Augmentation Techniques on Mammogram Images for CNN Based Breast Cancer Classification

P. Pratheep Kumar, V. Mary Amala Bai, and Geetha G. Nair

Abstract Deep learning is now the fastest expanding area of several medical image classification and identification. Convolutional neural networks (CNN) are the primary method used for classification across many deep neural networks (DNN). In breast cancer, classification using mammogram image has several challenges such as small dataset size and class imbalance issues. Small dataset issue is a major challenge while performing classification of medical images. Large set of training data is required to build a reliably performing machine learning model for classification. Practically, it is very difficult to generate a bench marked, pathologically tested, large set of medical images. To overcome this problem by proposed and implemented the image data augmentation is a method that can be used. We choose 115 breast mammography photographs with masses from the INbreast database for this study. The amount of breast mammography images was increased to 7732 image data by data augmentation. We utilize the preprocessing process to the breast mammography images, and then apply the CNN ideal is used to classify the images as benign or malignant. In this comparison, the quantitative analysis of classification performance between two processes such as before augmentation technique achieved 94.56% and after augmentation technique achieved better classification accuracy of 98.91%, respectively.

Keywords Breast cancer · Neural network · Convolutional neural network · Augmentation technique · INbreast database · And benign or malignant.

P. Pratheep Kumar (✉)
Department of Computer Science and Engineering, Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India

V. Mary Amala Bai
Department of Information Technology, Noorul Islam Centre For Higher Education, Kumaracoil, Tamil Nadu, India

G. G. Nair
Consultant in Obstetrics and Gynaecology, Women and Children Hospital, Thycaud, Thampanoor, Kerala, India

N. Marriwala et al. (eds.), *Emergent Converging Technologies and Biomedical Systems*, Lecture Notes in Electrical Engineering 841,
https://doi.org/10.1007/978-981-16-8774-7_53

1 Introduction

Breast cancer is a severe public health issue globally. It causes more than 1500 deaths in Switzerland alone annually. It is also the typical cancer mortality cause among women. Experts can prevent and cure breast cancer no matter at what stage it is discovered. However, earlier detection is presently the only effective choice available to reduce the illness's related physiological and psychological burden [1–3]. Mammography is the utmost sensitive method available for earlier detection of breast cancer. A discussion into mammography's efficiency to detect breast cancer early is a closed topic. Systematic mammography screening of women between 50 and 60 years is necessary to lower breast cancer mortality. Mammographic density, a robust breast cancer risk factor is increasingly being used to tailor preventive, and screening schemes. It is a primary determinant of mammography screening sensitivity and thereby of interval cancer rates. Mammography is the most important of all imaging methods to examine breast tissue, as it is efficient and accepted [4, 5]. Many computer-aided diagnosis (CAD) techniques were suggested to facilitate discovery of masses in mammograms, an important breast cancer display. The classification of tumour as benign or malignant, the features of these tumours are specified in Table 1.

CNNs produce an outcomes in a number of classification activities, but they still have a number of obstacles to face amid broad perspectives. They have problems both with over-setting and generalisation due to the vast scale of the networks touching millions of limits as well as the absence of sound workout data sets. Finally, the averting of the adversarial attacks [6] that could mislead the DNNs is a rising concern for researchers.

The researchers are battling to resolve these issues and to produce better outcomes by amending the design of the networks, designing and acquiring new learning algorithms. Lack of quality data in sufficient numbers, or an unequal level of class balance within the datasets is the most common issue [2]. The most efficient DNNs today are very large, so that a lot of data is required, which is often difficult to deliver. The famous CNN VGG16 architecture, for instance, consists of a total of 16 neuron layers with a total parameter of 138 million [7]. Moreover, ImageNet, the data set which

Table 1 Differences between benign and malignant mass tumour

Benign mass tumour	Malignant mass tumour
Benign mass are moving in nature	Malignant mass are fixed mass
Soft and clear round with besetment fibrous capsule	Irregular shaped with no capsule
Easy to remove the benign mass and does not recur again	Problematic to remove the malignant mass and may recur again
Tumour cells multiply slowly	Tumour cells multiply rapidly
The growing tumour by expanding and pushing away and against nearby tissue	The tumour growth by invading and destroying nearby tissue

holds more than 1 million pictures from 1000 non-overlapping categories, generally tests the efficiency of new architectures [8]. Data increase by data synthesis is one way of tackling this problem. The interest in multiplied data and the popularity of CNNs have been rapidly increased. The traditional and affinity-orientated transformations: creating new images through a rotation of the original image, zooming in and out, moving, applying distortions, changing the colour palette are the maximum standard and recognised operative practice for data extension. Although, advantages in some cases are not enough for simple classical operations to substantially improve the accuracy of the neural network or overawed the overfitting problem [9]. Furthermore, the current study of so-called CNN attacks has shown that deep neural networks can be easily misclassified by only limited rotations and image translations, addition of noise to images [10], or even altering a pixel in the image [11].

An algorithm presented for automated breast cancer segmentation in mammographic images scheme [12] resulted in a better classification performance. Application of thresholding technology and morphological preprocessing was the principal contribution to this algorithm in order to remove radiopaque products and labels and to separate the background area from the breast profile. The MIAS database for all mammographic images was extensively tested to show the validity of the proposed segmentation system. This database included 322 images with rectangular labels of high intensity. Bright scanning artefacts have been found in most database images. Achieved the detection accuracy about 99.06% using the high intensity square labelling. In the qualitative assessment, the method was precisely segmented throughout the breast region by covering all density classes in a large variety of digitized mammograms.

Currently developed methods for image enhancement are not only traditional methods and CNN methods. An interesting approach is a random technique proposed in [13], which can be quickly and relatively easily implemented but which gives good results in CNN generalization capacity. A noise-filled rectangle is painted randomly in a picture with the method, which leads to changes in the original pixels. As authors said it lessens the hazard of overpassing and makes the model additional robust by extending the data set to different levels of occlusion.

2 Proposed Methodology

The breast cancer classification and identification by using mammogram image proposed architecture is shown in Fig. 1. Initially we take the small breast cancer Mammography images of INbreast dataset, which has small amount of image data. As we have seen, small dataset cannot provide better classification rate. By this problem, we introduced the data augmentation technique to create the several images from small number of images by using different image augmentation scheme as flipping, cropping, noise injection, rotate, and random brightness augmentation. Primarily, we choose only 115 breast mammography images with masses and enhance the image quality by using CLAHE technique. After the data augmentation technique applied

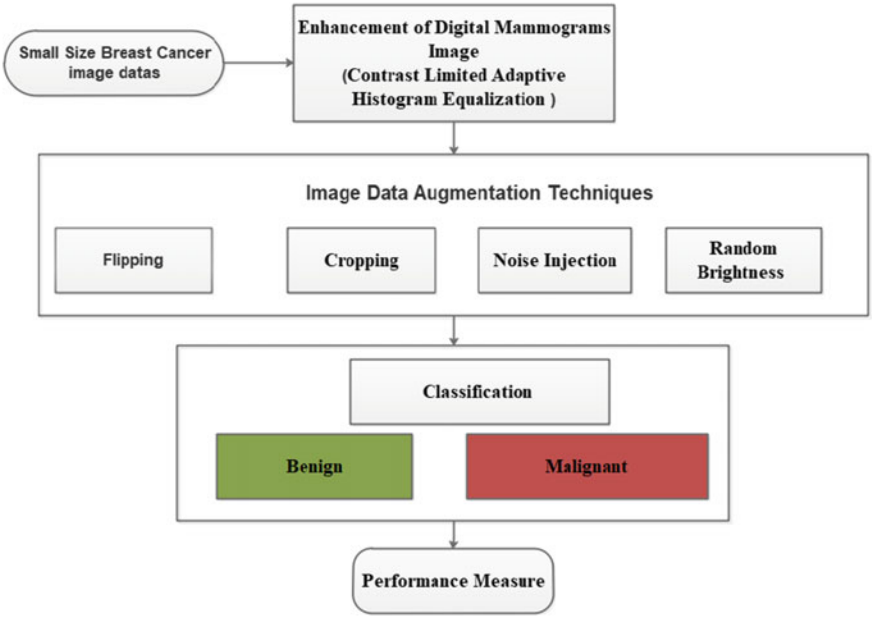

Fig. 1 Flow diagram of proposed method

to increase the image data to increase the learning rate. Further, the increased image data that is given as training and test data to a deep learning-based CNN classifier is ideal to classify the image a benign or malignant.

2.1 Data Description

Originally collected from the Centro Hospitalar S. Joao [SJSB] mammograms from the INbreast database that contain 115 cases with 410 images in total [14]. Of these, 90 were women with both breasts disease. There are four types, including the mass, calcification, asymmetry, and deformation of breast diseases detailed in the database. The images of this database contain Craniocaudal (CC) and Mediolateral oblique (MLO) views, and the breast density according to BI-RADS standards is divided into four categories, which are described in following Table 2 and also dataset sample is showed in Fig. 2.

Table 2 The benign and malignant class labels for breast density

Category	Number
Density-1 with Benign	12
Density-1 with Malignant	30
Density-2 with Benign	4
Density-2 with Malignant	32
Density-3 with Benign 1	13
Density-3 with Malignant	8
Density-4 with Benign	6
Total	115

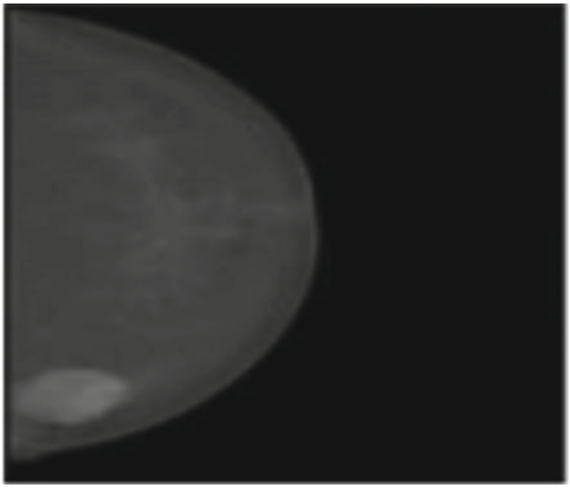

Fig. 2 Dataset sample image

3 Data Augmentation Process

Data augmentation is a vital measure of training discriminative CNNs. A variety of augmentation strategies, including flipping, cropping, noise injection, rotate, and random brightness augmentation are implemented.

3.1 Cropping

Cropping method may be used by cutting a central patch of the image in a practical manner for image with a diverse height and width. Random crops can also be used for the effect of translations very similar. In contrast, the size of the input is reduced by cutting, as a pixel ratio of (256, 256) to (126, 126) whereas translations maintain the spatial dimensions of the image. This may not be a label-preserving conversion according to the reduction threshold chosen for crops.

3.2 Noise Injection

A random value matrix generally from the Gaussian distribution consists of the injection of noise. Adding image noise can help to make CNNs more robust. Geometric changes are excellent solutions for the location differences found in training results. The distribution of training data from test data can be separated from a wide range of potential sources of bias. If there are positional distortions, such as the fact that every face is effortlessly focused in the frame, geometric changes can be a great solution. Besides great aptitude to overcome positional biases, geometric changes are beneficial, since they are simply carried out. The noise injected images is showed as in Fig. 3.

In order to start operations like horizontal flipping and rotation, many imaging processing libraries are available. Geometric transformations have the difficulties that include additional memory requirement, computer transformation, and extra training time. Some geometric transformations must be observed manually, such as reduction or random cropping, to ensure that the image label is not altered. Finally, the distances between training data and the test data are additional complex than the positional and translation variances in several application fields covered, such as medical image investigation. The scope of the application of geometric changes is therefore relatively limited.

Fig. 3 Noise injected image

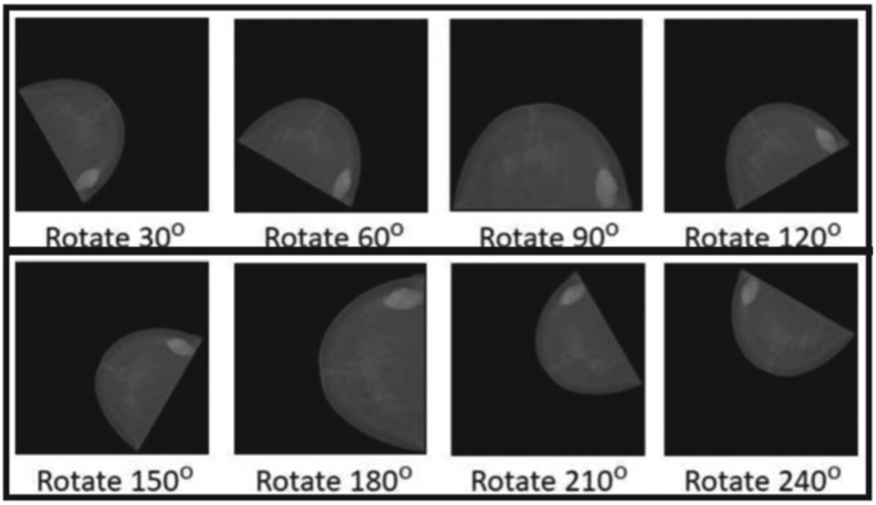

Fig. 4 Rotation of different angle image [14]

3.3 Rotation

In multi-angle-rotation data increase (modify = 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330°), and then horizontally and vertically rotate the original image and the 11-angle-rotation images. Not only does the process increase the number of samples, it also avoids overfitting. The different rotation of breast images is exposed in Fig. 4.

3.4 Flipping

In case of a vertical or horizontal breast image flip, an image flip means to reverse rows or columns of pixels. Vertical flips, we assume, capture a special medical image property, that is to say, invariance in vertical reflection. Normally only horizontal flips are used for natural pictures because vertical flips do often not represent natural pictures. However, a vertical flip of a mass would still result in a realistic mass (Fig. 5).

3.5 Random Brightness Augmentation

This is the important augmentation techniques, the brightness is randomly given to the image to create various random brightness image, which is showed in Fig. 6 (Table 3).

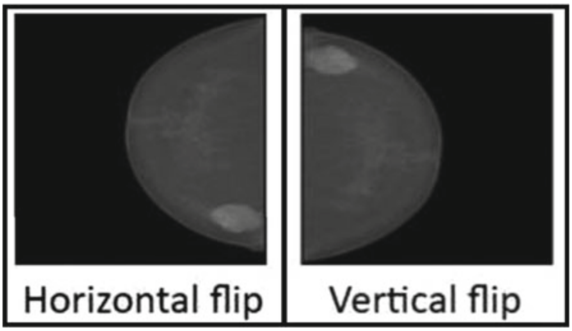

Fig. 5 Flipping breast cancer image [14]

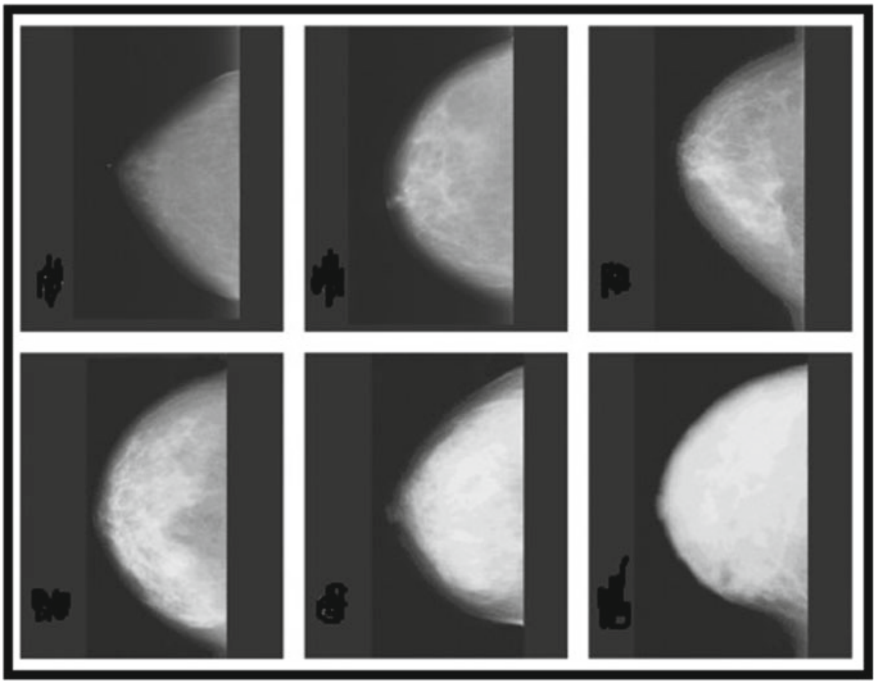

Fig. 6 Brightness augmentation images [14]

3.6 *Digital Mammograms Enhancement*

The CLAHE approach [15] is used to improve the contrast that some mammograms include, sometimes degraded, in some pictures. In proportion to the pixel intensity in the local intensity histogram, the intensity of a pixel converts into value within the display range. CLAHE is a case of adaptive histogram equalization (AHE) in which

Table 3 Dataset images after image augmentation [14]

Category number	All image	Training data	Testing data
Density-1 with Benign	874	701	173
Density-1 with Malignant	2170	1738	432
Density-2 with Benign	298	230	68
Density-2 with Malignant	2314	1843	471
Density-3 with Benign 1	946	739	207
Density-3 with Malignant	586	461	125
Density-4 with Benign	442	356	86
Density-1 with Benign	102	68	34
Total	7732	6136	1596

the images are improved to the highest contrast enhancement factor by the level of a consumer film. In this technique, improvements are made in small areas so that the over-improvement is very low compared with AHE due to noise or the effect of edge shadows.

Initially, the CLAHE method was established to decrease the shade and sound emitted by medical images in homogeneous areas [16]. The approach was used to develop digital mammograms and has shown good improvements in visual efficiency mammograms.

A small block input image I with M * N dimensions is separated. CLAHE is then used to increase each block's contrast. Bilinear interpolation is finally used to reconnect the next blocks to whole pictures. The steps mentioned in CLAHE are as follows.

(1) Patches of the images shall be divided into blocks of 8 * 8 size which are non-overlapping.
(2) Calculation of the histogram of any block.
(3) A histogram clip limit, t = 0.001, is set for enhancing patches in comparison.
(4) The histogram is redistributed after clipping the threshold value.
(5) The following transformation function modifies each block histogram:

$$\sum_{i=0}^{t} p_t(A_i) \tag{1}$$

where $p_t(A_i)$ is represent as the input patch image greyscale probability density function value and $p_t(A_i)$ is describe as

$$p_t(A_i) = \frac{m_i}{m} \tag{2}$$

where m_i is represent as the grey scale value of input pixel I and m is represent as the total sum of pixels in a block.

(6) Bilinear interpolation is used in any patch to association the next blocks. In the new histogram, the grey scale value of the patch is also modified.

We used the block size of 8 * 8 for our experiment, and the histogram clip limit is set to 0.001.

4 Convolution Neural Network for Classification Task

The relevance of CNN's findings has been shown in the classification of photographs. CNN has an architecture of multi-layered layer shadowed by a maximum layer of pooling. The sum of layers varies with the designer. A fully connected layer like MLP is fed the final maximum pooling layer output and then forwarded to Softmax.

The pooling layer is used to reduce the convolution layer's dimensionality. Average pooling, mean pooling, and full pooling are the most commonly used pooling layer algorithms. During preparation, a random disabling of the neurons is used for the discontinuation algorithm, usually with a 0.3–0.6 dropout ratio. The last layer of CNN is a soft max layer, which includes the output neuron by the sum of classes of the problem and is given a trust score.

The kernel sizes of 7 * 7 are used in both conv and max pooling layers. There are 16 kernels in the convolution layers, and 5 * 5 kernels in the second layer are included. Then the neural layer is completely linked. In the experiment the dropout ratio is 0.55. The layer of Softmax is used for classification CNN preparation. Figure 7 presents the complete network architecture of CNN.

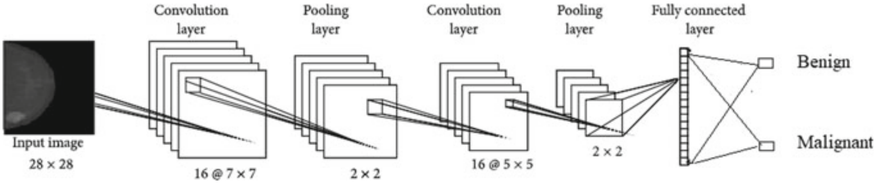

Fig. 7 Convolution neural network model

5 Simulation and Results

In this section, we discussed the performance of the CNN classifier by increase the input data set image by using data augmentation technique. In this simulation, experiment conducted by using software tool as python with system requirement of 4 GB RAM with 2 GHZ Intel i3 core processor. The validation of testing and training the data by using the breast cancer dataset image. The proposed system performance is estimated by using the different parametric metrics, which are explained in following section.

5.1 Performance Measures

The proposed system, in which different classifier performances is measure by using different parametric. The developed system is assessed using evaluation metrics such as TP, FP, TN, FN, sensitivity, precision, specificity, F-measure, and accuracy.

- TP—Sum of benign image is correctly categorized as noncancerous image.
- TN—Sum of malignant image is correctly categorized as cancerous image.
- FP—Sum of benign image is wrongly categorized as cancerous image.
- FN—Sum of malignant image is wrongly categorized as noncancerous image.

Sensitivity

Sensitivity is also called as recall. Sensitivity is distinct as the percentage of image with abnormal, whose output is positive and it is calculated using the Eq. 3 as

$$Sensitivity = TP/(TP + FN) \tag{3}$$

Specificity

Specificity, is defined as percentage of image with normal, whose output is negative and it is calculated using the Eq. 4 as

$$Specificity = TN/(TN + FP) \tag{4}$$

Classification Accuracy

Classification accuracy is defined as the sum of correctly classified images, which is separated by the total sum of images and then it is multiplied by 100 to turn it into a percentage. It is calculated using the Eq. 5 as

$$Classification\ Accuracy = (TP + TN)/(TP + FP + TN + FN) \times 100 \tag{5}$$

Precision

Precision is distinct as the sum of true positives, which is divided by the number of TP and false positives and it is calculated using the Eq. 6 as

$$Precision = TP/(TP + FP) \tag{6}$$

False Positive Rate

FPR is distinct as the sum of false positives, which is divided by the sum of false positives and true negative and it is calculated using the Eq. 7 as

$$FPR = FP/(FP + TN) \tag{7}$$

F-Measure

This is the kind of parameter measure, which association of recall and precision. The F-measure is determined by using the Eq. 8 as

$$F\text{-}measure = 2 * Recall * Precision/Recall + Precision \tag{8}$$

Mean Square Error (MSE)

Measure of fidelity of image. The parameter used to compare between the two images by providing quantitative or similarity rate. MSE calculation formula is expressed Eq. 9 as

$$MSE = \frac{1}{PQ} \sum \sum (f(i, j) - f^R(i, j))^2 \tag{9}$$

In Table 4, it represents that the performance of different augmentation technique with different parameter measures. In this analysis, the combination of entire augmentation technique achieved better performance than separate technique performance.

Table 4 Comparison analysis of different data augmentation technique with classification

Data augmentation methods	Precession (%)	Recall (%)	F-measure (%)	MSE (%)
Flipping	97.46	95.32	97.21	6.48
Cropping	96.40	95.59	96.44	5.67
Rotation	94.52	95.09	96.21	6.97
Noise injection	93.58	94.68	95.22	6.49
Random brightness	96.33	95.17	94.45	5.84
Combination of entire augmentation	98.49	97.92	98.64	4.63

Table 5 Comparison of classification accuracy with augmentation

Data augmentation methods	Accuracy (%)
Flipping	96.46
Cropping	95.21
Rotation	97.36
Noise injection	94.56
Random brightness	95.32
Combination of entire augmentation	98.91

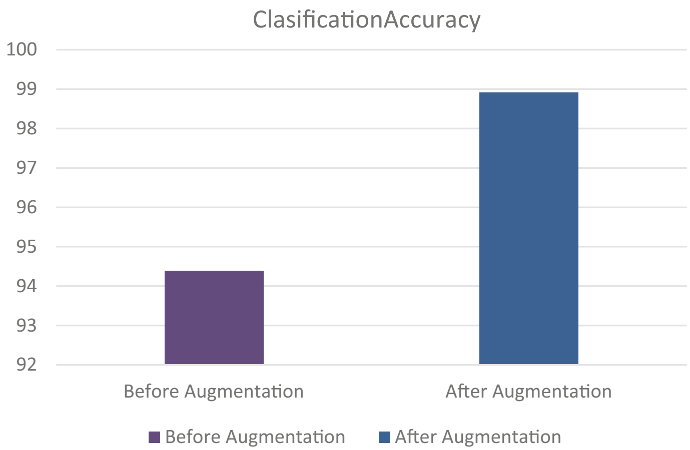

Fig. 8 Comparison of classification accuracy before and after augmentation

Table 5 represents that the accuracy performance of different augmentation technique; in flipping technique it reached 96.46%, cropping technique achieved 95.21% and rotation method reached 97.36% by the noise injection scheme it achieved the 94.565, which is the least value than other models. However, the combination of entire augmentation achieved the better classification accuracy of 98.91%, which is better accuracy performance than the other individual methods.

In Fig. 8, it shows the graphical representation of performance before and after augmentation. In without augmentation data technique, it achieved the least classification accuracy of 94.56%. Whereas, after the image data augmentation technique, the model achieved the better classification accuracy of 98.91%, respectively.

6 Conclusion

In this study, we did the analysis of learning ability of the training model. Using a small dataset, it may be poor due to the lack of potential useful learning information. Data augmentation techniques are implemented to increase the size of image data. In this augmentation technique, we included some methods as flipping, cropping, noise injection, rotate, and random brightness augmentation. In every technique, we separately analysed the classification accuracy by CNN classifier with different parametric measures. In CNN architecture, we implemented the proper kernel filter to achieve the maximum classification result. In this, we successfully create the maximum number of breast cancer images to train and test the model to achieve high classification accuracy as 98.91%. However, in this study the proposed model is to create or enhance image data to classify the tumour as two kinds such as benign or malignant. Further we design and implement hybrid architecture model to classify the breast images at various classification strategies. This work is further extended to clinically tested mammogram image samples supplied by VPS Lakeshore Hospital, Kochi, Kerala. We extend our sincere gratitude to the management and staff of the hospital.

References

1. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, Back T, Chesus M, Corrado GS, Darzi A, Etemadi M (2020) International evaluation of an AI system for breast cancer screening. Nature 577(7788):89–94
2. Pratheep Kumar P, Mary Amala Bai V, Geetha GN (2021) Challenges and solutions in CNN based diagnostic techniques on breast mammographic images: a survey. Int J Pharm Res 13(2)
3. Bahl M, Mercaldo S, Dang PA, McCarthy AM, Lowry KP, Lehman CD (2020) Breast cancer screening with digital breast tomosynthesis: are initial benefits sustained. Radiology 295(3):529–539
4. Shaikh TA, Ali R, Beg MS (2020) Transfer learning privileged information fuels CAD diagnosis of breast cancer. Mach Vis Appl 31(1):1–23
5. Karale VA, Singh T, Sadhu A, Khandelwal N, Mukhopadhyay S (2020) Reduction of false positives in the screening CAD tool for microcalcification detection. Sādhanā 45(1):1–12
6. Engstrom L, Tsipras D, Schmidt L, Madry A (2017) A rotation and a translation suffice: fooling CNNs with simple transformations. arXiv:171202779
7. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv:14091556
8. ImageNet (2018). http://image-net.org/. Accessed 05 Feb 2018
9. Wang J, Perez L (2017) The effectiveness of data augmentation in image classification using deep learning. Tech Rep 11(2):112–124
10. Jin J, Dundar A, Culurciello E (2015) Robust convolutional neural networks under adversarial noise. arXiv:151106306
11. Su J, Vargas DV, Kouichi S (2017) One pixel attack for fooling deep neural networks. arXiv:171008864
12. Mahmood M, Isa M, Ashidi (2015) A fully automated breast separation for mammographic images. In: 2015 IEEE international conference on biosignal analysis, processing and systems (ICBAPS), vol 12, no 3, pp 37–41

13. Zhong Z, Zheng L, Kang G, Li S, Yang Y (2017) Random erasing data augmentation. arXiv: 170804896
14. Huang ML, Lin TY (2020) Dataset of breast mammography images with masses. Data Brief 1(31):105–128
15. Khan SA, Hussain S, Yang S (2020) Contrast enhancement of low-contrast medical images using modified contrast limited adaptive histogram equalization. J Med Imaging Health Inf 10(8):1795–1803
16. Farhan AH, Kamil MY (2020) Texture analysis of mammogram using histogram of oriented gradients method. In: IOP conference series: materials science and engineering, vol 881, no 1. IOP Publishing, pp 21–49