Review

Image Augmentation Techniques for Mammogram Analysis

Parita Oza [1,2]*[0000-0001-9637-6157], **Paawan Sharma** [1], **Samir Patel** [1], **Festus Adedoyin**[3] **and Alessandro Bruno** [3]*[0000-0003-0707-6131]

1   Pandit Deendayal Energy University, Gandhinagar, India; parita.ophd19@sot.pdpu.ac.in,, Paawan.Sharma@sot.pdpu.ac.in, Samir.Patel@sot.pdpu.ac.in
2   Nirma University, Ahmedabad, India, parita.prajapati@nirmauni.ac.in
3   Bournemouth University, UK ; abruno@bournemouth.ac.uk, fadedoyin@bournemouth.ac.uk
*   Correspondence: parita.prajapati@nirmauni.ac.in; abruno@bournemouth.ac.uk

Abstract: Research in the medical imaging field using deep learning approaches has become progressively contingent. Scientific findings reveal that supervised deep learning methods' performance heavily depends on training set size, which expert radiologists must manually annotate. The latter is quite a tiring and time-consuming task. Therefore, most of the freely accessible biomedical image datasets are small-sized. Furthermore, it is challenging to have big-sized medical image datasets due to privacy and legal issues. Consequently, not a small number of supervised deep learning models are prone to overfitting and cannot produce generalized output. One of the most popular methods to mitigate the issue above goes under the name of data augmentation. This technique helps increase training set size by utilizing various transformations and has been publicized to improve the model performance when tested on new data. This article surveyed different data augmentation techniques employed on mammogram images. The study aims to provide insights into augmentation and deep learning-based augmentation techniques.

Keywords: Data Augmentation, Deep Learning, Medical Imaging, Mammograms

1. Introduction

Amongst various artificial intelligence fields, Deep Learning (DL) is widely adopted for the processing and analysis of radiological images. DL has been successfully applied to various Computer Vision tasks such as Object Segmentation, Detection, and Classification especially thanks to accuracy rates achieved by convolutional neural networks (CNNs). CNNs have the capabilities to automatically learn features through several network layers from a large set of labeled dataset [1]. Concerning the biomedical image analysis topic, CNNs have been successfully utilized for various tasks such as lesion or tumor classification, suspicious region detection, and abnormality detection [2–4]. DL-based solutions serve as a second opinion tool for expert radiologist and assist them in decision making and proper treatment planning [5]. To build a DL model that is able to generalise information from data, there needs a large amount of ground-truth data to avoid the model being very accurate only on the training dataset images. The latter goes under the name of overfitting [6,7] and represents a critical issue to overcome to have a model capable of delivering appropriate knowledge inference capabilities on a given application domain. Having high-quality and manually annotated data is a time-consuming and expert dependant task. That is quite common in the context of mammogram analysis [8–10]. One of the most challenging tasks for DL models is the generalisation, with generalisation being the capability of models to recognise those categories they were trained for on new data [11,12], The model with poor generalisation generally does not perform well due to high overfitting on the training set. Overfitting can be observed somehow in the plot showing validation accuracy at every epoch of the training phase [1]. Figure 1 shows the pictorial representation of models with and without overfitting. The training and validation loss
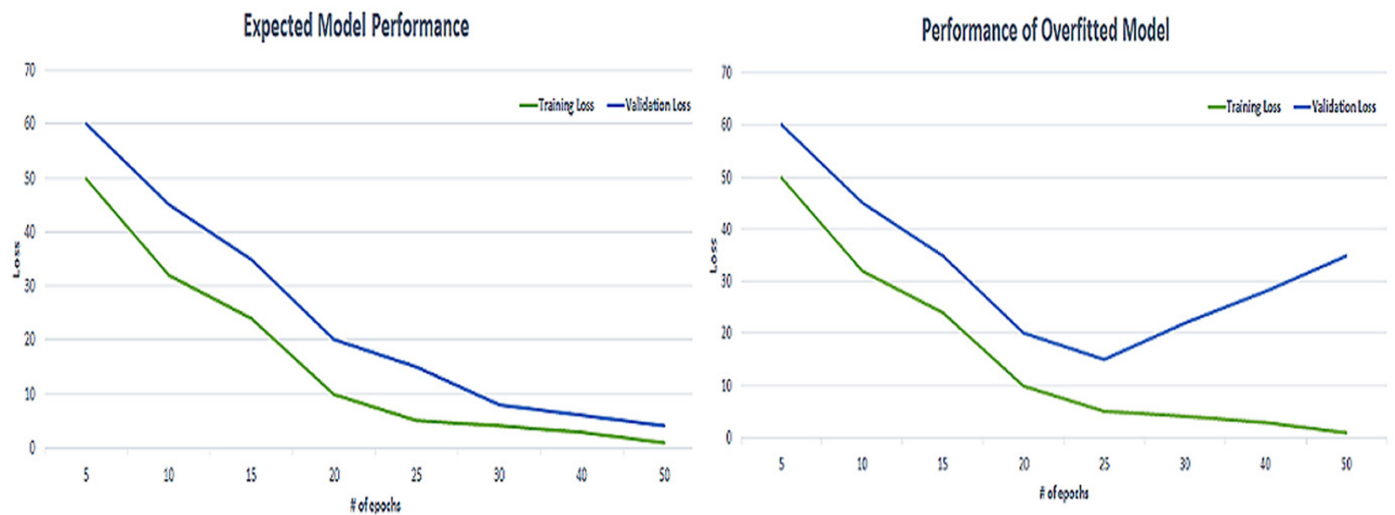
Figure 1. The left side plot shows the ideal trend of the model with training and validation error functions decreasing almost simultaneously. The plot on the right side shows the undesired effect of overfitting, having the training error decrease and, conversely, validation error increase suddenly.

curve is progressively and simultaneously reducing, which is a perfect circumstance, as shown in the figure 1 (left). The right side of the figure shows overfitting, in which the validation loss begins to grow after a certain number of epochs. In contrast, the training loss keeps decreasing. That is due to the model's inability to work effectively with unknown or new data. One of the reasons for this phenomenon might be a lack of enough training samples. The validation error of suitable DL models should continue to decrease along with the training error. Data augmentation methods can help achieve this task. Augmented data can characterise the inclusive set of input data points and minimise the distance between validation and training data. Data augmentation techniques apply alterations to training datasets to produce more samples. Moreover, this technique helps the model avoid learning features too specific to the original data, resulting in a more generalised model with improved performance on the test dataset. Class distribution imbalance in datasets is another common challenge. For instance, binary classification problems occur when one class (the minority class) holds considerably fewer samples than the other class (the majority class). Due to this, the model may get biased towards the majority class, possibly resulting in misclassification. Augmenting the minority class images may be used to mitigate the imbalance problem. Data augmentation is not the only approach to reduce the effect of overfitting and class imbalance. Other options for avoiding overfitting in Deep Learning models are also explored in the literature (see figure 2).

Batch Normalization: Batch Normalization can overcome the side-effect of overfitting by diminishing the internal covariate shift and instability in the distributions of Deeper networks' layer activations. For each mini-batch, batch normalisation standardises the inputs to a layer. This has the effect of bringing the learning process into balance. *Dropout:* Dropout applies during the training phase to get randomly selected neurons ignored. That avoids the so-called layer's "over-reliance" on a few inputs. As a consequence, It prevents neurons from co-adapting to training data. *Transfer Learning (TL):* Transfer learning improves models' performances on new and unknown data. The main point with TL is the employment of pre-trained models to be fine-tuned on a specific application domain using a small-sized dataset. *Pre-training:* Model pre-training is similar to that of TL, the only difference is here model architectures can be defined and weights are transferred. *Early-stopping:* It allows providing an arbitrarily large number of training epochs to suddenly stop training if the model does not perform well on the validation set.
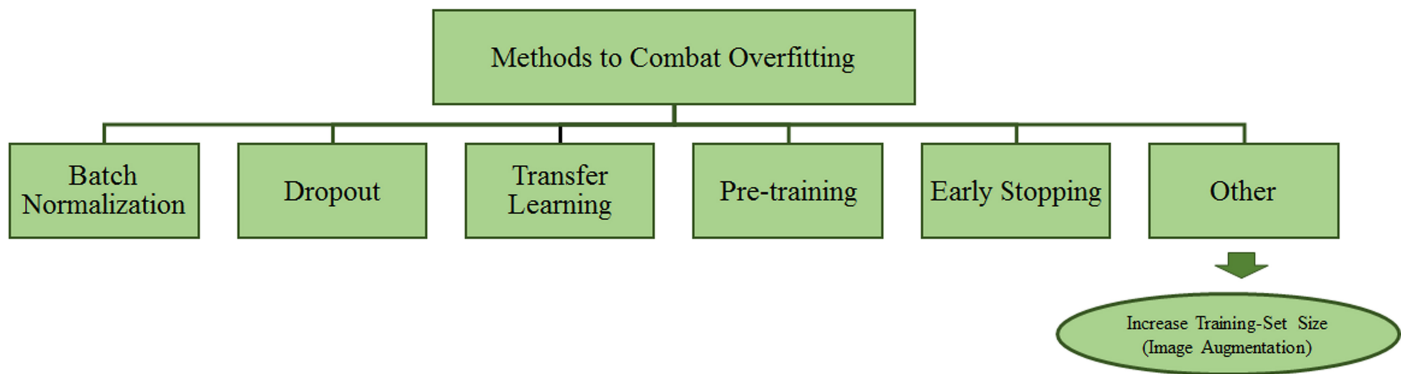
Figure 2. Methods to tackle overfitting

1.1. Research Contribution

Image augmentation techniques have been applied to mammogram datasets to increase the training set size, allowing data-hungry learners to benefit from more representative data. A review is conducted to summarise image augmentation techniques used in medical imaging applications such as deep learning-based breast cancer diagnosis. The two following main categories of image augmentation techniques are reviewed here: 1. Basic image augmentation techniques like geometric, colour space, intensity-based transformations, etc. 2. Deep learning-based augmentation as generative adversarial networks (GANs) and neural style transfer. The search terms used in the study are combinations of keywords such as "data augmentation", "image augmentation", "deep learning", "breast cancer", and "mammograms". Articles that do not utilize or discuss data/image augmentation were not considered in this study. The research mainly focuses on image augmentation for mammogram images. Therefore, articles whose subject is on other image modalities such as CT scan, Breast MRI, Breast ultrasounds, Histopathology, etc. are excluded. Articles on image augmentation used in the literature for breast image analysis applications are also summarised according to aspects such as dataset, model, technique, tasks performed, etc. This paper aims to give quick access to the research field and form an appropriate groundwork on the domain. This work examines several articles from various conferences, books and indexed journals out of scientific databases such as Scopus, IEEE, Web of science and PubMed. In the scientific literature, comprehensive and insightful surveys on image augmentation methods are present; some are specific to medical images. For example, the authors in [1] suggested several Data Augmentation solutions as ways to tackle models overfitting due to low-sized datasets. Another article [5] presents a thorough evaluation of the data augmentation methods employed in the broad topic of medical image analysis topic. In further detail, the authors focused on CT and MRI. Yet another article reports recent advancements in data-augmentation techniques for brain MRI [13] by examing the papers submitted to the Multimodal Brain Tumor Segmentation Challenge (BraTS 2018 edition [14]).

1.2. Paper Topology

The paper is structured as follows: Section 1 provides background and context for image augmentation within the broad topic of deep learning-based CAD system for medical imaging in . Section 2 delves into various image augmentation techniques used in practice. Deep learning-based image augmentation methods are showcased in section 3. Insights

into test-time augmentation are provided in section 4. Discussions and conclusions sections 5 and 6 respectively end the paper.

2. Basic Image Augmentation Techniques

Data augmentation encompasses a wide range techniques by inserting random variations into the existing training samples while preserving class labels. The purpose of data augmentation is to improve the model knowledge inference capability.

One of the most meaningful principles adopted in data augmentation relates to the physical phenomenon of vibrations and perturbations of a state. Perturbations take the form of slightly changed versions of images. Consequently, it increases the dataset size, allowing the network to infer knowledge from a more significant number of images. The network can learn more robust characteristics since it is constantly exposed to new, slightly modified copies of the input data. Therefore, when using deep learning in computer vision tasks, three types of data augmentation are the most likely; 1. Dataset generation and expansion. 2. On-the-fly data augmentation. 3. Amalgamation of Dataset generation and on-the-fly data augmentation. As widely covered in the scientific literature, supervised DL models [15] need a large amount of training data to unleash their knowledge inference capabilities fully. Therefore, exploring other paths in scenarios lacking data is necessary. In the worst-case scenario, only one image is available, and data augmentation comes into play to produce a complete image collection. The task is carried out by applying random transformations (rotation, flipping etc.) and other effects to the original image. Then, the newly generated images feed the DL model during the training phase. Methods like generation and expansion can generate N number of images. However, these approaches are not exempt from flaws: Using images produced by these methods, model's generalization ability is not improved rather only training set is increased by creating more examples and every newly created sample is based on minimal dataset.

On-the-fly data augmentation (sometimes also called in-place) is the second type of data augmentation [16]. On-the-fly data augmentation helps DL model training see new variations of images at each epoch. It takes image batches as input then applies a series of random transformations and other effects on each image in the batch. It finally returns a randomly altered image batch.

2.1. Geometric Transformations

In geometric transformation, an original image undergoes various transformations such as translation, rotation, scaling, flipping, or resizing to increase the training dataset size [5]. These conventional data augmentation techniques produce somewhat correlated images [17] and hence offer significantly less improvement to the model training and generalization over test data. However, these transformations lead to a significant increase in the training dataset; therefore, they are widely used in the domain [13]. This section presents the most commonly used geometric transformations for computer-aided breast cancer diagnosis.

Flipping: Flipping generates a mirror image of an image with both horizontal or vertical axes. The horizontal axis is more preferred over vertical flipping because the top and bottom parts of an image may not be interchangeable always [13]. However, flipping cannot always be a label-preserving transformation (e.g. MNIST dataset) [1]. In datasets such as DDSM and CBIS-DDSM, most of the breast profiles are on the left side of the mammograms. Making uniform direction of the breast in mammograms makes padding easier to perform during preprocessing steps. This section discusses some common methods of geometric transformations. The section also briefly discusses articles where these methods have been imparted to increase the training size to weaken the effect of overfitting in deep learning models for breast cancer diagnosis.

Rotation: Images are rotated leftward or rightward across an axis within the range [1°,359°]. The rotation angle determines the safety of this augmentation technique. The possibility of keeping the label post-transformation is known as a Data Augmentation

method's safety. The label of an image may no longer be preserved with an increase in rotation degree. For eg. rotation transformation is possibly safe on medical image datasets (X-ray, mammograms, Breast MRI etc..) as well as on images of other datasets like ImageNet [18], but not on images of 9 and 6 for digit identification task.

Translation: Translation applies to image augmentation to prevent positional bias [1]. This transformation translates the whole image by a given translation vector along a specific direction. It helps the network learn geographically invariant properties rather than focusing on features present in a single spatial location [13]. In the case of breast mammograms, translation of images can generate suitable augmented images. After the translation, padding, or pixel replication usually comes into play to fill out the leftover space. The process keeps the image dimensions [1].

Scaling: Scaled versions of images are added to the training set; the deep neural network can learn valuable deep features regardless of their original scale. Furthermore, scaling can be applied using scaling factors over different directions. For example, breast lesions may vary in size; this transformation can bring practical augmented images into the training dataset.

Figure 3 shows examples of geometric transformation applied to MIAS images.

Limited dataset size is one of the most common barriers in the medical research domain, and therefore the scientific literature provides a wide range of techniques to handle this issue. Costa et al. in [19] employed data augmentation to create new images based on their original clinical mammography dataset, and they compared the results using various CNN architectures. Authors have used geometric transformations such as rotation by varying degrees, flipping and adding Poisson noise. The model performs better when more regions of interest are added to the training step using data augmentation techniques. A new convolution neural network (CNN) model for identifying architectural distortion is proposed by Oyelade et al. in [8],, which uses data augmentation to improve its performance. Methods such as rotation, flipping, shearing and scaling are implemented and used to increase the training size. "Deep learning algorithms can improve performance by expanding their training set with synthetic examples." cha et al. practically proved that in [20]. Horizontal and Vertical Flipping methods were used by Omonigho et al. in the work [21] to augment the training set. By augmenting the training set with scaling, horizontal flip, rotation by degree 90°, 180°, 270°), authors could achieve 95.70% overall accuracy on the modified Alexnet model. In another study [22], Rahman et al. showed how specific pre-processing, transfer learning, and data augmentation approaches may help overcome the dataset size bottleneck in medical imaging applications.

Geometric transformations such as reflection, translation, random scaling and random rotations were applied to DDSM mammogram datasets. Shi et al. [23] implemented a customised CNN to classify BI-RADS [24] density of mammogram images. MIAS dataset was augmented using various transformations such as zooming, flipping, rotation and shifting. The authors carried out five-fold cross-validation of the model, which yielded an average test accuracy of 83.6%. Image augmentation can expand the size of training sets. Still, it is paramount to keep a certain level of variety between the images. Therefore, Khan et al. [25] developed a mammogram classification system and adopted random horizontal and vertical shifts, random shear and zoom as data augmentation techniques. Zhang et al. [26] performed data augmentation through reflection and rotation. Initially, each original image underwent horizontal flipping, then original and reflected images were rotated by 90°, 180°, and 270° degrees, respectively. As a result, the dataset increased eight times in size. The authors evaluated seven different architectures and concluded that models built and optimised using data augmentation and transfer learning had a lot of potential for automatic breast cancer detection. Bruno et al. [12] extracted patches from mammogram datasets such as MiniMIAS [27] and their own freely accessible dataset called SuREMaPP. Image transformations such as translations, horizontal reflections, and crop were employed in the study to generated augmented patches. Figure 4 shows an example of patch and augmented patches generated through various geometric transformations.
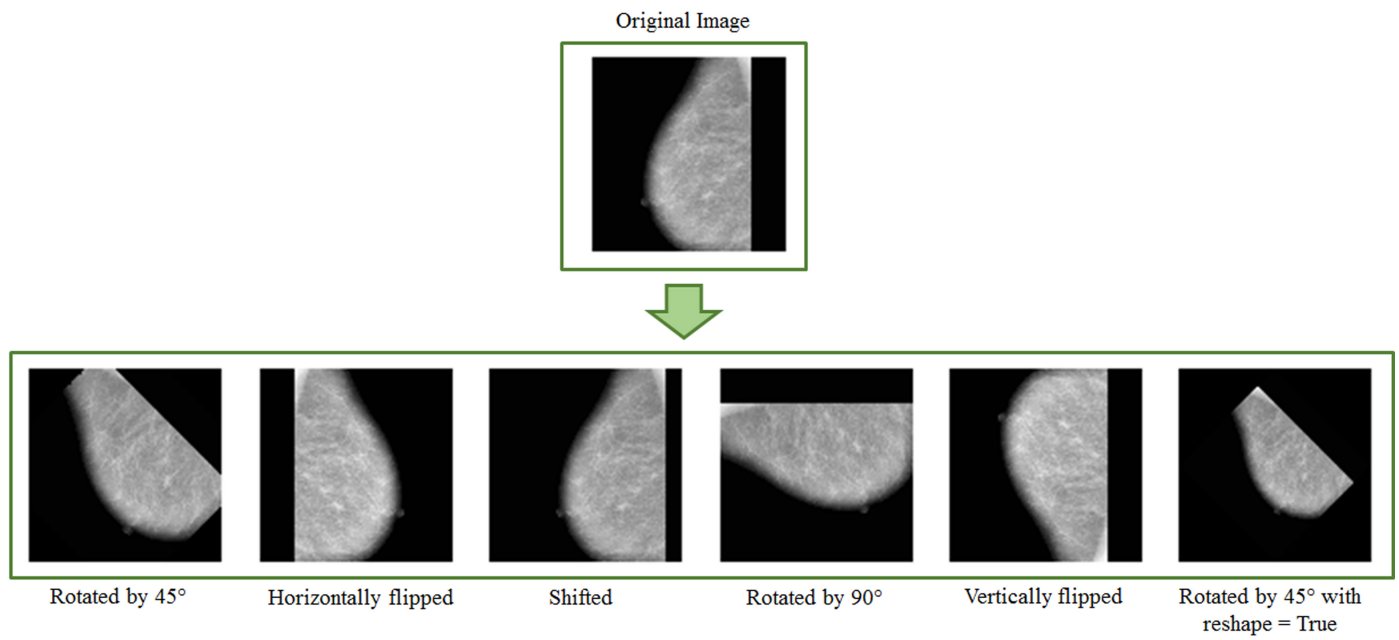
Figure 3. Example of images after applying geometric transformation
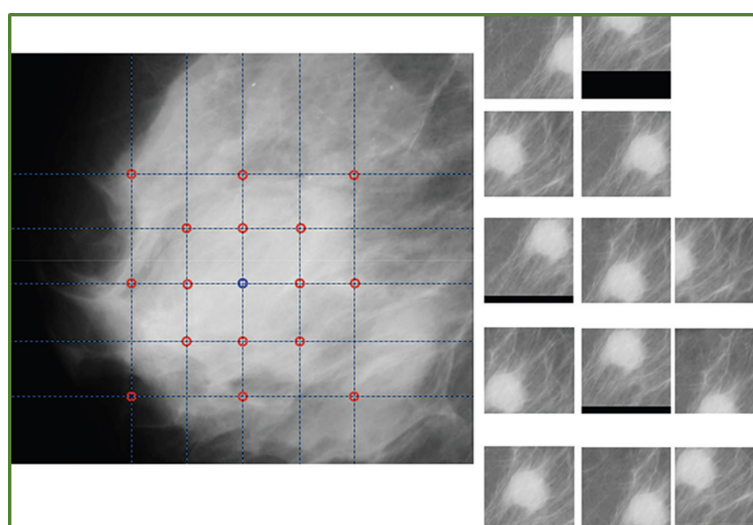

Figure 4. (a) An example of patch of Mammogram (b) A sample of patches generated with geometric transformation [12]

2.2. Pixel level Augmentation:

Position augmentation alters geometric shape of original images. However, some augmentation methods do not modify geometric properties of pictures but only pixel intensity values. This type of transformation is quite helpful for research in medical imaging fields, as medical images are obtained with several technologies and imaging modalities; hence they can be essentially assorted in pixel intensities [13]. In pixel-level augmentation, intensities of pixels are perturbed with random noise and a given probability, also called random intensity variation. In addition, a pixel-level augmentation modifies the brightness of an image. Among others, gamma correction (and all its variants), image blurring, image sharpening represent forms of pixel-level augmentation [28–30]. Another method that proved reliable applies to colour channels spaces. Isolating a single colour channel, such as R, G, or B, is the first step for colour augmentation consisting of deriving a colour histogram that describes the image allows further advanced colour augmentations. Mammograms are grayscale images. Grayscale mammograms are turned into pseudo-colour pictures to assess the effectiveness of Mask R-CNN. The latter is carried out using multi-scale morphological sifting, which boosts mass-like patterns. Mask R-CNN is then used with transfer learning to detect and segment masses on pseudo-colour images at the same time [31].

2.3. Other

Random erasing is another data augmentation technique [32]complementary to the previously described ones. The main goal of this technique is to make a model robust against occlusions in images. One of the most meaningful features is the learning phase being parameter-free.

Kernel filters rely on spatial filtering techniques to sharpen or smooth pixel values. It uses MxM size filtering masks. Along with transformation such as padding, flipping, and cropping Kang et al. [33] used a kernel filter swapping pixel values with a $n \times n$ sliding window. The experiments were carried out on four different datasets such as SVHN [34], MNIST [35], CIFAR-10 [36] ,and STL-10 [37]. A different Data Augmentation method involves combining images by averaging their pixel values. The visual results generated by this method are not easy to perceive to human eyes. Still, the main point is to generate further samples to have the model be able to generalise information from datawill[1]. However, various kernel filters such as gaussian blur, mean filter, median filter, Laplacian filter etc., can be employed for data augmentation purposes. Figure 5 shows some images generated with various kernel filters.

The enhanced dataset was utilised to train five state-of-the-art models using an augmentation strategy that increased both the size and variance of the dataset by Adedigba et al. in [38]. Along with geometric transformations, authors also have used Various other methods such as gaussian blurring and additions of white noise to augment the training set. It was shown that DensNet has achieved highest training and validation accuracy (99.01% and 99.99%, respectively). To increase and balance the available database at train time, artificially produced mammograms and data augmentation techniques are applied by Yemini et al. in [39]. The receiver operating characteristics (ROC) curve is used to assess the proposed scheme's performance. Along with flipping transformation, authors of this work used gaussian noise and also Changed image brightness to generate new images from the original samples.

3. Advanced Augmentation Techniques

Several methods to generate new data have been developed to overcome the issues associated with basic data augmentation procedures. Deep learning based advanced augmentation methods can generate synthetic images by learning representations of images. emphGenerative adversarial network (GAN) and its successive proposed variations represent the most widely described DL networks for data augmentation. GAN belongs to the family of unsupervised deep learning algorithms capable of extracting hidden underlying
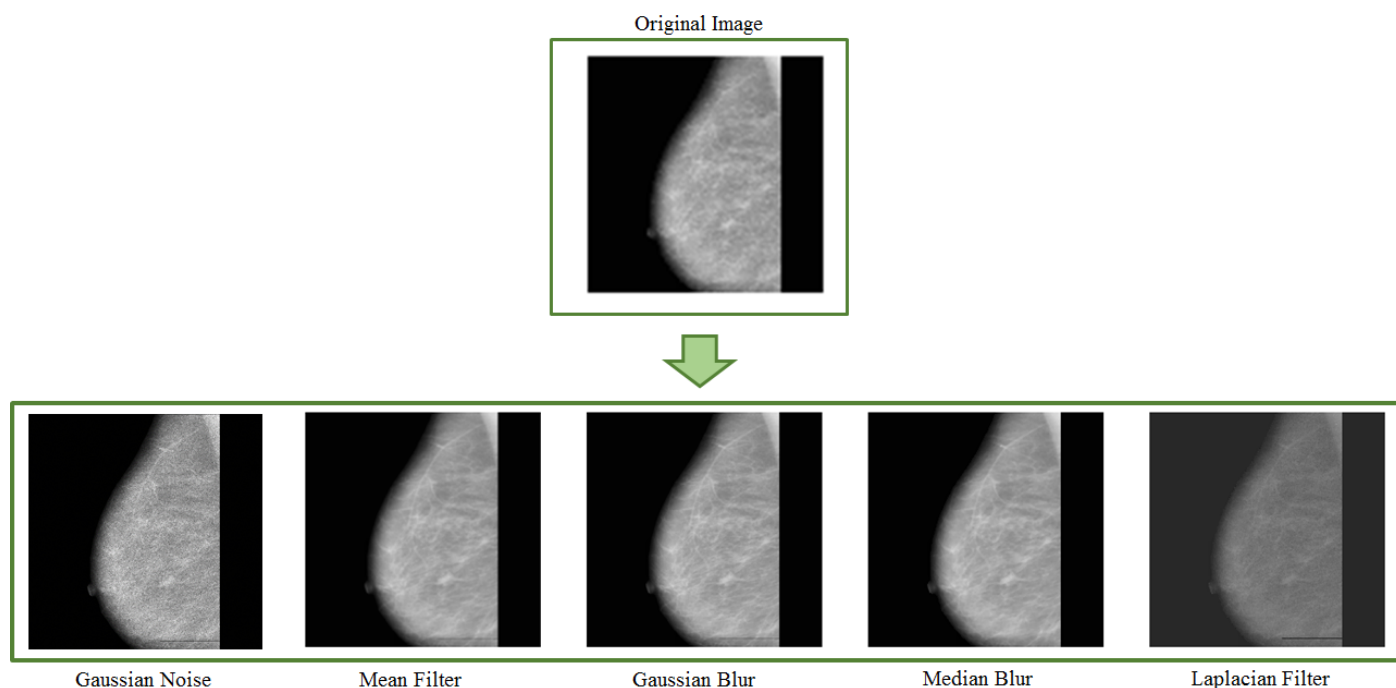
Figure 5. Example of data augmentation based on various filters and noise

properties from data and employing them in decision-making. To develop a more gener-alised model, these approaches are sometimes integrated with supervised algorithms. GAN is a widely used data augmentation approach to detect patterns and variances in image samples from the training dataset [40,41]. The fundamental goal of a GAN is to develop new image samples (by a generator) that the discriminator will not be able to tell apart from the original ones (Both these network branches compete against each other and gradually learn to produce better results.)[42]. In regards to mammogram augmentation, Shen et al. [43] provide a unique strategy based on GANs for generating varied mass images and then performing contextual infilling by inserting synthetic masses into normal mammograms. Furthermore, their system automatically annotates created mass from patches. Shen et al. [43] employed augmentation methods, which improved the performances of the detection technique. Some mammography pictures are transformed into mammogram images with mass findings, as shown in figure 6. Annotations with bounding-box label is also shown.

GANs have also been used in the literature of the domain for breast mass detection [41], mass classification [44] as well as mass segmentation [45].

Deep Learning proved effective even in mixing styles out of different images. Neural Style Transfer is meaningfully representative of the quality levels achieved on this [46]. The overall goal is to alter visual representations formed in CNNs [47]. *Neural Style Transfer* is well known for its uses in creative application domains, but it can also be used to augment data. The technique manipulates the sequential representations across a CNN to transfer the style of one image to another while keeping the original content [1]. Gatys et al. [48] first proposed NST, which typically takes two input images: a content image C to be transferred and a style reference image S, and then executes feature learning of the feature representations of Fl(C) and Fl(S) in layer l of a neural style transfer network [49]. However, if the image styles from different datasets are are way too far, it may cause a wide domain gap undermining deep learning models' capacities to target a specific scenario of interest. Wang et al. [49] proposed multi-resolution and multi-reference neural style transfer network to address the problem of style diversity in mammograms. With very high resolution, the network can normalize styles from several vendors (eg. GE healthcare (GE) and United Imaging Healthcare (UIH) to the same style baseline.
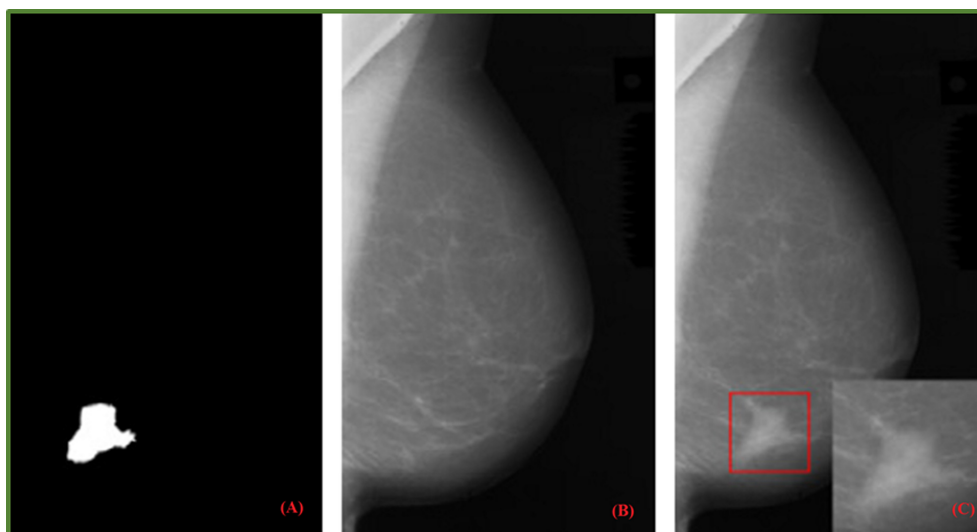
Figure 6. Given Mask Image (A) Normal mammogram image (B) Generated mammogram image with synthetic mask (C) [43]

Generative Adversarial Networska are reliant on two main components, namely, generator and discriminator. Scientific literature shows they are also used to learn noise augmentations. In adversarial training, one model classifies examples while another adds noise to them to deceive the classifier. The adversarial model is then given a loss function by the classification model, allowing it to improve itself to create better noise. Including images from adversarial training might help models acquire more robust features that are less sensitive with noise distortions. Although it has been proven that using adversarial search to inject noise improves performance on adversarial cases, it is unclear if this is effective for decreasing overfitting. That is still currently an open challenge, and has researchers investigate the link between adversarial attack resistance and actual performance on test datasets [1]. Adversarial training (including GAN-based and other adversarial learning networks) is employed by several DL-based augmentation systems [50,51]. In simple words, GANs help out generate new images from a given dataset. Nevertheless, the realistic level of artificially generated images for medical scenarios is still under investigation. It is still unclear if the resulting synthetic images can accurately depict realistic radiological characteristics in medical imaging [5].

A deep neural system to support mammogram tumor recognition was proposed in [52]. The authors used GANs to augment input images and having the model able to reconstruct test images correctly. The research was carried out utilising a large-sized database containing around 10000 mammographic images from the DDSM dataset. Wu et al. A class-conditional GAN (ciGAN) is trained to conduct contextual in-filling, which is subsequently used to synthesise lesions onto healthy screening mammograms in . Figure 7 shows samples generated using ciGAN for both cancerous to non-cancerous and non-cancerous to cancerous transformation. Examples of GAN synthesising a non-cancerous patch from a cancerous lesion can be seen in the first row. Conversely, the second row shows GAN synthesising a cancerous lesion on a non-cancerous patch using randomly selected segmentations from previous cancerous patches.

Swiderski et al. [52] remarked that a ResNet-50 classifier trained on GAN-augmented data could produce better AUROC than a model trained solely on traditionally augmented data. Another work in [20] investigated two main aspects: overfitting mitigation effectiveness of data augmentation with synthetic mammograms breast mass identification accuracy rate improvement. "In silico" procedural analytic breast and breast mass modelling algorithms were used to create synthetic mammograms. They were then projected into mammographic pictures using simulated X-ray projections. A novel approach for detecting abnormal and normal mammograms has been presented by Ramadan et al. [53]. They combined a cheat sheet containing standard features retrieved from the ROI with
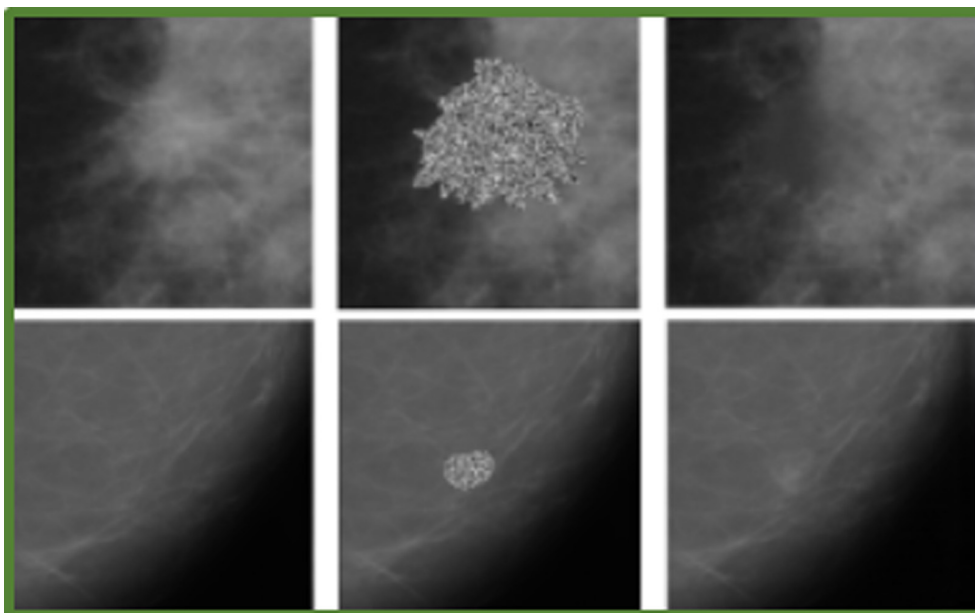
Figure 7. Samples generated by ciGAN:Each row comprises the original image, the ciGAN input, and the sample generated for the opposite class (from left to right). First row shows a transformation of cancerous patch to non-cancerous patch and second row shows transformation of non-cancerous patch to cancerous patch [9]

Data Augmentation to boost CNN performances in breast cancer detection. As a result, the accuracy rate improved by at least 12.2% and precision by at least 2.2%. Authors of [54] investigated conducted research on the capability of GANs to generate medical images as close to real ones as possible. Specialist doctors were involved in checking out GANs effectiveness on this task. Some promising results showed that current developments in GAN-based image synthesis might successfully apply to high-resolution medical imaging. Figure 8 provides examples of original and synthetic mammograms from [54].

Users may change or enrich existing datasets by effortlessly putting a genuine breast mass or micro-calcification cluster retrieved from a source digital mammography into a different place on another mammogram. This approach was used in [55] with the authors presenting findings of a reader experiment that compared the realism of inserted lesions to clinical lesions. Using the receiver operating characteristic (ROC) technique, radiologist ratings showed that injected lesions cannot be consistently discriminated from clinical lesions. Based on the identification of masses in the projections, Authors of [56] assessed the usage of data augmentation and the selection of non-overlapping areas of interest (ROI). Zhang et al. [26] combined data augmentation and transfer learning techniques with CNN models to improve the performance of the classifiers for mammogram images.

Table 1 summarizes basic and advanced augmentation techniques with their strength and limitations.

Table 2 presents a summary of methods that adopted image augmentation strategies to improve the model performance and counter overfitting.

4. Test-time Augmentation (TTA):

Over the last few years, a new image augmentation technique has increasingly caught researchers' interest. It goes under the name of TTA, standing for Test time Augmentation. Wang et al. [71] provided the scientific community with a mathematic formulation of TTA. They present TTA as an inference problem with hidden parameters and prior distributions. Therefore, images are considered as results of an elaboration process with hidden parameters. The final goal is to evaluate structure-wise uncertainty associated with image transformations and noise. Other than the previously mentioned techniques, TTA creates various augmented images of the test set, feeds these augmented images to the trained
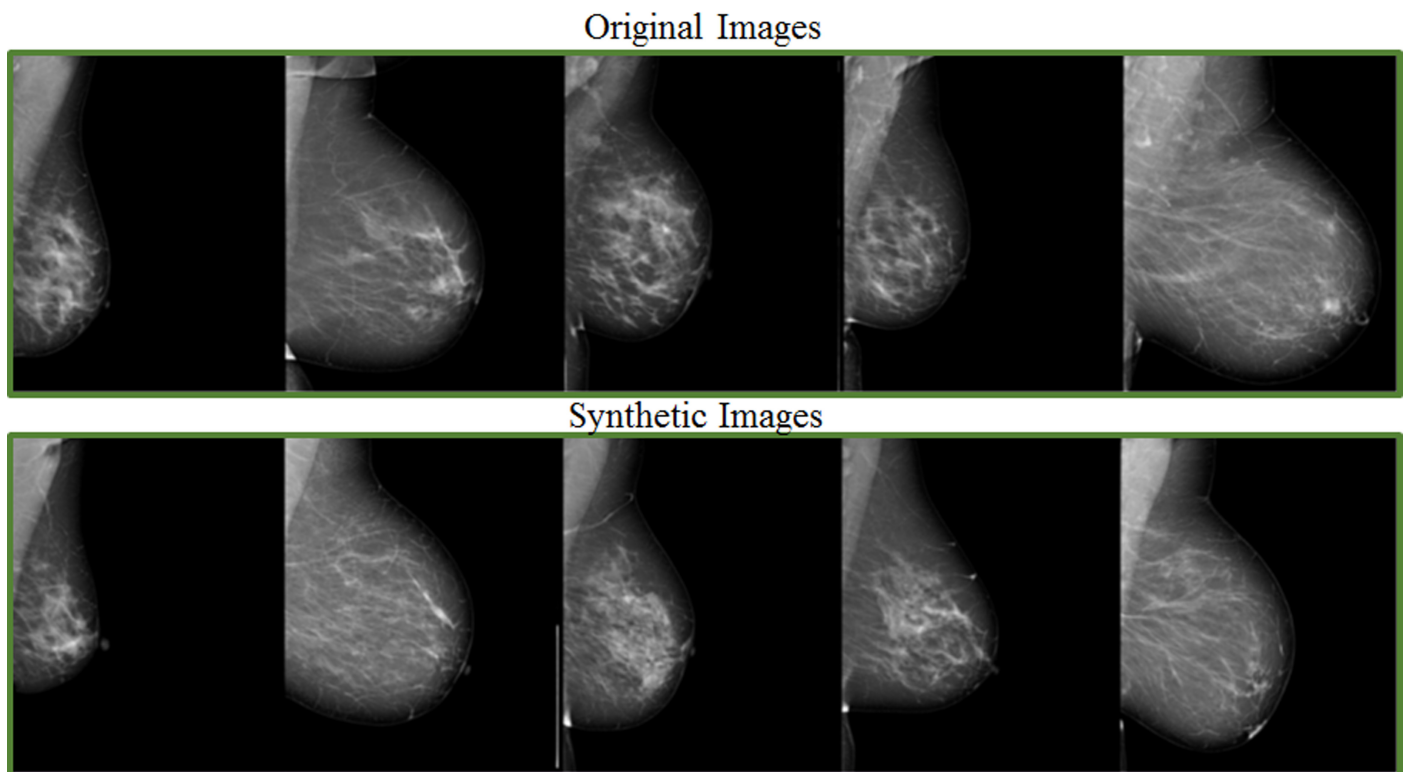
Figure 8. Randomly sampled examples of original and synthetic mammograms [54]

Table 1: Summary of Basic and Advanced Image Augmentation Techniques

Sr No	DA Technique	Sub Category	Label Preserving	Strength	Limitation
1	Geometric Transformation	Flipping	No	Good solutions for positional bias present in training data. Easy implementation	Additional memory, Transformation compute cost, Additional training time, Manual observation
		Cropping	Not always		
		Rotation	Not always		
		Translation	Yes		
2	Noise Injection	-	Yes	Allows model to learn more robust	Difficult to decide amount of noise to be added
3	Kernel Filters	-	Yes	Good to generate sharpen and blurred images	Similar to CNN mechanism
4	Mixing Images	-	No	-	Makes not much sense from human perspective. Not suitable for medical images
5	Random Erasing	-	Not always	Analogous to dropout regularization. Designed to combat image recognition challenges due to occlusion, A promising technique to guarantee a network pays attention to the entire image, not a subset of it	Some manual intervention may be necessary depending on the dataset and application
6	Adversarial Training	-	Yes	Help to illustrate weak decision boundaries better than standard classification metrics	Less explored
7	Generative Adversarial Training	-	Yes	GANs generate data that looks similar to original data	Harder to train, Generating results from text or speech is very complex.
8	Neural Style Transfer	-	-	Improves the generalization ability of simulated datasets	Efforts needed to select style, Additional memory, transformation cost

Table 2: Summary of articles using Image augmentation

| Ref	Task performed	Model	Dataset	Model Performance	Data Aumentation Approach
[8]	AD detection	Deep CNN and Deep CNN (Augmented CNN-SW+)	Private	AUC: 0.83 ± 0.14	Rotation by 90, 180 and 270 degrees, mirroring and adding Poisson noise
[8]	AD detection	Deep CNN	MIAS, DDSM, INBreast	Accuracy: 93.75 %	Rotation, flipping, shear, scaling etc.
[20]	Mass detection	Faster R-CNN	CBIS-DDSM	Sensitivity: 0.833 ± 0.038	Horizontal and Verticle Flipping
[49]	Mass detection	mr2NST	mammograms from GE and UIH	-	Neural Style Transfer
[57]	BI-RADS Classification	AlexNet	INBreast	Accuracy: 83.4	Image co-registration
[21]	Tumor detection	Modified AlexNet	MIAS	95.70%	Scaling, horizontal flip, rotation (90, 180, 270)
[22]	Mass Classification	InceptionV3 and ResNet50	DDSM	Accuracy: InceptionV3-79.6 ResNet50-85.71	Geometric Transformation
[58]	Mammogram classification	Pre-trained CNN Architectures	Private	-	Reflection and Rotation
[23]	BI-RADS classification	CNN	MIAS	Accuracy: 83.6%	Flip, rotation, shift and zoom
[38]	Mammogram Classification	Pre-trained CNN Architectures	MIAS	Accuracy:99.01%	Gaussian blurring, horizontal flipping, internal refection and mild addition of white noise
[39]	Mass detection	Google Inception-V3	INBreast	ROC: 0.86	Gaussian noise, Flipping, Changing image brightness
[59]	Mass Classification	VGG based DCNN	INBreast, CBIS, BCRP	-	elastic deformations
10	Mass Classification	DCNN	MIAS, INBreast, DDSM	Conventional DA techniques - 88% GAN- 94%	GAN
[60]	Mass Classification	AlexNet, InceptionV3	INBreast, CBIS-DDSM	Accuracy: INBreast: Alexnet-0.9892, InceptionV3-0.9919 CBIS-DDSM: Alexnet -0.6138, InceptionV3 0.8142	rotation, flipping,shearing
[10]	Tumor Classification	DCNN	MIAS, DDSM, INBreast	Accuracy-94%	GAN
[61]	Lesion Classification	ResNet50, VGG16, VGG19	CBIS-DDSM	Accuracy-90.4%	Geometric transformation, Contrast and brightness adjustment
[62]	Abnormality Classification	Meta Learning , REsnet101	CBIS-DDSM	Accuracy: Meta Learning-76%, Resnet101-71%	Geometric transformations
[25]	Mammogram Classification	VGGNet, GoogleNet, Resnet	CBIS-DDSM, MIAS	AUC-0.932	Geometric transformations
[63]	Mammogram Classification	Residual Networks	INBreast	Specificity-0.89	Rotation , Translation
[64]	Mass detection	InceptionV3	INBreast	ROC-0.91	Geometric transformations, Contrast and brightness adjustment,
[26]	Mammogram Classification	Alexnet, Resnet	Private	-	Geometric transformations
[65]	AD detection	Alexnet, SVM	CBIS-DDSM, DDSM, MIAS	Accuracy–92	Geometric transformations, TTA
[66]	Mammogram detection and classification	YOLO	INBreast	Accuracy-89.6	Rotation, Flipping
[67]	* Build datasets of breast mammography	Alexnet, Densenet, Shufflenet	INBreast	-	Rotation, Flipping
[68]	Mass Detection	Faster R-CNN	OMI-DB	TPR: 0.99 ± 0.03 at 1.17 FPI - malignant 0.85 ± 0.08 at 1.0 FPI - benign	Horizontal Flipping
[69]	Breast cancer diagnosis	Pre-trained CNN Architectures	CBIS- DDSM, BCDR, INBreast,MIAS	F1 Score for MIAS 0.907 ± 0.150	
[70]	Breast cancer classification	DCNN	MIAS	Accuracy-90.50	Feature wise data augmentation
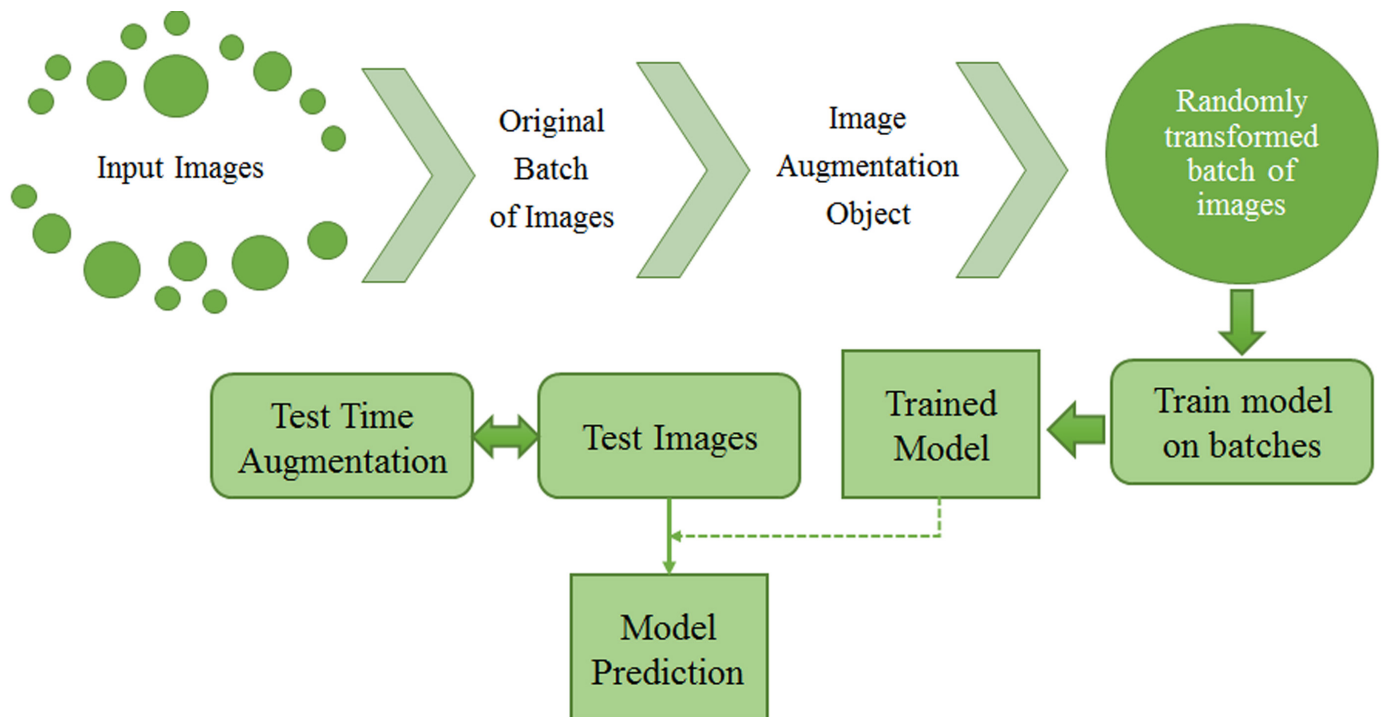
Figure 9. Train and test-time data augmentation.

model, and finally returns an ensemble of those predictions to get a more assertive response [72]. Figure 9 shows the process of both train and test time augmentation, while in figure 10 test-time data augmentation framework is depicted. TTA has conveyed new possibilities to the medical imaging field by measuring the strength, and network consistency as practical issues [73]. TTA can be used for those methods which modify an incoming example with affine, pixel-level, or elastic transformations in the case of lesion classification from mammograms. The research community has focused on training data augmentations, while data transformation before inference has yet to be fully explored. TTA combines numerous inference findings utilising various data augmentations to categorise one image (see figure 10). Kim et al. [72] presented a TTA method that is instance-aware and based on loss predictor. They improved image classification performance with the dynamic use of TTA transformations. The authors of [74] employed Test Time Augmentation for U-Net [75] to tackle medical image segmentation. Another study [76] employed TTA with the model making predictions on five, 224 × 224 image patches, as well as horizontally reflected patches (for a total of ten patches), and then averaging the outputs on over the ten patches with the softmax layer. An inference approach called Mixup Inference (MI), reliant on simple geometric intuitions, was proposed by Pang et al.[77]. The method mixes inputs with additional random samples. Vedalankar et al. [65] addressed the analysis of architectural distortion in mammograms with an integrated solution based on AlexNet and SVM. However, the solution heavily relies on TTA as the data augmentation technique on mammogram images.

5. Discussion

This section considers data augmentation and its employment in mammogram analysis and related tasks. The paper spans the main data augmentation approaches as listed in table 1. Most of them build on geometric transformations, noise injection, kernel filters, mixing images, random erasing, generative adversarial training, neural style transfer. In figures 3, 4 and 5 examples of basic geometric transformations and image filtering are given to show how simple operations allow increasing datasets' volumes. Conversely, when advanced image augmentation methods are tackled, things gradually start becoming more complex from theoretical and computational perspectives. In figures 6, 7 and 8,
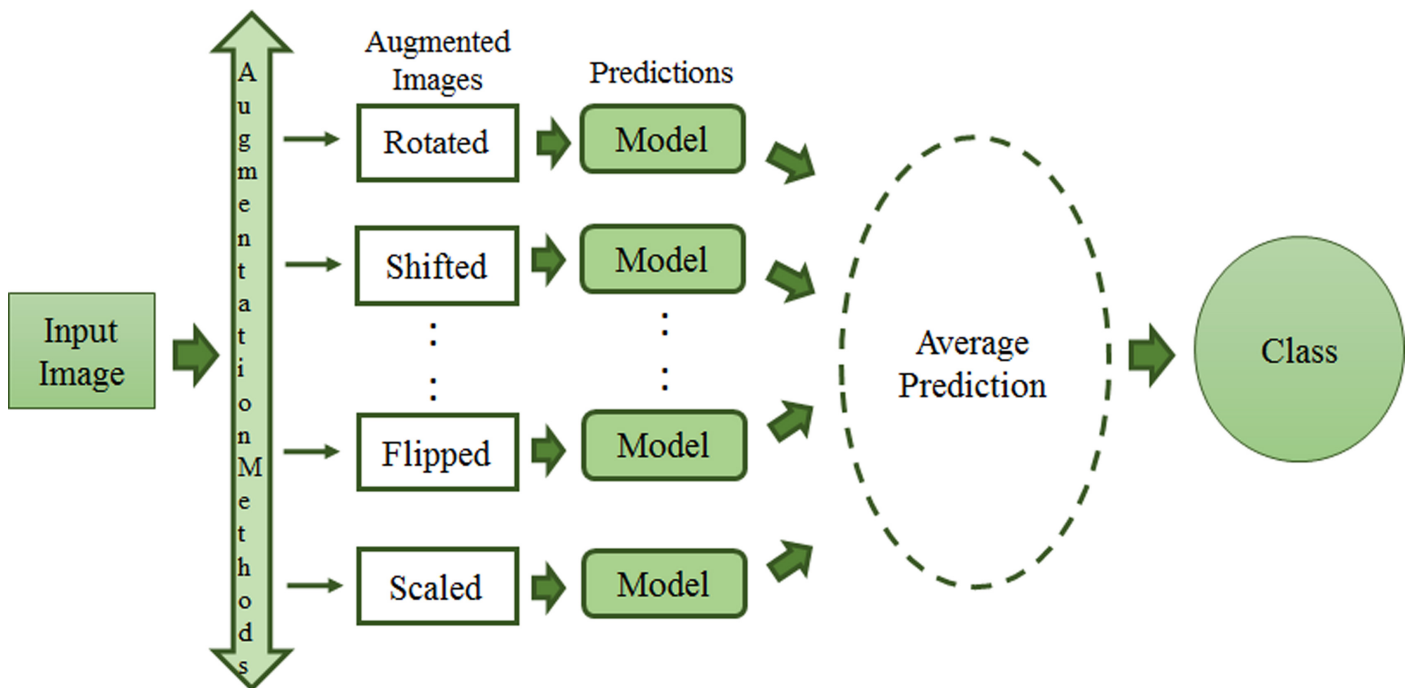
Figure 10. Test-time data augmentation framework

pictures respectively taken from Shen et al.'s[43], Wu et al.'s[9] and Korkinot et al.'s [54] demonstrate the level of refinement achieved by more recently introduced deep learning techniques such as GANs. It is easy to comment on how hard discriminating synthetic images from real images can be, especially to non-experts. Apart from the considerations mentioned above, it is necessary to span the performances of those models that heavily rely on data augmentation as shown in table 2. Around thirty methods tackling tasks such as mammogram classification, suspicious region segmentation and micro-calcification identification are compared according to several parameters. Although several methods achieve decent accuracy rates over several datasets, three main points are to be highlighted: 1. Oyelade and Ezugwu [8] achieved a 93.75% accuracy rate on anomaly detection from mammograms using a CNN-based technique and basic data augmentation techniques (rotation by 90, 180 and 270 degrees, mirroring and additive Poisson noise). 2. Conditional infilling GANs for data augmentation in mammogram classification by Dhivya et al. [10] averagely scored 94% accuracy over MIAS, INBreast and DDSM, which include images having different spatial resolution and acquiring device properties. The same method gets to an 88% accuracy rate when only basic data augmentation techniques are adopted. 3. Razali et al. [60] reached an excellent 99% accuracy rate on InBreast and DDSM with basic augmentation techniques on two datasets. However, it would be worth investigating any further improvement with advanced data augmentation techniques.

However, after spanning all methods in table 2, it is pretty noticeable how advanced mammogram augmentation impacts the not negligible accuracy rate improvement by 6% over three different datasets. Investigating all elements causing an increase in accuracy on a specific task is not trivial. Therefore, further experimental campaigns are needed to ensure fair comparisons having different tools trained with basic or advanced data augmentation.

6. Conclusions

This paper aims to provide insights into the broader area of the mammogram image analysis from a data augmentation perspective. The first sections introduce the main theoretical concepts in a more general sense. Then, around thirty methods tackling mammogram analysis tasks are compared in table 2 according to several parameters. Many approaches have been proposed on the topic over the last few years. Although some

performances are excellent, some further investigations are necessary to draw a line on the impact of data augmentation on the information generalisation capabilities of supervised deep learning paradigms. Some evidence shows a decisive increase in accuracy rates from basic to advanced augmentation techniques, especially the GANs-based ones. Current trends in computer vision see more new methods building on self-supervised and semi-self supervised paradigms. Purely supervised learning approaches combined with advanced data augmentation should run against self-supervised and semi self-supervised learning methods to balance computational costs, accuracy rates, and information generalisation capabilities.

Abbreviations

The following abbreviations are used in this manuscript:

CAD   Computer-Aided Diagnosis
BI-RADS   Breast Imaging Reporting and Database System
AI   Artificial Intelligence

References

1. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *Journal of Big Data* **2019**, *6*, 1–48.
2. Oza, P.; Sharma, P.; Patel, S. Machine learning applications for computer-aided medical diagnostics. Proceedings of Second International Conference on Computing, Communications, and Cyber-Security. Springer, 2021, pp. 377–392.
3. Pillai, R.; Oza, P.; Sharma, P. Review of machine learning techniques in health care. In *Proceedings of ICRIC 2019*; Springer, 2020; pp. 103–111.
4. Oza, P.; Sharma, P.; Patel, S.; Kumar, P. Deep convolutional neural networks for computer-aided breast cancer diagnostic: a survey. *Neural Computing and Applications* **2022**, pp. 1–22.
5. Chlap, P.; Min, H.; Vandenberg, N.; Dowling, J.; Holloway, L.; Haworth, A. A review of medical image data augmentation techniques for deep learning applications. *Journal of Medical Imaging and Radiation Oncology* **2021**.
6. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *nature* **2015**, *521*, 436–444.
7. Patel, H.J.; Oza, P.; Agrawal, S. AI Approaches for Breast Cancer Diagnosis: A Comprehensive Study. International Conference on Innovative Computing and Communications. Springer, 2022, pp. 393–419.
8. Oyelade, O.N.; Ezugwu, A.E. A deep learning model using data augmentation for detection of architectural distortion in whole and patches of images. *Biomedical Signal Processing and Control* **2021**, *65*, 102366.
9. Wu, E.; Wu, K.; Cox, D.; Lotter, W. Conditional infilling GANs for data augmentation in mammogram classification. In *Image analysis for moving organ, breast, and thoracic images*; Springer, 2018; pp. 98–106.

10. Dhivya, S.; Mohanavalli, S.; Karthika, S.; Shivani, S.; Mageswari, R. GAN based Data Augmentation for Enhanced Tumor Classification. 2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP). IEEE, 2020, pp. 1–5.

11. Oza, P.; Shah, Y.; Vegda, M. A Comprehensive Study of Mammogram Classification Techniques. In *Tracking and Preventing Diseases with Artificial Intelligence*; Springer, 2022; pp. 217–238.

12. Bruno, A.; Ardizzone, E.; Vitabile, S.; Midiri, M.; others. A novel solution based on scale invariant feature transform descriptors and deep learning for the detection of suspicious regions in mammogram images. *Journal of Medical Signals & Sensors* **2020**, *10*, 158.

13. Nalepa, J.; Marcinkiewicz, M.; Kawulok, M. Data augmentation for brain-tumor segmentation: a review. *Frontiers in computational neuroscience* **2019**, *13*, 83.

14. Bakas, S.; Reyes, M.; Jakab, A.; Bauer, S.; Rempfler, M.; Crimi, A.; Shinohara, R.T.; Berger, C.; Ha, S.M.; Rozycki, M.; others. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. *arXiv preprint arXiv:1811.02629* **2018**.

15. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data* **2021**, *8*, 1–74.

16. Rosebrock, A. Keras ImageDataGenerator and Data Augmentation. https://pyimagesearch.com/2019/07/imagedatagenerator-and-data-augmentation/, 2019.

17. Shin, H.C.; Tenenholtz, N.A.; Rogers, J.K.; Schwarz, C.G.; Senjem, M.L.; Gunter, J.L.; Andriole, K.P.; Michalski, M. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. International workshop on simulation and synthesis in medical imaging. Springer, 2018, pp. 1–11.

18. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248–255. doi:10.1109/CVPR.2009.5206848.

19. Costa, A.C.; Oliveira, H.C.; Vieira, M.A. Data augmentation: effect in deep convolutional neural network for the detection of architectural distortion in digital mammography. Proceedings of the 24. Brazilian congress on medical physics, 2019.

20. Cha, K.H.; Petrick, N.A.; Pezeshk, A.X.; Graff, C.G.; Sharma, D.; Badal, A.; Sahiner, B. Evaluation of data augmentation via synthetic images for improved breast mass detection on mammograms using deep learning. *Journal of Medical Imaging* **2019**, *7*, 012703.

21. Omonigho, E.L.; David, M.; Adejo, A.; Aliyu, S. Breast cancer: tumor detection in mammogram images using modified alexnet deep convolution neural network. 2020 International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS). IEEE, 2020, pp. 1–6.

22. Rahman, A.S.A.; Belhaouari, S.B.; Bouzerdoum, A.; Baali, H.; Alam, T.; Eldaraa, A.M. Breast mass tumor classification using deep learning. 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT). IEEE, 2020, pp. 271–276.

23. Shi, P.; Wu, C.; Zhong, J.; Wang, H. Deep learning from small dataset for BI-RADS density classification of mammography images. 2019 10th International Conference on Information Technology in Medicine and Education (ITME). IEEE, 2019, pp. 102–109.

24. Magny, S.J.; Shikhman, R.; Keppke, A.L. Breast, Imaging, Reporting and Data System (BI RADS). *StatPearls [Internet]* **2020**.

25. Khan, H.N.; Shahid, A.R.; Raza, B.; Dar, A.H.; Alquhayz, H. Multi-view feature fusion based four views model for mammogram classification using convolutional neural network. *IEEE Access* **2019**, *7*, 165724–165733.

26. Zhang, X.; Zhang, Y.; Han, E.Y.; Jacobs, N.; Han, Q.; Wang, X.; Liu, J. Whole mammogram image classification with convolutional neural networks. 2017 IEEE international conference on bioinformatics and biomedicine (BIBM). IEEE, 2017, pp. 700–704.

27. SUCKLING J, P. The mammographic image analysis society digital mammogram database. *Digital Mammo* **1994**, pp. 375–386.

28. Agarwal, M.; Mahajan, R. Medical images contrast enhancement using quad weighted histogram equalization with adaptive gama correction and homomorphic filtering. *Procedia computer science* **2017**, *115*, 509–517.

29. Sahnoun, M.; Kallel, F.; Dammak, M.; Mhiri, C.; Mahfoudh, K.B.; Hamida, A.B. A comparative study of MRI contrast enhancement techniques based on Traditional Gamma Correction and Adaptive Gamma Correction: Case of multiple sclerosis pathology. 2018 4th International

Conference on Advanced Technologies for Signal and Image Processing (ATSIP). IEEE, 2018, pp. 1–7.

30. Galdran, A.; Alvarez-Gila, A.; Meyer, M.I.; Saratxaga, C.L.; Araújo, T.; Garrote, E.; Aresta, G.; Costa, P.; Mendonça, A.M.; Campilho, A. Data-driven color augmentation techniques for deep skin image analysis. *arXiv preprint arXiv:1703.03702* **2017**.

31. Min, H.; Wilson, D.; Huang, Y.; Liu, S.; Crozier, S.; Bradley, A.P.; Chandra, S.S. Fully automatic computer-aided mass detection and segmentation via pseudo-color mammograms and mask r-cnn. 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). IEEE, 2020, pp. 1111–1115.

32. Zhong, Z.; Zheng, L.; Kang, G.; Li, S.; Yang, Y. Random erasing data augmentation. Proceedings of the AAAI conference on artificial intelligence, 2020, Vol. 34, pp. 13001–13008.

33. Kang, G.; Dong, X.; Zheng, L.; Yang, Y. Patchshuffle regularization. *arXiv preprint arXiv:1707.07103* **2017**.

34. Netzer, Y.; Wang, T.; Coates, A.; Bissacco, A.; Wu, B.; Ng, A.Y. Reading Digits in Natural Images with Unsupervised Feature Learning **2011**.

35. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **1998**, *86*, 2278–2324.

36. Krizhevsky, A.; Hinton, G.; others. Learning multiple layers of features from tiny images **2009**.

37. Coates, A.; Ng, A.; Lee, H. An analysis of single-layer networks in unsupervised feature learning. Proceedings of the fourteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2011, pp. 215–223.

38. ADEDIGBA, A.P.; ADESHINAT, S.A.; AIBINU, A.M. Deep learning-based mammogram classification using small dataset. 2019 15th International Conference on Electronics, Computer and Computation (ICECCO). IEEE, 2019, pp. 1–6.

39. Yemini, M.; Zigel, Y.; Lederman, D. Detecting masses in mammograms using convolutional neural networks and transfer learning. 2018 IEEE International Conference on the Science of Electrical Engineering in Israel (ICSEE). IEEE, 2018, pp. 1–4.

40. Shen, L.; Margolies, L.R.; Rothstein, J.H.; Fluder, E.; McBride, R.; Sieh, W. Deep learning to improve breast cancer detection on screening mammography. *Scientific reports* **2019**, *9*, 1–12.

41. Desai, S.D.; Giraddi, S.; Verma, N.; Gupta, P.; Ramya, S. Breast Cancer Detection Using GAN for Limited Labeled Dataset. 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), 2020, pp. 34–39. doi:10.1109/CICN49253.2020.9242551.

42. Oza, P.; Sharma, P.; Patel, S.; Bruno, A. A Bottom-Up Review of Image Analysis Methods for Suspicious Region Detection in Mammograms. *Journal of Imaging* **2021**, *7*. doi:10.3390/jimaging7090190.

43. Shen, T.; Hao, K.; Gou, C.; Wang, F.Y. Mass Image Synthesis in Mammogram with Contextual Information Based on GANs. *Computer Methods and Programs in Biomedicine* **2021b**, *202*, 106019. doi:https://doi.org/10.1016/j.cmpb.2021.106019.

44. S., D.; S, M.; S, K.; S, S.; R, M. GAN based Data Augmentation for Enhanced Tumor Classification. 2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP), 2020, pp. 1–5. doi:10.1109/ICCCSP49186.2020.9315189.

45. Singh, V.K.; Rashwan, H.A.; Romani, S.; Akram, F.; Pandey, N.; Sarker, M.M.K.; Saleh, A.; Arenas, M.; Arquez, M.; Puig, D.; Torrents-Barrena, J. Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network. *Expert Systems with Applications* **2020**, *139*, 112855. doi:https://doi.org/10.1016/j.eswa.2019.112855.

46. Jing, Y.; Yang, Y.; Feng, Z.; Ye, J.; Yu, Y.; Song, M. Neural Style Transfer: A Review. *IEEE Transactions on Visualization and Computer Graphics* **2020**, *26*, 3365–3385. doi: 10.1109/TVCG.2019.2921336.

47. Gatys, L.A.; Ecker, A.S.; Bethge, M. A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576* **2015**.

48. Gatys, L.A.; Ecker, A.S.; Bethge, M. Image style transfer using convolutional neural networks. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2414–2423.

49. Wang, S.; Huo, J.; Ouyang, X.; Che, J.; Xue, Z.; Shen, D.; Wang, Q.; Cheng, J.Z. mr2 NST: Multi-resolution and Multi-reference Neural Style Transfer for Mammography. International Workshop on PRedictive Intelligence In MEdicine. Springer, 2020, pp. 169–177.

50. Gao, M.; Fessler, J.A.; Chan, H.P. Deep Convolutional Neural Network with Adversarial Training for Denoising Digital Breast Tomosynthesis Images. *IEEE Transactions on Medical Imaging* **2021**.

51. Shen, R.; Yao, J.; Yan, K.; Tian, K.; Jiang, C.; Zhou, K. Unsupervised domain adaptation with adversarial learning for mass detection in mammogram. *Neurocomputing* **2020**, *393*, 27–37.
52. Swiderski, B.; Gielata, L.; Olszewski, P.; Osowski, S.; Kołodziej, M. Deep neural system for supporting tumor recognition of mammograms using modified GAN. *Expert Systems with Applications* **2021**, *164*, 113968.
53. Ramadan, S.Z. Using convolutional neural network with cheat sheet and data augmentation to detect breast cancer in mammograms. *Computational and Mathematical Methods in Medicine* **2020**, *2020*.
54. Korkinof, D.; Heindl, A.; Rijken, T.; Harvey, H.; Glocker, B. MammoGAN: High-resolution synthesis of realistic mammograms. International Conference on Medical Imaging with Deep Learning–Extended Abstract Track, 2019.
55. Pezeshk, A.; Petrick, N.; Sahiner, B. Seamless lesion insertion in digital mammography: methodology and reader study. Medical Imaging 2016: Computer-Aided Diagnosis; Tourassi, G.D.; III, S.G.A., Eds. International Society for Optics and Photonics, SPIE, 2016, Vol. 9785, pp. 134 – 139. doi:10.1117/12.2217056.
56. Zeiser, F.A.; da Costa, C.A.; Zonta, T.; Marques, N.; Roehe, A.V.; Moreno, M.; da Rosa Righi, R. Segmentation of masses on mammograms using data augmentation and deep learning. *Journal of digital imaging* **2020**, *33*, 858–868.
57. Domingues, I.; Abreu, P.H.; Santos, J. Bi-rads classification of breast cancer: a new pre-processing pipeline for deep models training. 2018 25th IEEE international conference on image processing (ICIP). IEEE, 2018, pp. 1378–1382.
58. Zhang, X.; Zhang, Y.; Han, E.Y.; Jacobs, N.; Han, Q.; Wang, X.; Liu, J. Classification of whole mammogram and tomosynthesis images using deep convolutional neural networks. *IEEE transactions on nanobioscience* **2018**, *17*, 237–242.
59. Castro, E.; Cardoso, J.S.; Pereira, J.C. Elastic deformations for data augmentation in breast cancer mass detection. 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). IEEE, 2018, pp. 230–234.
60. Razali, N.F.; Isa, I.S.; Sulaiman, S.N.; Karim, N.K.A.; Osman, M.K. High-level Features in Deeper Deep Learning Layers for Breast Cancer Classification. 2021 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE). IEEE, 2021, pp. 170–175.
61. Alkhaleefah, M.; Chittem, P.K.; Achhannagari, V.P.; Ma, S.C.; Chang, Y.L. The influence of image augmentation on breast lesion classification using transfer learning. 2020 International Conference on Artificial Intelligence and Signal Processing (AISP). IEEE, 2020, pp. 1–5.
62. Wang, Y.; Song, M.; Tian, X. Meta-learning Based Breast Abnormality Classification on Screening Mammograms. 2021 International Conference on Computer Engineering and Application (ICCEA). IEEE, 2021, pp. 153–156.
63. Bagchi, S.; Mohd, M.N.H.; Debnath, S.K.; Nafea, M.; Suriani, N.S.; Nizam, Y. Performance Comparison of Pre-trained Residual Networks for Classification of the Whole Mammograms with Smaller Dataset. 2020 IEEE Student Conference on Research and Development (SCOReD). IEEE, 2020, pp. 368–373.
64. Lu, H.C.; Loh, E.W.; Huang, S.C. The classification of mammogram using convolutional neural network with specific image preprocessing for breast cancer detection. 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD). IEEE, 2019, pp. 9–12.
65. Vedalankar, A.V.; Gupta, S.S.; Manthalkar, R.R. Addressing architectural distortion in mammogram using AlexNet and support vector machine. *Informatics in Medicine Unlocked* **2021**, *23*, 100551.
66. Aly, G.H.; Marey, M.; El-Sayed, S.A.; Tolba, M.F. YOLO based breast masses detection and classification in full-field digital mammograms. *Computer methods and programs in biomedicine* **2021**, *200*, 105823.
67. Huang, M.L.; Lin, T.Y. Dataset of breast mammography images with masses. *Data in brief* **2020**, *31*, 105928.
68. Agarwal, R.; Díaz, O.; Yap, M.H.; Llado, X.; Marti, R. Deep learning for mass detection in Full Field Digital Mammograms. *Computers in biology and medicine* **2020**, *121*, 103774.
69. Chougrad, H.; Zouaki, H.; Alheyane, O. Multi-label transfer learning for the early diagnosis of breast cancer. *Neurocomputing* **2020**, *392*, 168–180.
70. Ting, F.F.; Tan, Y.J.; Sim, K.S. Convolutional neural network improvement for breast cancer classification. *Expert Systems with Applications* **2019**, *120*, 103–115.

71. Wang, G.; Li, W.; Aertsen, M.; Deprest, J.; Ourselin, S.; Vercauteren, T. Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. *Neurocomputing* **2019**, *338*, 34–45.

72. Kim, I.; Kim, Y.; Kim, S. Learning loss for test-time augmentation. *arXiv preprint arXiv:2010.11422* **2020**.

73. Wang, G.; Li, W.; Ourselin, S.; Vercauteren, T. Automatic brain tumor segmentation using convolutional neural networks with test-time augmentation. International MICCAI Brainlesion Workshop. Springer, 2018, pp. 61–72.

74. Isensee, F.; Petersen, J.; Klein, A.; Zimmerer, D.; Jaeger, P.F.; Kohl, S.; Wasserthal, J.; Koehler, G.; Norajitra, T.; Wirkert, S.; others. nnu-net: Self-adapting framework for u-net-based medical image segmentation. *arXiv preprint arXiv:1809.10486* **2018**.

75. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Springer, 2015, pp. 234–241.

76. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* **2012**, *25*, 1097–1105.

77. Pang, T.; Xu, K.; Zhu, J. Mixup inference: Better exploiting mixup to defend adversarial attacks. *arXiv preprint arXiv:1909.11515* **2019**.

Short Biography of Authors

Parita Oza is working as an Assistant Professor in Computer Science and E Institute of Technology, Nirma University, Ahmedabad. She has pursued he and Communication Technology from Nirma University. She is involved in undergraduate and postgraduate level. She has published several research international conferences and journals. She is pursuing her PhD from Pa University (Gandhinagar, India). Her research area includes Image Processin Medical Imaging. She also serves as a reviewer, technical program committ chair for international conferences.

Dr Paawan Sharma received his PhD (Engineering) from Homi Bhabha Nat M.Tech. (Communication Systems) from SVNIT, Surat and B.E. (ECE) fro jasthan, Jaipur. His research area is multi-disciplinary, spanning application in various domains such as signal processing, embedded systems, patter vision, and Artificial Intelligence. He has 3 Indian patents (published) and m He has guided/co-guided 3 PhD Scholars focusing on disaster managemen grid analysis. He has also worked in Wipro Technologies in the VLSI domain IEEE and a Member of ACM, ISSIA, IAPR.

Dr Samir Patel is currently an Associate professor and Head of the Departm and Engineering at Pandit Deendayal Energy University, Gandhinagar, Ind degree from Nirma University in October 2012. His area of interest inclu Parallel computing, Data Mining and Image processing. He has over 22 years experience and has published 34 papers in National/International conferenc he has authored a book on "An Integration of Image processing and Data M Watermarking". He is awarded the pedagogical Innovation award (PIA-2015 council. He has got the research interest to work on different funded projects Machine Learning.

Dr Alessandro Bruno is a Lecturer in Computing at Bournemouth Univ
Computing and Informatics (Poole, United Kingdom). He is serving as Co-Inv
research project. He received his PhD in Computer Engineering in 2012 fr
(Italy), defending a thesis titled 'Advanced Texture and Keypoints Analys
Inspections'. His research activities mainly focus on Computer Vision, P
Artificial Intelligence. He worked as a postdoctoral research fellow at Pale
(Computer Vision and Image Processing) group, INAF (Italian National Ins
Bournemouth University's Department of Computing and Informatics (UI
a research visitor at the UCL (University College London) Mullard Space
addition, he worked as a Research Associate at NCCA (National Centre fo
from Bournemouth University. He also serves as a member of the prog
international scientific conferences, academic and guest editor for internatio