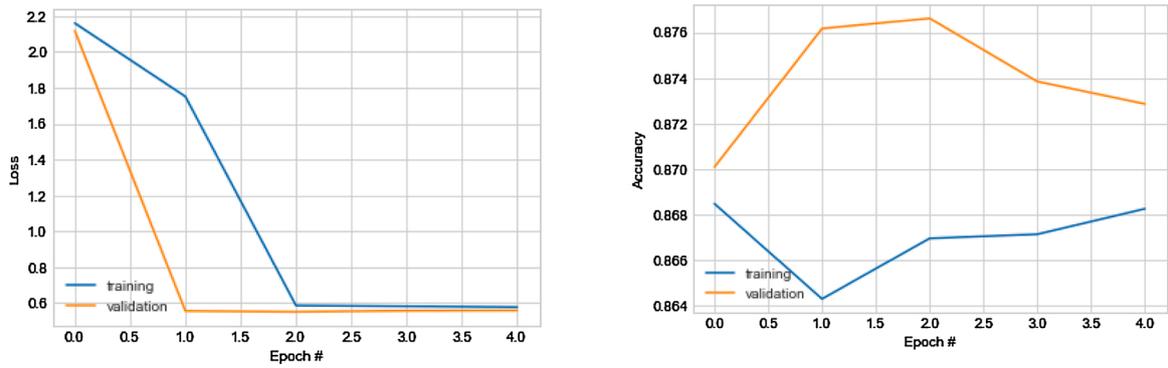

Fig. 9. Pattern of change in loss function and accuracy of training and evaluation when the proposed model is applied to DDSM + CBS dataset.
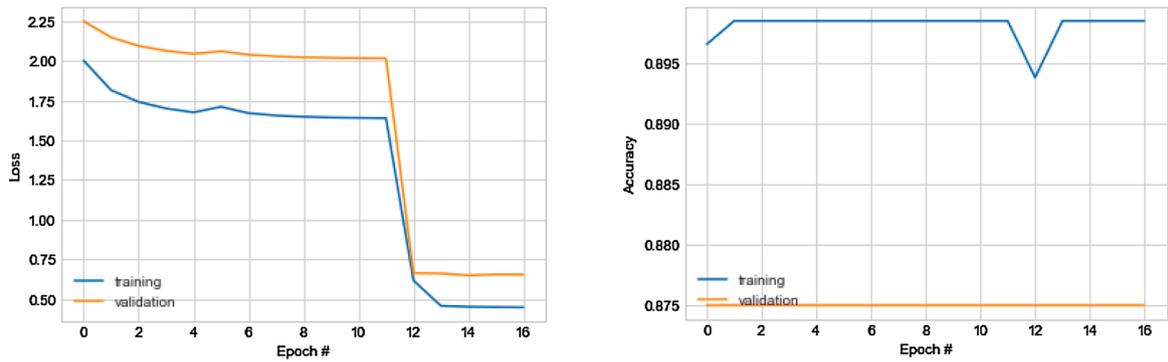

Fig. 10. Pattern of change in loss function and accuracy of training and evaluation when the proposed model with data augmentation is applied to DDSM + CBS dataset.
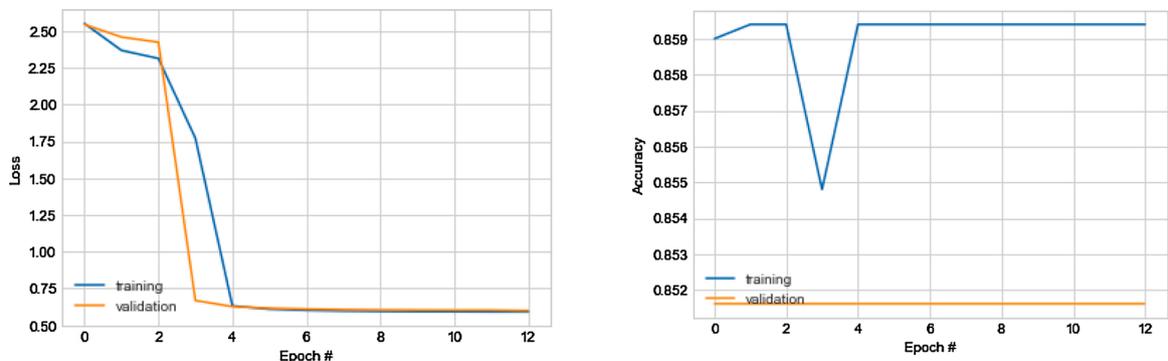

Fig. 11. Accuracy and loss of training and evaluating MIAS dataset with data augmentation defined in item (i).
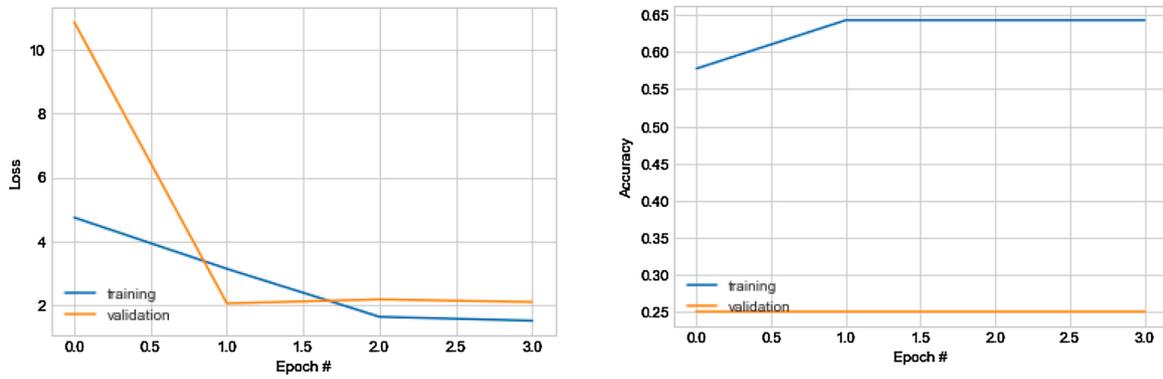
Fig. 12. Pattern of change in loss function and accuracy of training and evaluation MIAS whole image dataset (1024 × 1024) dataset with data augmentation.
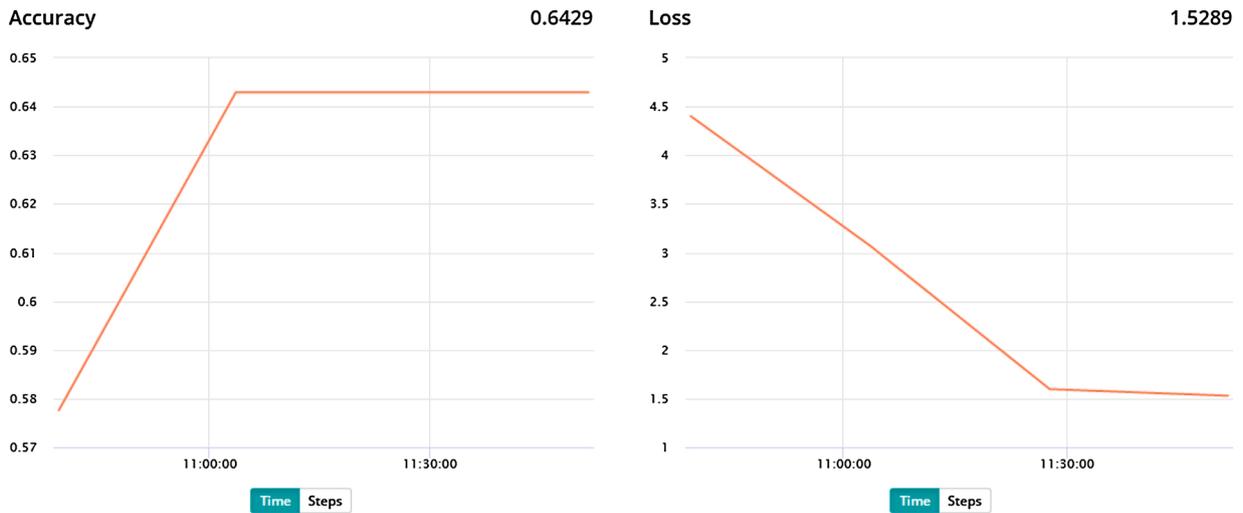

Fig. 13. Pattern of change in accuracy of training and loss function INbreast whole image dataset (2560 × 3328) dataset with data augmentation.

are captured in Fig. 9.

Similarly, the proposed model revealed that when data augmentation was applied to the DDSM + CBS dataset, it was able to improve its accuracy and decrease its lost value, as shown in Fig. 10. Using data augmentation technique, the model yielded 89.50 % and 87.50 % during training and evaluation respectively, while the loss values changed from

0.68 to 0.49 from training to evaluation.

This study also investigated the effect of varying transformation operations on the DDSM + CBS dataset, which is shown as follows:

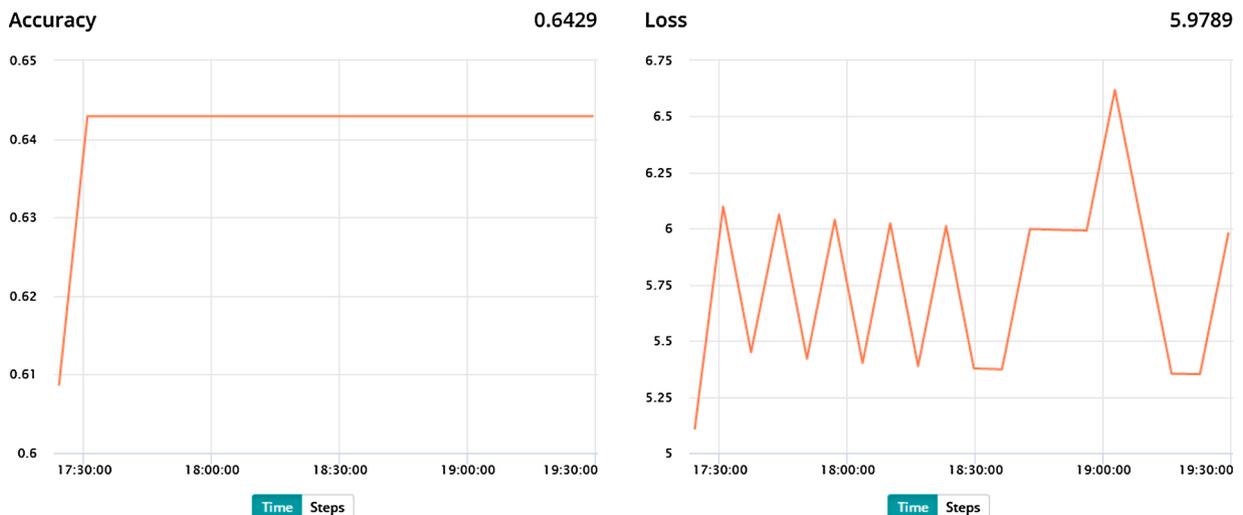i Vertical and horizontal flip, zoom range of 0.50 rotations at 180 degrees, and results are shown in Fig. 11.

Fig. 14. Pattern of change in accuracy and loss function of MIAS whole image dataset (1024 × 1024) using a CNN model with higher depth than the one proposed in this study, that is, besides the use of data augmentation.
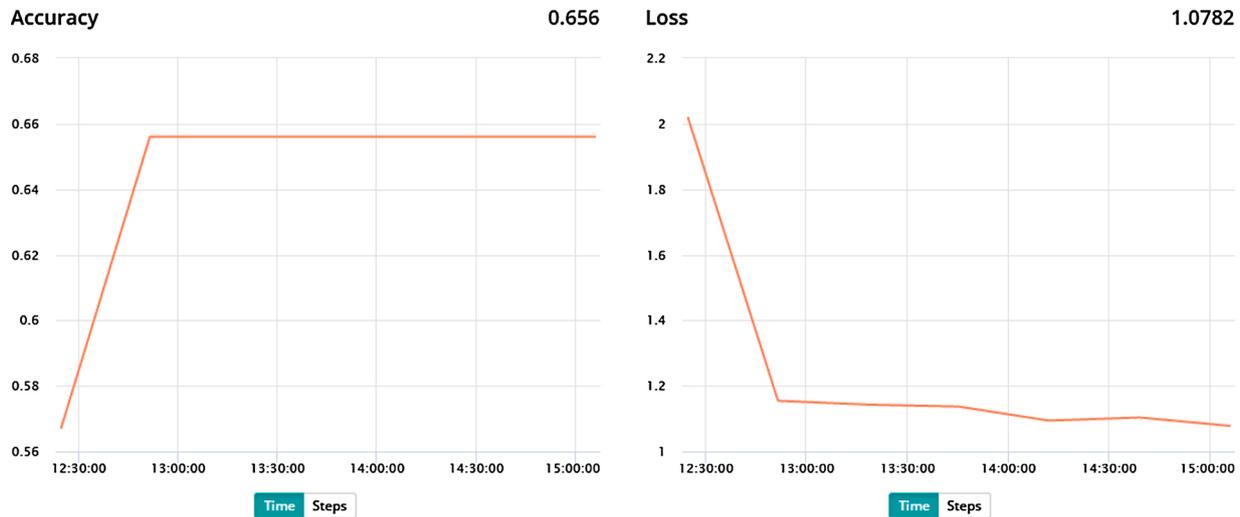
Fig. 15. Pattern of change in accuracy and loss function of INbreast whole image dataset (2560 × 3328) using a CNN model with higher depth than the one proposed in this study, in addition to the use of data augmentation.

Table 8
Proposed architectural distortion-based CNN with data augmentation.

Model Description	Dataset	Accuracy
Proposed CNN model	MIAS 299 × 299 dataset	93.75 %
	DDSM + CBS dataset	87.50 %
	MIAS 299 × 299 dataset	93.75 %
Proposed CNN model + data augmentation	DDSM + CBS dataset	90.62 %
	INbreast dataset 2560 × 3328	64.29 %
	MIAS 1024 × 1024	64.29 %

Table 9
Comparative analysis of the proposed architectural distortion CNN with popular CNN architectures.

CNN	Number of Weights	Batch Size	Learning Rate (FT, SC)	Best Model Iterations (Fine Tuning, from scratch)
Proposed CNN model	668,837	32	0.001	8, 20
GoogLeNet	10,299,840	32	10^{-5}, 10^{-5}	12, 12
AlexNet	56,866,848	32	10^{-5}, 10^{-5}	665
ResNet-50	23,512,128	32	10^{-5}, 10^{-4}	4104
VGG-16	134,256,320	32	10^{-5}, 10^{-5}	9, 58

5.3. Performance on whole/large images

We discovered that our model performs well when the input sizes of datasets are served as small patches (preferable 299 × 299) rather than whole images (larger than 512 × 512). This was established based on the experimentation carried out on our model using the MIAS and INbreast whole mammogram image datasets.

5.3.1. MIAS whole image dataset (1024 × 1024)

We served our proposed model with samples of 1024 × 1024 sized whole MIAS images and found the accuracy to have dropped sharply from expectations. The accuracy of 64.29 % and loss value of 5.9789, as shown in Fig. 12, revealed that our model appears to be too shallow for all the features of the images to be learned. We acknowledge that both the experimentation on MIAS whole image dataset (1024 × 1024) and INbreast whole image dataset (2560 × 3328) were performed even with data augmentation operation yet we obtained a reduced performance compared to those previously reported in the sections above.

5.3.2. INbreast whole image dataset (2560 × 3328)

Similarly, the model does not appear to appreciate in performance under the INbreast 2560 × 3328 dataset. This dataset also under-performed compared to the small-sized samples as it yielded maximum accuracy and loss values of 64.29 % and 1.5298, respectively, as shown in Fig. 13. We were interested in investigating why the performance of the sets of datasets dropped so much to those low values of accuracies, and why the loss values were so high.

We experimented with a deeper model by increasing the depth of our model to unravel the reason for the sharp drop in performance. That investigation did not yield any better performance, as shown in Fig. 14 for MIAS whole images and Fig. 15 for INbreast whole images.

In summary, we present the best performance obtained by the

Table 10
Comparing the contributions and performances of similar approaches.

| Authors	Contribution	Dataset	Accuracy
Proposal	Deep learning architecture for effective detection of architectural distortion using data augmentation	DDSM + CBIS, INbreast, and MIAS	93.75 %
Abbas [62]	DeepCAD: multilayer deep-learning architecture DCNN: AlexNet.	600 ROI: masses 300, benign300	91.5%
Ragab et al. [59]	Fine-tuned to classify two classes instead of 1000 classes	DDSM and CBIS-DDSM; Data Augmentation: rotation	73.6 %,
Hang et al. [46]	GlimpseNet	DDSM	66.2 %
Jadoon et al. [63]	CNN-DW and CNN-CT with augmented data set	IRMA (DDSM and MIAS) and Lawrence Livermore National Laboratory (LLNL) and Rheinisch Westfälische Technische Hochschule (RWTH)	83.74 %

proposed model in Table 8 both in the data augmentation cases and non-augmentation cases.

Furthermore, we attempted to compare the performance of our model in detecting architectural distortion in mammograms with similar researches which might have used the same dataset. The result of the
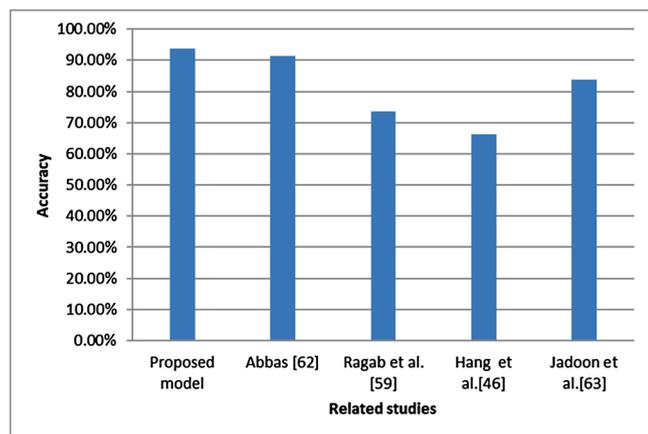
Table 11

Comparing the performances of accuracies of popular architectures with our proposed model based on CBIS-DDSM dataset.

Architectures	Accuracy based on CBIS-DDSM dataset
Proposed architecture	87.29 %
GoogLeNet	59.80 %
AlexNet	65.60 %
ResNet-50	62.70%
VGG-16	58.00%
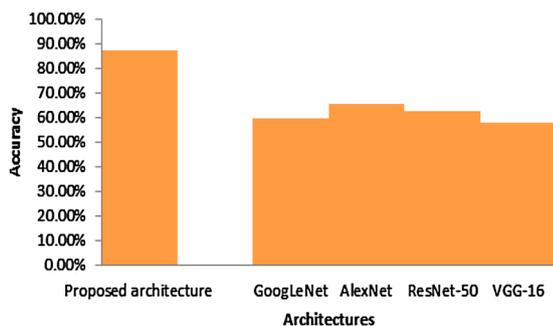
(a)

(b)

Fig. 16. A representation of graphical comparison of the proposed CNN model/architecture with (a) similar models characterizing architectural distortion in digital images and (b) state-of-the-art architectures, both in terms of accuracy.

comparison is listed in Tables 9 and 10. We discovered that our model yielded outstanding performance in comparison with other similar approaches.

In Table 11, we make a comparative analysis of some popular architecture with our proposed model using the well-known CBIS-DDSM dataset. We felt it necessary to present the performance evaluation over similar architectures using the same dataset. This will present an unbiased evaluation of the achievement obtained in this paper. Performance comparison based on accuracy with other major architectures revealed that we obtained 87.29 % accuracy, which translates to 21.69 % accuracy improvement over the optimal architecture (AlexNet).

In Fig. 16 (a) and (b), we graphically analyse the performance of our proposed model in comparison with other state-of-the-art CNN models for architectural distortion and generic CNN architectures.

Although similar learning models tend to train over a large number of epochs with the hope attaining a desirable accuracy, we, however, discovered that the number of epochs is not as significant as the validation and training error. Our approach in this study was to keep watch on our training and validation error; as long as it keeps dropping training

should continue until it stabilizes. For instance, if the training/validation error starts increasing, that might be an indication of overfitting. We, therefore, attempted our training with a number of epochs as high as possible and terminated training based on the error rates in addition to using an early stopping call back.

We understand that designing an optimal and performance efficient CNN model which can accept variable sized inputs (especially large images) requires some artistic work rather than the scientific method. However, we decided to subject our proposed model to both small-sized (299×299) input and large-sized inputs (1024×1024 and 2560×3328) to study, investigate and report performance of the model. This will help researchers understand the intricacy and challenges of such a method. Our position on this approach was, therefore, drawn from practical experience learnt from this study. Although we do not claim that such models are unrealistic, however some compromise may be attained with learning model depth and computational cost.

6. Conclusion

Characterization of abnormalities in mammograms for the purpose of detection of malignant breast tissue has generated significant interest in deep learning research. Although more study and state-of-the-art performances have been achieved in the detection of micro-calcifications and masses, only a few efforts are directed to the detection of architectural distortion. The negligence of study in the detection of architectural distortion is largely due to its subtle nature which presents it as irrelevant and rarely occurring in breast images. However, we observed that detection of this abnormality enhances early revealing of breast cancer. In this paper, we proposed a novel deep learning architecture with **Conv-Conv-Pool** layering style and aimed at effectively detecting the existence of architectural distortion in mammograms. We have improved the performance of our model by augmenting the data using standard augmentation technique. The proposed model strengthens the conventional claims that data augmentation enhances the performance of learning models when applied in a novel way. In addition, we applied our proposed model on several benchmarked datasets to obtain an interesting result. The result obtained showed that our proposed model outperforms similar models adapted to the task of detection and classification of architectural distortion from digital breast images. Moreover, this study demonstrates rigorous experimentation with benchmark datasets by investigating and reporting their performances when applied to the proposed model. Finally, an outstanding output of this study is the careful selection of data argumentation parameters using the traditional approach. All these demonstrates an interesting performance of the proposed approach in this study. In future, we plan to apply GAN-based image generation approach as a replacement to our augmentation technique. This is intended to boost the performance of the proposed model further. Besides this, we plan to work around the proposed model by deepening it and adjusting necessary parameters to provide support for whole (images) mammograms like those from INbreast dataset.

CRediT authorship contribution statement

Olaide N. Oyelade: Conceptualization, Methodology, Software, Data curation, Writing - original draft. **Absalom E. Ezugwu:** Supervision, Writing - review & editing.

Declaration of Competing Interest

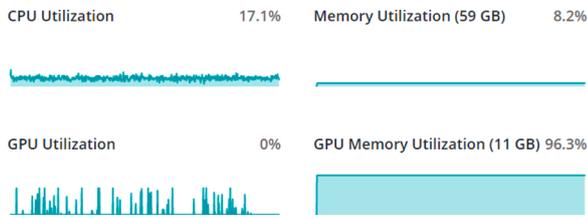The authors report no declarations of interest.

Fig. A1. Comparison of GPU/CPU utilization for DDSM + CBS dataset (299×299) on GPU Tesla V100.

Fig. A2. CPU/memory utilization for MIAS dataset (299×299) on Standard CPU configuration under data augmentation.

Fig. A3. CPU/memory utilization for MIAS dataset (299×299) on Standard CPU configuration not under data augmentation.
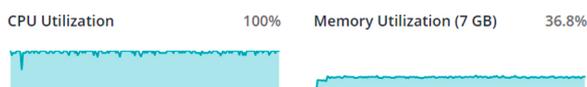

Fig. A4. CPU/memory utilization for DDSM + CBS dataset (299×299) on High-Performance CPU configuration under data augmentation.

Fig. A5. CPU/memory utilization for DDSM + CBS dataset (299×299) on Standard CPU configuration not under data augmentation.

Appendix A. Computational Performance of the Experimentation Resources

A Computational resource usage metrics

For each of the jobs we executed in the training mode, we observed the CPU and GPU performance of each job and captured them to demonstrate resource utilization in the computational environment. For example, the GPU or CPU Memory utilization metric might indicate whether we can increase or decrease the batch size of our jobs to ensure that we are fully utilizing your GPU/CPU. It can also help debug failed jobs due to out-of-memory (OOM) errors. We received CPU utilization which showed the percentage of CPU usage; memory utilization which showed the percentage of RAM usage; and disk utilization which showed the percentage of SSD usage. Additionally, for the jobs we executed on GPU-powered configurations, we received GPU utilization which showed the percentage of GPU usage by the training job; and GPU memory utilization which showed the percentage GPU Memory by the training job. Fig. 1 below describe the metrics of these computational resources with respect to our jobs.

Fig. A1–A5

References

[1] Cancer Health Center (CHC): Understanding Cancer – Diagnosis and Treatment, 2019. Retrieved 4th September from: https://www.webmd.com/cancer/understanding-cancer-treatment#1.

[2] C. Dang, T.A. Gilewski, A. Surbone, L. Norton, Holland-Frei cancer medicine. Growth Curve Analysis. Holland-frei Cancer medicine. 6th Edition, 6th edition, Publication of the American Cancer Society, 2003.

[3] M.P. Sampat, G.J. Whitman, M.K. Markey, A.C. Bovik, "Evidence based detection of spiculated masses and architectural distortion", in: San Diego, in: J. M. Fitzpatrick, J.M. Reinhardt (Eds.), Proceedings of SPIE Medical Imaging 2005: Image Processing, 5747, 2005, pp. 26–37.

[4] How Common Is Breast Cancer? Retrieved 22nd April 2019, from https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html.

[5] J. Ferlay, H.R. Shin, F. Bray, D. Forman, C. Mathers, D.M. Parkin, Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008, Int. J. Cancer 127 (12) (2010) 2893–2917, https://doi.org/10.1002/ijc.25516, 2010 Dec 15.

[6] Davies Adeloye, Olaperi Y. Sowunmi, Wura Jacobs, Rotimi A. David, Adeyemi A. Adeosun, Ann O. Amuta, Sanjay Misra, Muktar Gadanya, Asa Auta, Michael O. Harhay, Kit Yee Chan, Estimating the incidence of breast cancer in Africa: a systematic review and meta-analysis, J. Glob. Health 8 (June 1) (2018), https://doi.org/10.7189/jogh.08.010419, 2018.

[7] WHO: Breast Cancer: Prevention and Control, 2019. Retrieved 3rd September 2019, from: http://www.who.int/cancer/detection/breastcancer/en/index1.html.

[8] T.W. Freer, M.J. Ulissey, Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center, Radiology 220 (3) (2001) 781–786, https://doi.org/10.1148/radiol.2203001282, 2001.

[9] A.-A. Nahid, Y. Kong, Involvement of machine learning for breast cancer image classification: a survey, in: Computational and Mathematical Methods in Medicine, 29, 2017.

[10] P. Xi, C. Shu, R.A. Goubran, Abnormality detection in mammography using deep convolutional neural networks, 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA) (2018) 1–6.

[11] W. Zhu, X. Xie, Adversarial Deep Structural Networks for Mammographic Mass Segmentation, ArXiv, 2016 abs/1612.05970.

[12] L. Tsochatzidis, L. Lena Costaridou, I. Pratikakis, Deep learning for breast Cancer diagnosis from mammograms—a comparative study, J. Imaging 5 (37) (2019), https://doi.org/10.3390/jimaging5030037, 2019.

[13] R.M. Rangayyan, Computer-aided Detection of Subtle Signs of Early Breast Cancer: Detection of Architectural Distortion in Mammograms. International Conference on Biomedical Engineering and Assistive Technologies (BEATS 2012), Dr. B.R. Ambedkar National Institute of Technology, Jalandhar, India, 2012. December, 2012.

[14] D. Abdelhafiz, C. Yang, R. Ammar, et al., BMC Bioinformatics 20 (Suppl 11) (2019) 281, https://doi.org/10.1186/s12859-019-2823-4.

[15] M.A. Durand, S. Wang, R.J. Hooley, M. Raghu, L.E. Philpotts, Tomosynthesis-detected architectural distortion, Manage. Algor. Radiol. Pathol. Correlat. 36 (2) (2016) 311–321, https://doi.org/10.1148/rg.2016150093.

[16] Vandana Dialani, Priscilla J. Slanetz, Ronald L. Eisenberg, S. Gaur, V. Dialani, P. J. Slanetz, R.L. Eisenberg, Architectural distortion of the breast, AJR Am. J. Roentgenol. 201 (Nov 5) (2013) W662–670, https://doi.org/10.2214/AJR.12.10153.

[17] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Rev. Insight Nat. 5 (21) (2015) 436–444, https://doi.org/10.1038/nature14539.

[18] Arunava, Convolutional Neural Network: An Introduction to Convolutional Neural Networks, Retrieved 4th September 2019, from:, 2018 https://towardsdatascience.com/convolutional-neural-network-17fb77e76c05.

[19] H. Shin, H. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. Summers. Deep convolutional neuralnetworks for computer-aided detection: Cnn architectures,dataset characteristics and transfer learnings. IEEE Trans.Medical Imaging, 35(5):1285–1298.

[20] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Neural Infor. Process. Syst. 25 (2012), https://doi.org/10.1145/3065386.

[21] C. Szegedy, et al., Going deeper with convolutions, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 1–9, https://doi.org/10.1109/CVPR.2015.7298594.

[22] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, 2016, pp. 770–778, https://doi.org/10.1109/CVPR.2016.90.

[23] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, CoRR (2014) abs/1409.1556.

[24] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE (1998) 2278–2324, https://doi.org/10.1109/5.726791, vol. 86,11Nov. 1998.

[25] A.C. Costa, H.C.R. Oliveira, J.H. Catani, N. de Barros, C.F.E. Melo, M.A.C. Vieira, Data augmentation for detection of architectural distortion in digital mammography using deep learning approach, Comput. Vis. Pattern Recognit. (2018) arXiv:1807.03167.

[26] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015) 2818–2826.

[27] C. Szegedy, S. Ioffe, V. Vanhoucke, A. Alemi, Inception-v4, Inception-ResNet and the impact of residual connections on learning, AAAI (2016).

[28] F. Chollet, Xception: deep learning with depthwise separable convolutions. Published, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017).

[29] S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, Aggregated residual transformations for deep neural networks, Published in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017).

[30] M. Jasionowska, A. Gacek, Wavelet convolution neural network for classification of spiculated findings in mammograms, Inf. Technol. Biomed. (2019) 199–208, https://doi.org/10.1007/978-3-030-23762-2_18.

[31] G.Bloch I. Palma, S. Muller, et al., Spiculated lesions and architectural distortions detection, in: J. Mart (Ed.), Digital Breast Tomosynthesis Datasets, 2010, pp. 712–719. IWDM 2010, LNCS 6136.

[32] A. Kamra, V.K. Jain, S. Singh, S. Mittal, Characterization of architectural distortion in mammograms based on texture analysis using support vector machine classifier with clinical evaluation. Society for imaging informatics in medicine 2015, J. Digit. Imaging 29 (2015) 104–114, https://doi.org/10.1007/s10278-015-9807-3.

[33] Q. Zheng, M. Yang, J. Yang, Q. Zhang, X. Zhang, "Improvement of generalization ability of deep CNN via implicit regularization in two-stage training process", IEEE Access,6, pp. 15844-15869.

[34] M. Cogswell, F. Ahmed, R. Girshick, L. Zitnick, D. Batra, Reducing overfitting in deep networks by decorrelating representations, ICLR'15 (2015).

[35] A. Antoniou, A. Storkey, H. Edwards, Data Aug-mentation Generative Adversarial Networks, arXiv preprint arXiv, 2017, 1711.04340.

[36] J. Feng, Q. Bu, F. Liu, M. Zhang, Y. Ren, Y. Lv, Breast mass detection in digital mammogram based on gestalt psychology, J. Healthc. Eng. 2018 (2018) 13, https://doi.org/10.1155/2018/4015613.

[37] P. Teare, M. Fishman, O. Benzaquen, E. Toledano, E. Elnekave, Malignancy detection on mammography using dual deep convolutional neural networks and genetically discovered false color input enhancement, J. Digit. Imaging 4 (30) (2017) 499–505.

[38] M. Maqsood, F. Nazir, U. Khan, F. Aadil, H. Jamal, I. Mehmood, O. Song, Transfer learning assisted classification and detection of alzheimer's disease stages using 3D MRI scans, Sensors (Basel 19 (June 11) (2019) 2645, https://doi.org/10.3390/s19112645, 2019Jun 11.

[39] S. M. L. de Lima, A. G. da Silva-Filho, and W. P. dos Santos, "Detection and classification of masses in mammographic images in a multi-kernel approach," Computer Methods and Programs in Biomedicine, 134. 11–29.

[40] L. Perez and J. Wang. The effectiveness of data augmentationin image classification using deep learning. arXiv preprint arXiv:1712.04621.

[41] Thomas Neff ;, Data Augmentation in Deep Learning Using Generative Adversarial Networks. MASTER`S THESIS to Achieve the University Degree of Diplom-ingenieu Master's Degree Programme Information and Computer Engineering, 2018.

[42] T. Kooi, A. Gubern-Merida, J.J. Mordang, R. Mann, R. Pijnappel, K. Schuur, et al., A comparison between a deep convolutional neuralnetwork and radiologists for classifying regions of interest in mammography, in: International Workshop on Digital Mammography, 2016, Springer, 2016, pp. 51–56.

[43] R.M. Rangayyan, S. Banik, J.L. Desautels, Detection of architectural distortion in prior mammograms via analysis of oriented patterns, J. Vis. Exp. (78) (2013), e50341, https://doi.org/10.3791/50341.

[44] R. Ben-Ari, A. Akselrod-Ballin, L. Karlinsky, S. Hashoul, Domain specific convolutional neural nets for detection of architectural distortion in mammograms, Biomedical Imaging (ISBI 2017) 2017 IEEE 14th International Symposium on IEEE (2017) 552–556.

[45] R. Bakalo, J. Goldberger, R. Ben-Ari, A dual branch deep neural network for classification and detection in mammograms. Computer Vision and Pattern Recognition, arXiv, 2019, 1904.12589.

[46] W. Hang, Z. Liu, A. Hannun, GlimpseNet: Attentional Methods for Full-Image Mammogram Diagnosis, Stanford AI Lab Internal Report, Retrieved 31 December 2019, from, Stanford University, Stanford, CA, USA, 2017, http://cs231n.stanford.edu/reports/2017/pdfs/517.pdf.

[47] Y. Qiu, S. Yan, R.R. Gundreddy, Y. Wang, S. Cheng, H. Liu, et al., A new approach to develop computer-aided diagnosis scheme of breast mass classification using deep learning technology, J. X-Ray Sci. Technol. (Preprint) (2017) 1–13.

[48] Z. Jiao, X. Gao, Y. Wang, J. Li, A deep feature based framework for breast masses classification, Neurocomputing 197 (2016) 221–231, https://doi.org/10.1016/j.neucom.2016.02.060.

[49] I. Bakkouri, K. Afdel, Breast tumor classification based on deep convolutional neural networks, Advanced Technologies for Signal and Image Processing (ATSIP). International Conference on IEEE 2017 (2017) 1–6.

[50] A. Dubrovina, P. Kisilev, B. Ginsburg, S. Hashoul, R. Kimmel, Computational mammography using deep neural networks, Comput. Methods Biomech. Biomed. Eng. Imaging Vis. (2016) 1–5.

[51] R.K. Samala, H.P. Chan, L.M. Hadjiiski, M.A. Helvie, Cha KH. Richter CD. Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms, Phys. Med. Biol. 23 (62) (2017) 8894.

[52] N. Antropova, B.Q. Huynh, M.L. Giger, A deep feature fusion methodology for breast Cancer diagnosis demonstrated on three imaging modality datasets, Med. Phys. 44 (October10) (2017) 5162–5171, https://doi.org/10.1002/mp.12453.

[53] H. Chougrad, H. Zouaki, O. Alheyane, Convolutional Neural Networks for Breast Cancer screening: Transfer Learning with Exponential Decay, arXiv preprint arXiv, 2017, 171110752.

[54] M.M. Jadoon, Q. Zhang, I.U. Haq, S. Butt, A. Jadoon, Three-class mammogram classification based on descriptive CNN features, Biomed Res. Int. 2017 (2017) (2017) 11–19, https://doi.org/10.1155/2017/3640901. Article ID 3640901.

[55] F. Jiang, H. Liu, S. Yu, Y. Xie, Breast mass lesion classification in mammograms by transfer learning. Proceedings of the 5th International Conference on Bioinformatics and Computational Biology, ACM, 2017, pp. 59–62.

[56] K. Sharma, B. Preet, Classification of Mammogram Images by Using CNN Classifier, 2016, https://doi.org/10.1109/ICACCI.2016.7732477, 2743–9.

[57] M. Marchesi, Megapixel Size Image Creation Using Generative Adversarial Networks, ArXiv, 2017 abs/1706.00082.

[58] I.C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M.J. Cardoso, J.S. Carddoso, INreast: toward a full-field digital mammographic database. Faculdade De Medicina, Alameda Prof. Hernˆani Monteiro, Universidade do Porto, 2012, pp. 236–248, https://doi.org/10.1016/j.acra.2011.09.014. Acad Radiol 2012; 19.

[59] D.A. Ragab, M. Sharkas, S. Marshall, J. Ren, Breast cancer detection using deep convolutional neural networks and support vector machines, PeerJ 7 (2019) e6201, https://doi.org/10.7717/peerj.6201.

[60] Murali.S. Minavathi, M.S. Dinesh, Model based approach for detection of architectural distortions and spiculated masses in mammograms, IJCSE 3 (11) (2011) 3534–3546.

[61] Floyed Server, 2019. Retrieved 12 December 2019 from, https://www.floydhub.com/jobs.

[62] Q. Abbas, DeepCAD: a computer-aided diagnosis system for mammographic masses using deep invariant features, Computers 4 (5) (2016) 28, https://doi.org/10.3390/computers5040028. Computers2016, 5(4), 28.

[63] M.M. Jadoon, Q. Zhang, I.U. Haq, S. Butt, A. Jadoon, Three-class mammogram classification based on descriptive CNN features, Biomed Res. Int. 2017 (2017) (2017) 11, https://doi.org/10.1155/2017/3640901, 3640901.

[64] X. Zhang, Y. Zhang, E.Y. Han, N. Jacobs, Q. Han, X. Wang, J. Liu, Classification of whole mammogram and tomosynthesis images using deep convolutional neural networks, IEEE Trans. Nanobioscience 17 (July 3) (2018) 237–242, https://doi.org/10.1109/TNB.2018.2845103. Epub 2018 Jun 72018.

[65] D. Ribli, A. Horváth, Z. Unger, et al., Detecting and classifying lesions in mammograms with deep learning, Sci. Rep. 8 (2018) 4165, https://doi.org/10.1038/s41598-018-22437-z.

[66] L. Shen, L.R. Margolies, J.H. Rothstein, et al., Deep learning to improve breast Cancer detection on screening mammography, Sci. Rep. 9 (2019) 12495, https://doi.org/10.1038/s41598-019-48995-4.

[67] Fei Gao, Teresa Wu, Jing Li, Bin Zheng, SD-CNN: a shallow-deep CNN for improved breast cancer diagnosis, Comput. Med. Imaging Graph. 70 (2018), https://doi.org/10.1016/j.compmedimag.2018.09.004.

[68] S. Eleftherios Trivizakis Georgios, D. Ioannidis Vasileios, Z. Melissianos Georgios, Aristidis Tsatsakis Demetrios A. Papadakis, Kostas Marias Spandidos, A Novel Deep Learning Architecture Outperforming 'off-the-shelf' Transfer Learning and Feature-based Methods in the Automated Assessment of Mammographic Breast Density. Oncology Reports, 2019, pp. 2009–2015, https://doi.org/10.3892/or.2019.7312.

[69] Arthur C. Costa, Helder C.R. Oliveira, Lucas R. Borges, Marcelo A.C. Vieira, Transfer learning in deep convolutional neural networks for detection of architectural distortion in digital mammography, Proc. SPIE 11513, 15th International Workshop on Breast Imaging (IWBI2020) (2020), https://doi.org/10.1117/12.2564348, 115130N (22 May).

[70] Yue Li, Zilong He, Yao Lu, Xiangyuan Ma, Yanhui Guo, Zheng Xie, Zeyuan Xu, Weiguo Chen, Haibin Chen, Deep learning on mammary glands distribution for architectural distortion detection in digital breast tomosynthesis, Phys. Med. Biol. (2020), https://doi.org/10.1088/1361-6560/ab98d0.

[71] Muhammad Attique Khan, Imran Ashraf, Majed Alhaisoni, Robertas Damaševičius, Rafal Scherer, Amjad Rehman, Syed Ahmad Chan Bukhari, Multimodal brain tumor classification using deep learning and robust feature selection: a machine learning application for radiologists, Diagnostics (Basel) 10 (August 8) (2020) 565, https://doi.org/10.3390/diagnostics10080565, 2020.

[72] Muhammad Attique Khan, Seifedine Kadry, Majed Alhaisoni, Yunyoung Nam, Yudong Zhang, Venkatesan Rajinika, Sarfraz Muhammad Shahzad, Computer-aided gastrointestinal diseases analysis from wireless capsule endoscopy: a framework of best features selection, IEEE Access 8 (2020) 132850–132859, https://doi.org/10.1109/ACCESS.2020.3010448, 2020.

[73] S. Zahoor, Ikram Ullah Lali, M.A. Khan, K. Javed, W. Mehmood, Breast Cancer detection and classification using traditional computer vision techniques: a comprehensive review, Curr. Med. Imaging Rev. (April) (2020) 6, https://doi.org/10.2174/1573405616666200406110547. Epub ahead of print. PMID: 32250226.

[74] Amna Liaqat, Muhammad Attique Khan, Muhammad Sharif, Mamta Mittal, Tanzila Saba, K. Suresh Manic, Feras Nadhim Hasoon Al Attar, Gastric tract infections detection and classification from wireless capsule endoscopy using computer vision techniques: a review, Curr. Med. Imaging Rev. 16 (2020) 1, https://doi.org/10.2174/1573405616666200425220513.

[75] Abdul Majid, Muhammad Attique Khan, Mussarat Yasmin, Amjad Rehman, Abdullah Yousafzai, Usman Tariq, Classification of stomach infections: a paradigm of convolutional neural network along with classical features fusion and selection, Microsc. Res. Tech. 83 (5) (2020), https://doi.org/10.1002/jemt.23447.

[76] Muhammad Irfan Sharif, Jian Ping Li, Muhammad Attique Khan, Muhammad Asim Saleem, Active deep neural network features selection for segmentation and recognition of brain tumors using MRI images, Pattern Recognit. Lett. 129 (1) (2020) 181–189.

[77] M. Attique Khan, S. Rubab, Asifa Kashif, Muhammad Imran Sharif, Nazeer Muhammad, Jamal Hussain Shah, Yu-Dong Zhang, Suresh Chandra Satapathy, Lungs Cancer classification from CT images: an integrated design of contrast based classical features fusion and selection, Pattern Recognit. Lett. 129 (1) (2020) 77–85.

[78] American Cancer Society, Breast CancerFacts & figures 2019-2020, Am. Cancer Soc. J. CA (2020). Retrieved from https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2019-2020.pdf on 26th September 2020.

[79] Rebecca L. Siegel, Kimberly D. Miller, Ahmedin Jemal, Cancer statistics, A Cancer J.Clin. 70 (1) (2020), https://doi.org/10.3322/caac.21590.

[80] National Cancer Insitute: Surveillance, Epidemiology, and End Results Program. Cancer Stat Facts: Female Breast Cancer, 2020. Retrieved from https://seer.cancer.gov/statfacts/html/breast.html on 26th September.

[81] A. A Lorek, A. Śliwczyński, B. Więckowska, B. Stawowski, J. Dagiel, J. Gawrychowski, Analysis of diagnostic methods for focal lesions in breast glands with use of open surgical biopsies and core-needle biopsies in Poland, Med. Sci. Monit. 24 (2018) 4974–4981, https://doi.org/10.12659/MSM.908658.

[82] Marek Kowal, Paweł Filipczuk, Andrzej Obuchowicz, J.ózef Korbicz, Roman Monczak, Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images, Comput. Biol. Med. 43 (10) (2013) 1563–1572, https://doi.org/10.1016/j.compbiomed.2013.08.003.

[83] Nithya Rajendran, B. Santhi, Classification of normal and abnormal patterns in digital mammograms for diagnosis of breast Cancer, Int. J. Comput. Appl. 28 (6) (2011), https://doi.org/10.5120/3391-4707.

[84] Sidney M. L de Lima, Abel G. da Silva-Filho, Wellington Pinheirodos Santos, Detection and classification of masses in mammographic images in a multi-kernel approach, Comput. Methods Programs Biomed. 134 (October 2016) (2013) 11–29.

[85] Swagota Bera, Monisha Sharma, Bikesh Singh, Feature extraction and analysis using gabor filter and higher order statistics for the JPEG steganography, Int. J. Appl. Eng. Res. Dev. 13 (5) (2018) 2945–2954. ISSN 0973-4562 (2018).

[86] S. Khan, M. Hussain, H. Aboalsamh, et al., A comparison of different Gabor feature extraction approaches for mass classification in mammography, Multimed. Tools Appl. 76 (2017) 33–57, https://doi.org/10.1007/s11042-015-3017-3.

[87] Yufeng Zheng, Breast Cancer detection with Gabor features from digital mammograms, Algorithms 3 (1) (2010), https://doi.org/10.3390/a3010044.

[88] Murali S. Minavathi, M.S. Dinesh, Model based approach for detection of architectural distortions and spiculated masses in mammograms, IJCSE 3 (11) (2011) 3534–3546.

[89] Salabat Khan, Muhammad Hussain, Hatim A. Aboalsamh, George N. Bebis, A comparison of different Gabor feature extraction approaches for mass classification in mammography, Multimed. Tools Appl. 76 (1) (2017).

[90] Naser Safdarian1, Mohammad Reza Hediyehzadeh, Detection and classification of breast Cancer in mammography images using pattern recognition methods, Multidiscip. Cancer Investig. 3 (4) (2019) 13–24, https://doi.org/10.30699/acadpub.mci.3.4.13.

[91] Olaide Nathaniel Oyelade, Absalom El-Shamir Ezugwu, A state-of-the-art survey on deep learning methods for detection of architectural distortion from digital mammography, IEEE Access 8 (2020) (2020) 148644–148676, https://doi.org/10.1109/ACCESS.2020.3016223.

[92] Jan KukačkaVladimir GolkovDaniel CremersDaniel Cremers, Regularization for deep learning: a taxonomy, Conference Paper at ICLR 2018 (2018).

[93] Nusrat Ismoilov Sung-Bong Jang, A comparison of regularization techniques in deep neural networks, Symmetry 10 (11) (2018) 648, https://doi.org/10.3390/sym10110648.