RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt2619

Spatial Transformer Network with Deep Features for Improved Breast Cancer Image Classification

Talari Swapna [a], V. Vijaya Kumar [a, *]

[a] Department of Computer Science and Engineering, Anurag University, TS, India

* Corresponding Author Email: drvvk144@gmail.com

DOI: https://doi.org/10.54392/irjmt2619

Received: 07-09-2025; Revised: 09-01-2026; Accepted: 16-01-2026; Published: 27-01-2026

Abstract: Breast cancer continues to be an utmost worldwide health issue, accounting for a notable number of cancer-related deaths among women every year. premature and precise identification is vital to improve prognosis and reduce mortality. However, automated classification of mammographic images is challenged by low-resolution data, spatial distortions, and the lack of fine-grained texture characteristics critical for identifying malignant tissues. In this research, we present a revolutionary deep learning framework that conveys these limitations through an integrated architecture combining a Super-Resolution Generative Adversarial Network (SRGAN), a Spatial Transformer Network (STN), and a ResNet50-based image classifier. A significant advancement in our architecture is the incorporation of embedding Bilinear Interpolation (BI) and Lanczos interpolation (LZ) within the STN module. Unlike traditional fixed-kernel Lanczos filters, our learnable version enables the network to adapt interpolation weights during training, allowing it to better handle spatial variations and preserve critical anatomical details. The SRGAN enhances the resolution of mammographic images, aiding in improved feature extraction, while the ResNet50 backbone performs classification with refined spatial awareness. The enhanced images are then passed through the STN for spatial normalization, followed by classification using a fine-tuned ResNet50 backbone. We evaluated the proposed pipeline on the MIAS (Mammographic Image Analysis Society) and Digital Database for Screening Mammography (DDSM) dataset, a benchmark in breast cancer imaging research. The model demonstrated superior classification performance, achieving 97.08% accuracy, 95.91% precision, 95.08% specificity, and 95.48% f-measure on DDSM dataset and 96.58% accuracy, 94.62% specificity, 94.59% sensitivity, and an f-measure of 94.59% on MIAS datatset. These outcomes show a significant improvement over comparable approaches that employ traditional interpolation or omit spatial correction mechanisms.

Keywords: Breast Cancer (BC) Diagnosis, Mammographic Image Classification, Deep Learning, Learnable Interpolation, Super-Resolution, Spatial Transformation, Resnet50.

1. Introduction

Globally, breast cancer is considerably widespread malignancies in females [1, 2]. Recent investigations indicate that it is the second most common type of cancer, behind lung cancer [3, 4]. It is a most common cancer in emerging nations. In order to mitigate the death rate, early-stage preventive interventions are desperately needed. However, inadequate resources and human error raise the possibility of an inaccurate diagnosis. Recent developments in computer vision and Deep learning (DL) have aided in automating the process of distinguishing benign from malignant instances ones that include cancerous cells [5]. For this reason, scientists have developed mammography to identify breast cancer early on. Prior research [6, 7] indicates that mammography lowers mortality by about 40%. Nevertheless, mammography has a number of drawbacks. Among these limitations, underdiagnosis of prognostic breast cancer and false positive (cancer not present) are major issues. Furthermore, routine mammography screening can help detect breast cancer early. However, frequent problems including false negatives, low screening rates, and needless biopsies are noted [8].

Motivation: DL is a viable solution to address these issues, resulting in increased screening accuracy, fewer false negatives, and fewer needless biopsies. Hidden features and relationships that are unseen to the naked eye can be found by DL algorithms [9]. The accessibility of such data is, however, challenging and occasionally unattainable in a number of medical imaging domains due to privacy concerns. Additionally, the medical industry naturally faces the imbalanced class issue. Using data augmentation to artificially expand the sample size is a popular strategy for

addressing over fitting [10]. Generative Adversarial Network (GAN)-based data augmentation has been used in numerous recent researches [11-15]. Prior research [16-18] has demonstrated that the use of GANs for data augmentation modifies the training distribution and improves classifier performance on datasets associated with breast cancer. GANs [11] have been utilized to create artificially generated computed images in health care area in recent years. GANs have also been used in non-image data domains, like transcriptome data for cancer staging, subtyping, and diagnosis. The synthetic production of highly realistic samples using GAN-based data augmentation has advanced quickly [12]. GANs have been used by researchers in medical imaging, comprising MRI and tomography [13, 14]. Techniques for image super-resolution (SR) are bringing emphasis to the restoration of poor-quality images that contain lost pixels. To ensure quantitative performance of the applied SR procedures to elongate the images in a manner that makes the texture qualities and feature details clear, researchers are increasingly turning to GANs because of their ability to realistically reconstruct the images [19]. The research [19], therefore, presented a SRGAN based SR solution approach for reconstructing histopathological images of breast cancer and employed a wide residual block. The research also used a self-attentional layer to capture various significant characteristics in both the discriminator and generator models.

Research Gaps: The SRGAN enhances the Low Resolution (LR) mammogram images by reconstructing High Resolution (HR) versions that expose fine diagnostic details such as micro calcifications, speculations, and lesion boundaries, which are often blurred in datasets like MIAS may be caused due to critical diagnostic structures like micro calcifications, masses, and margins are very small and easily get blurred or lost in LR images, but it still may lacks spatial distortions that may be caused due to different acquisition techniques, patient movement while screening etc.,. To address spatial distortion the research [20] proposed STN. This step increases the visibility of subtle abnormalities and provides richer, noise-reduced feature representations to the subsequent modules. Next, the STN performs automatic geometric correction by learning optimal transformations that adjust for rotation, scale, translation, and positioning variations commonly found in mammograms, which may be caused due to patient movement during scan, variations in breast compression. Previous research [20] used BI which is smooth but blur the fine details. This research experimented using LZ and found that LZ retain fine details but prone to artifacts.

Contribution: Therefore, in the proposed work, with aim to refine the image resolution SRGAN models are trained on two data sources and conducted experiments. Additionally, spatial transformation techniques were used for removing spatial distortions in the images. It was found that, in terms of performance, SRGAN with spatial transformation using learnable interpolation performs better than SRGAN without attention mechanisms in our investigation. At last, ResNet-50 is used for feature extraction and classification and achieved a satisfactory outcome for the classification using the MIAS and DDSM data sources.

Thus, the pipeline SRGAN-STN-ResNet-50 provides a robust mammogram classification. The notable development in this framework is combining both LZ and BI resampling techniques. From the previous research, it is known that BI, though it suppresses noise it cannot retain high frequency textures as it averages the values from the neighboring pixels, and even though LZ retains sharpness, the noise may confuse the model as the LZ interpolated values may vary significantly from nearby original values. . Therefore, in this research, we embedded BI and LZ interpolation which can reduce artifacts while preserving edges.

Considering the aforementioned issues of previous research, the present research proposed an enhanced hybrid approach of using SRGAN for improving image resolution, STN for performing spatial alignment, and ResNet-50 for classifying the images. The encouraging experimental results demonstrate how well the model works to improve image resolution and alleviate spatial challenges in mammogram images.

2. Related Work

This section examines the recent research on GAN, SRGAN, STN techniques for breast cancer classification. The following is an overview of few relevant studies:

A GAN-based augmentation approach was put forth in [21]. Their research used five classification algorithms and feature selection on medical datasets enhanced with the least-squares GAN (LS-GAN) technique. Through efficient data classification, empirical studies have shown that support vector machines (SVM) outperform alternative techniques. Their data augmentation method, which was based on LS-GAN, provided a practical way to enhance the functionality of medical classification models.

In order to train binary classifiers, well-known GAN techniques were proposed on synthesized data [22]. The accuracy of these classifiers was then compared to scenarios in which only real data were used. This study suggests that GAN-generated data are highly applicable due to two factors: data privacy laws and the unavailability of medical data. This study's primary goal was to ascertain how GAN-generated data can enhance classification accuracy, mainly for insufficient datasets. The paper [22] presents a strategy that takes into account a larger dataset that includes

both synthetic and actual data in order to achieve this goal. The results showed that, for both bigger and smaller data amounts, the data produced using sophisticated GAN techniques.

Deep learning has shown encouraging improvements in mammography cancer categorization, according to [23]. But the biggest obstacles to more progress are problems like unequal class distribution and data shortages. GANs were employed as a data augmentation technique for classification networks so as to solve this problem, and used a semi-supervised learning architecture based on U-Net. The authors created lesions on mammography regions that looked normal and then extracted the lesions from the patches where they were found.

According to research [24], taking into account augmented mammography patches improved the model's overall performance. The investigation outcomes revealed that the augmented GAN-based regime outperformed the baseline with an AUC of 0.846.

The problem of minimal labeled data in breast cancer mammography image classification was attempted to be resolved in another work [7]. They chose a small tagged dataset to achieve this. In order to improve the classification accuracy using a DCNN algorithm, the authors presented a technique that uses DC-GAN to synthesize images and supplement them with actual images [7]. They employed a dataset of 74 malignant and 213 normal pictures for this purpose. Batch sizes of 32 and 4 were used for the GAN-based training. According to their research, the accuracy of 25 artificially created photos in batch sizes of 4 and 32 was compared; the accuracy of the 32-image batch was 87% higher. Three experiments were conducted so as to validate the procedure. These findings included visual testing, a similarity study, and CNN-based performance assessments with and without DC-GAN.

Wu *et al.* [12] contended in their study that deep learning frameworks have demonstrated outstanding outcomes for the identification of BC in mammograms. However, these models' effective outcomes have been hampered by limited data challenges. GANs have been utilized to artificially enhance mammography datasets in order to overcome this difficulty [12]. For this, they trained a class conditional GAN to do contextual infilling. Lesions from healthy screening mammography were then synthesized by the authors. The scientists showed that GANs are capable of producing artificial mammography patches of superior quality. Additionally, their empirical investigation assessed an enhanced data source to enhance the classification performance of BC.

Therefore, Ledg *et al*. [19] proposed the SRGAN model to create the missing points. Realistic perceptual loss functions, which produce more quantifiable results than per-pixel loss in generated images, can be distinguished using perceptually based algorithms [25]. In order to customize SR difficulties, Ledig *et al*. have defined a critical perceptual similarity.

In order to obtain perceptual loss, this model makes use of the cutting-edge VGG-19 model [26]. Furthermore, the model's depiction of the adversarial loss is predicated on the chances of all training samples for the discriminator. Additionally, before sending the most significant features to the up-sampling layers, this model can extract them by using residual blocks [26] in the generator. Table 1. Below summarizes GANs related research.

3. Methodology

The complete pipeline of SRGAN-STN-ResNet-50 is end-to-end trainable and designed to improve classification accuracy by enhancing the input images and allowing geometric transformations to be learned dynamically. This research has assembled LZ and BI to preserve texture and edge details to maintain smooth transitions and suppress artifacts, enabling improved classification accuracy on BC image datasets. The SRGAN begins by artificially creating LR images from the original HR ground truth images. This down sampling is produced by making use of bicubic interpolation, which make sure that the LR images are a degraded yet realistic version of the HR images. These LR images are then delivered as input to the SRGAN's generator network, whose fundamental target is to rebuild SR images that sharply take after the original HR images in both content and perceptual quality.

Table 1. Study of BC classification based on GANs

[7]	DCGAN	DDSM	Used GAN for augmenting the dataset. Model efficiency can be increased by enhancing image resolution.
[27]	LS-GAN	UCI	Used SVM for classification. Pretrained models are needed for deep feature extraction.
[28]	GAN	Optimam-Mammography Image Database	Used U-net for high resolution, and fine details may be lost while up-sampling

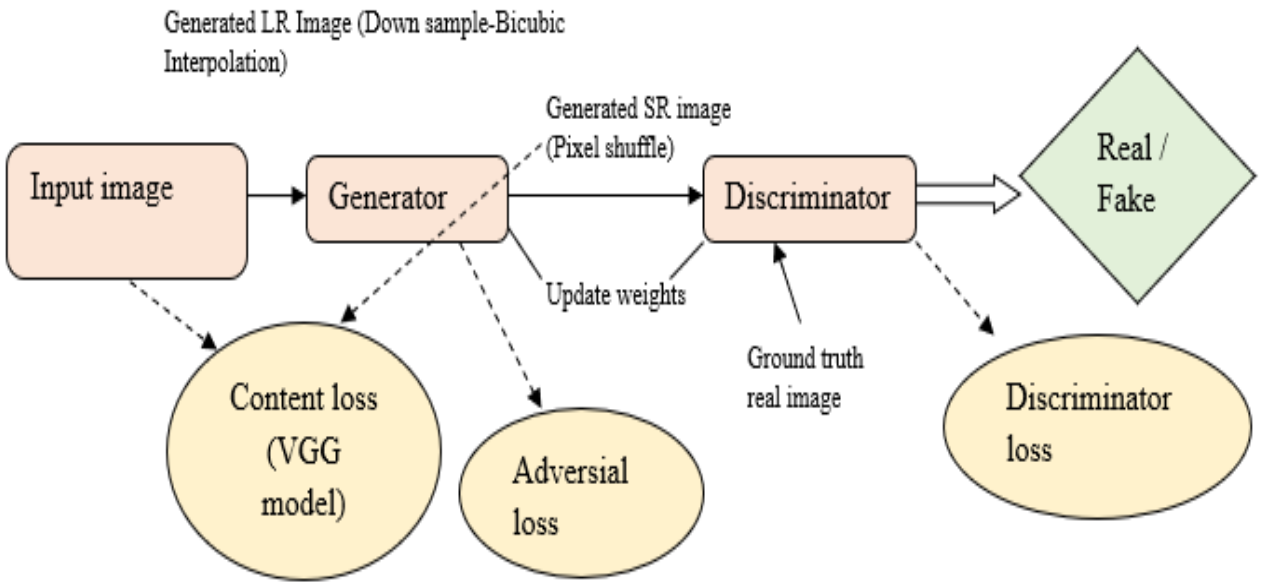

Figure 1. Block Diagram of SRGAN Training Procedure with Generator and Discriminator

To govern the generator towards creating additional realistic and visually convincing results, a discriminator network is used. The discriminator behaves as a quality evaluator by differentiating between real HR images and the generated SR images. Its contribution lies in producing adversarial feedback that encourages the generator to recapture fine textures, high-frequency details, and natural image structures that are generally missed in traditional upscaling methods. The sequential training procedure of the SRGAN architecture is depicted in Figure 1, which also shows how the discriminator, generator, and perceptual loss functions interact to provide realistic super-resolution. The discriminator attempts to maximize its capacity to discriminate actual and fraudulent mammogram images, whereas the generator tries reducing perceived disparity among generated SR and original HR images. This process of SRGAN outperforms traditional interpolation-based methods in terms of improving the visual fidelity and realism of the reconstructed SR images.

In SRGAN, the design of suitable loss functions is significant to achieve an equity between perceptual quality and natural image realism, in place of simply focusing on pixel-level similarity. Pixel-wise losses frequently resulted from traditional super-resolution techniques. To address this, the SRGAN uses loss functions that match the reconstruction procedure to both the statistical distribution of natural images and human visual perception. The deep feature representations taken from a pre-trained VGG network retain significant structures and semantic features. According to the discriminator, the generator is guided by the adversarial loss to generate realistic high-frequency features and textures that are visually identical to the original HR images. In addition, the discriminator is trained utilizing its own loss function, which motivates it to accurately discern between the SR images produced by the generator and actual HR images. The SRGAN uses an adversarial learning framework to formalize this training, optimizing the discriminator D and generator G with conflicting goals. Let $Y_{HR\in}R^{3\times h\times w}$ denote the HR image and $Y_{LR\in}R^{3\times h\times w}$ the down sampled version. The generator G learns a mapping: $\hat{Y}_{SR=}G(Y_{LR})$. Where $\hat{Y}_{SR}$ denotes an SR image generated from the LR BC images using the generator network G, based on the SRGAN framework. The loss functions are designed to effectively balance perceptual quality and visual realism in the SRGAN framework.

The discriminator $Disc$ is trained as a binary classifier to differentiate between real HR images $Y_{HR\in}R^{3\times h\times w}$ and the SR images of BC generated by the generator$\hat{Y}_{SR}$. Its loss function is given by the following equation (1):

$$Disc_{loss} = E[\log Disc(Y_{HR})] + E[\log(1 - Disc(G(Y_{SR})))] \quad (1)$$

For real HR images, $\log Disc(Y_{HR})$ should reach to 1 (real) and $\log(1 - Disc(G(Y_{SR}))$, should reach to 0 (fake) for the discriminator to be rewarded. Through the optimization of this loss, the discriminator enhances its capacity to accurately categorize both created and real images, giving the generator insightful adversarial feedback. Discriminator is therefore solely utilized to enhance the generator. This enhances the generator's realism in creating SR images. The discriminator's adversarial feedback forces the generator to create more realistic textures and features that look more natural, hence indirectly improving the generator. The SRGAN produces feature maps of mammogram images that are perceptually compelling, rich in fine details, and

free from the blurring effects typical of traditional approaches by cooperatively optimizing these complimentary goals.

The generator uses the adversarial loss to create visually realistic high-frequency textures of BC images and details for the discriminator. It gauges how well the generator deceives the discriminator, or how authentic the fake SR images appear. It motivates the generator to produce SR images that are identical to actual HR images shown in equation (2).

$$Y_{adv} = -E\left[\log Disc\left(G\left(Y_{SR}\right)\right)\right] \quad (2)$$

Perceptual loss guarantees that the resulting image retains significant structures and semantic features, whereas adversarial loss encourages realism. In place of raw pixel comparison, a pre-trained VGG network's deep feature representations are employed. Comparing the feature representations of the created image and the actual HR image is the usual method for calculating content loss. This research used pre-trained networks (VGG) to extract these properties. It compares VGG features while maintaining the proximity of SR images to HR. Comparison aids in the preservation of fine textures and features of BC images, which are essential for the reconstruction of high-quality images for BC classification shown in equation (3).

$$Y_{VGG}^{SR} = \frac{1}{W_{m,n},H_{m,n}} \sum_{\mathrm{m}=1}^{W_{m,n}} \sum_{\mathrm{n}=1}^{H_{m,n}} \left(\phi_{\mathrm{m,n}}(Y_V^{SR}\right) \quad (3)$$

Lastly, the ultimate goal of the generator balances realistic texture generation and structural fidelity by combining perceptual and adversarial concepts. The definition of the generator loss function is shown in equation (4).

$$Y_{Genloss} = Y_{VGG}^{SR} + \beta Y_{adv} \quad (4)$$

where the weighting factor β regulates the adversarial term's contribution.

To increase the quality of LR mammographic images, SRGAN is employed. A CNN that learns to upsample LR images to HR versions is employed in the generator. The network of the SRGAN-Generator architecture of the proposed work is shown in Figure 2. It is made up of a Parametric ReLU (PReLU) activation to add non-linearity after a Conv2d layer to extract broad contextual characteristics from input LR image.

The Generator uses PReLU as it starts from random noise or degraded input, it needs flexible nonlinearities to capture fine details. PReLU is a learnable version of ReLU. It allows the model to decide how much negative information to keep.

The formula of PReLU is seen in equation (5)

$$f(x) = \begin{cases} x \; if \; x > 0 \\ ax \; otherwise \end{cases} \quad (5)$$

Here, a is a learned parameter. This is learned during backpropagation. So, the network learns a suitable value depending on the problem domain. PReLU introduces learnable parameters for negative slope, allowing the generator to adaptively learn nonlinear mappings needed for high-quality image reconstruction. In ReLU, the positive side is fine, but the negative side always outputs zero. This leads to the dying ReLU problem. Therefore, PReLU is used to retain and propagate those small but important negative signals during up sampling, avoiding dead neurons and recreating high frequency details. The network then moves on to a series of eight residual blocks, each have two 3x3 convolutional layers, batch normalization layers, and a PReLU activation between them to improve gradient flow and preserve image details. After that, the model does up sampling using Pixel-Shuffle blocks. Each block is made up of a Pixel Shuffle operation that reorganizes channels into higher spatial resolution, a Conv2D layer, and a PReLU implementation. The Pixel Shuffle blocks gradually upscale the image. Finally, a Conv2D layer produces the HR BC mammogram images. High-quality image super-resolution is achieved by this architecture's effective combination of sub-pixel convolution (Pixel Shuffle) and residual learning.

In a CNN, the discriminator distinguishes between the SR images produced by the generator and the real HR images. Next, a Leaky ReLU (LReLU) activation is applied, which extracts basic low-level features like edges and textures while maintaining gradient flow for negative values. It balances adversarial learning, which results in more appropriate discrimination between HR and real mammogram images. Discriminator uses LReLU to stay stable while classifying HR and generated SR. LReLU allows a small fixed negative slope (0.01) for negative input, ensuring non zero gradients even for negative features, which are essential for robust adversarial training. The formula of LReLU is shown in equation (6)

$$f(x) = \propto x \quad (6)$$

Where ∝ is a small fixed constant (0.01). The discriminator's job is to distinguish real from fake. LReLU helps to remain sensitive to a wide range of feature distributions (including negative ones) without learning too much (unlike PReLU). Since the discriminator does not need to reconstruct features, but only judge realism, a fixed LReLU slope is often faster and more stable than PReLU in this part.

The Conv2D refines these properties further with a second layer. In view to guarantee a uniform size for FC layers regardless of the input resolution, a second Conv2D layer captures more complex visual structures. Its output is subsequently reduced to a fixed spatial dimension using adaptive average pooling.

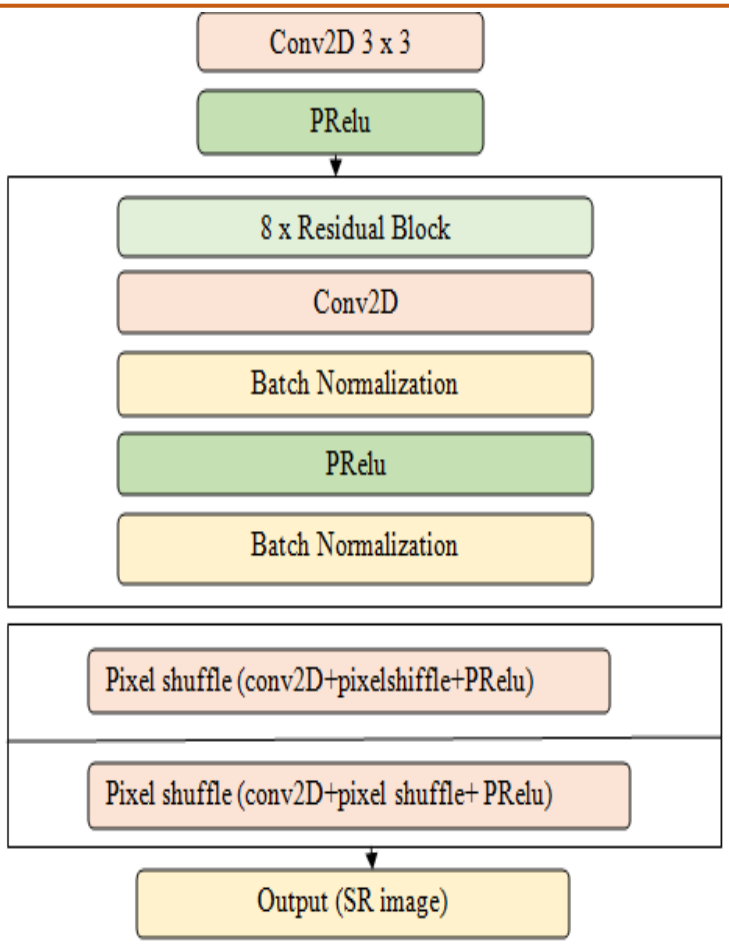

Figure 2. SRGAN Generator Architecture

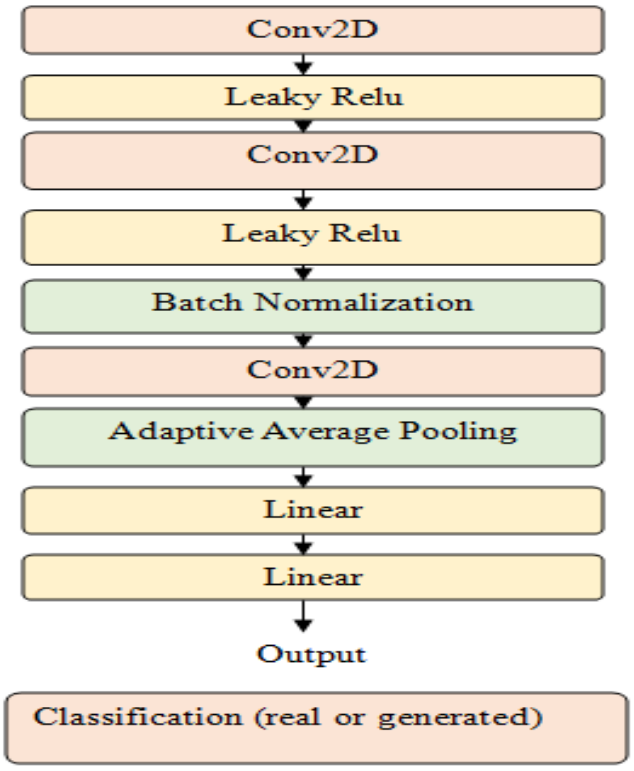

Figure 3. SRGAN Discriminator Architecture

In order to gradually reduce the dimensionality and produce a final scalar output, the pooled features are then flattened and processed through two linear layers. This output shows the likelihood that the input is authentic or fraudulent. Simply stated, the discriminator guides the generator to create more realistic SR images by learning to identify high-frequency details and textures. Figure 3 illustrates the proposed SRGAN discriminator architecture.

3.1 Spatial Transformer Network with Learnable Interpolation

To account for geometric misalignments and affine distortions in the images, this research incorporates an STN module. It comprises a localization network, a grid generator, and a sampler. The STN is integrated to learn geometric transformations that improve classification robustness. The STN architecture has two essential modules: the localization network, which predicts transformation parameters, and the grid sampler, which applies the predicted transformation to the input image. The localization network is a lightweight CNN that learns to predict an optimal 2×3 affine transformation matrix for each input image. It starts with a series of convolutional layers that extract low to mid-level spatial features, such as edges, corners, and textures. This affine matrix is fed into the grid sampler, which creates a normalized sampling grid that specifies how each output pixel should correspond to a certain point in the input image. BI and LZ interpolation are the two interpolation modes used by the grid sampler in our implementation. A learnable alpha parameter is introduced to allow more flexible resampling by adaptively blending the two modes. In particular, the output is transformed and calculated as follows: Given an input SR image $\hat{Y}_{SR}$, the STN learns a transformation matrix $\theta \in R^{2\times3}$ as shown in equation (7).

$$\theta = f_{loc}(\hat{Y}_{SR}) \quad (7)$$

Which is used to create a sampling grid G, and produce the transformed image $\hat{Y}_{STN}$ using grid sampling as shown in equation (1.8).

$$\hat{Y}_{STN} = GridSample(\hat{Y}_{SR}, G, BI) + (1 - W_{lz}).GridSample(\hat{Y}_{SR}, G, LZ) \quad (8)$$

Here $W_{lz} \in [0,1]$ is a learnable blending parameter initialized to 0.5, allowing adaptive weighting between BI and LZ. Here, $W_{lz}$is a trainable parameter constrained between 0 and 1. If $W_{lz} \rightarrow 1$, the STN prioritizes BI for smoother but slightly blurrier warping. If $W_{lz} \rightarrow 0$, then STN supports LZ. This fused sampling approach ensures to retain high frequency details of the generated spatially transformed features. Following a pretrained model, ResNet-50 is used for deep feature extraction.

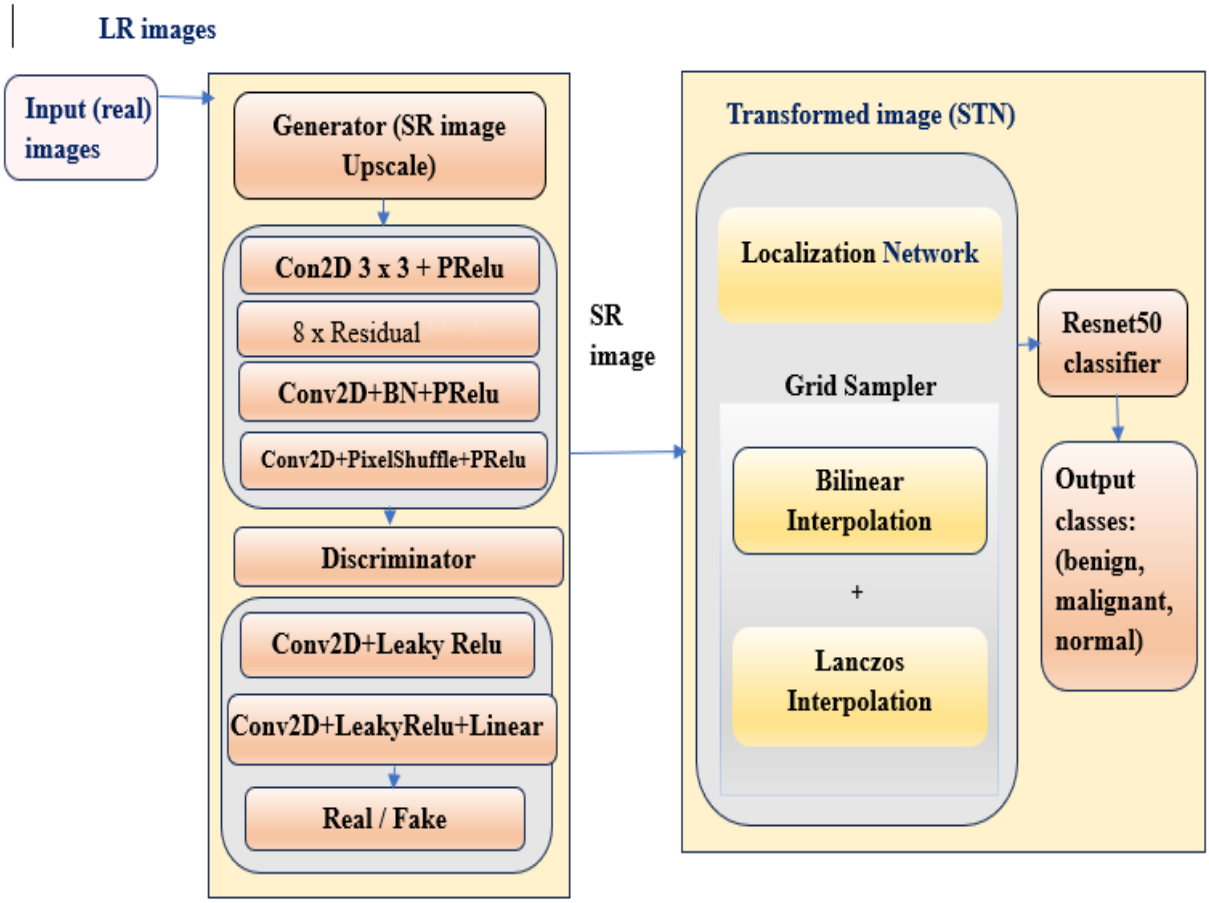

Figure 4. Overview of the Proposed Classification Architecture

The transformed mammogram image is forwarded to a ResNet-50 backbone, where initial layers are used for feature extraction and deeper layers (Layer3 and Layer4) remain trainable to adapt to domain-specific features. The resulting feature map is sent through a sequence of FC layers (2048→256→128→N) to perform classification across medical image classes. This architecture is specifically constructive for tasks such as BC image classification from datasets like MIAS, where resolution, spatial orientation, and fine details significantly influence diagnostic accuracy. Figure 4. illustrates the specific details of the proposed SRGAN-STN-ResNet-50 architecture for BC classification.

3.1.1 Dataset Description

The MIAS database of mammograms provided the digital mammography images used in this work for the empirical investigation of the suggested model. There are 322 digital mammography pictures in the "Portable Gray Map" (PGM) database maintained by the MIAS. Each image is 1024 pixels by 1024 pixels in size. The DDSM database contains a total of 2620 scanned images. The sample mammogram images are shown in Figure 5.

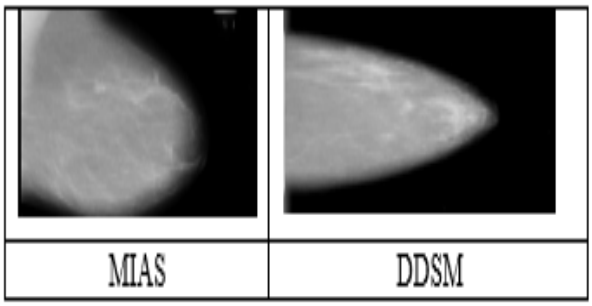

Figure 5. Sample Images from MIAS and DDSM data sources

4. Results and Discussion

The proposed method has collected Mammogram images from the Mammographic Image Analysis Society (MIAS) and Digital Database for Screening Mammography (DDSM) databases collected from kaggle to classify images as benign, malignant, or normal.

To examine the efficiency of the SRGAN-STN-ResNet-50, this approach is compared with other pretrained models and GANs that have already been established. The hybrid approach SRGAN-STN-ResNet-50 has attained an accuracy of 97.08% after 100 epochs with a learning rate of 0.3 and an error rate of 0.036 on the DDSM dataset and an accuracy of 96.58% after 100 epochs with a learning rate of 0.4and an error rate of 0.033. The datasets in this experiment have been considered 80% for training and the remaining 20% for testing. The problem of over-fitting is resolved with cross-validation of 10 folds. A few examinations on MIAS and DDSM mammogram images are produced to assess the efficacy of the suggested framework, and the results are displayed in Table 7. The confusion matrix of the proposed classification is shown in Tables 2 and 3.

We have conducted a few examinations on both datasets to assess the efficacy of our suggested framework. These results include a comparison of the proposed system and the performance of the existing works, are illustrated in Figures 8 and 9. Therefore, based on performance criteria, we can conclude that this method can be employed for breast cancer classification. These evaluation parameters can be expressed mathematically as follows:

Accuracy: The accuracy of the model is assessed by calculating the fraction of correctly classified instances to all instances. We used the equation that follows to do it:

$$\mathrm{Acc} = \frac{TP+TN}{TP+FP+TN+FN} \quad (9)$$

Precision: This metric computes the fraction of correctly identified positive results to the overall number of positive outcomes, including both true and false positives, so as to assess the accuracy of positive predictions. The following is the formula.

$$Prec = \frac{TP}{TP+FP} \quad (10)$$

Recall: It assesses the model's potential to precisely detect positive examples. The computation involves dividing the total number of true and false negative results by the number of true positives.

$$\mathrm{rec} = \frac{TP}{TP+FN} \quad (11)$$

F1-score: The F1-score gives a fair estimation of the model's performance by integrating precision and recall into a single metric.

$$\mathrm{F1^{-}Score} = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (12)$$

The model's performance is determined using the true positive (TP), true negative (TN), false positive (FP), and false negative (FN).

The research also evaluated F1-score, recall, and precision in addition to accuracy on the proposed model, and is listed in Table 4. The comparison between SRGAN-STN-ResNet50 and STN without pretrained models, ResNet-50 without transfer learning, is illustrated in Tables 5 and 6, and the graphs are illustrated in Figures 6 and 7.

In this section, we illustrate a comparison of our proposed model with existing works. The performance of the proposed work achieved a high accuracy of 97.08%, sensitivity of 95.91%, precision of 96.17% and F1-score of 95.48% on the DDSM data set and accuracy of 96.58%, sensitivity of 95.91%, precision of 96.17% and F1-score of 95.48% on the MIAS data set.

Table 2. Confusion Matrix of Proposed System with SRGAN-STN-ResNet50 on the MIAS dataset

Actual \ Predicted	Normal	Benign	Malignant
normal	325	10	12
benign	12	410	12
malignant	7	4	408

Table 3. Confusion Matrix of Proposed System with SRGAN-STN-ResNet50 on the DDSM dataset.

Actual \ Predicted	Normal	Benign	Malignant
normal	330	15	10
benign	10	410	10
malignant	5	5	410

Table 4. The performance metrics on SRGAN-STN-ResNet-50

Dataset	Accuracy	Precision	Recall	F1-score
MIAS	96.58	94.62	94.59	94.59
DDSM	97.08	95.91	95.08	95.48

Table 5. Performance Analysis of Models on the MIAS dataset

Models	Accuracy (%)	Sensitivity (%)	Precision (%)	F-measure (%)
STN	86	84	83	83
ResNet-50	94	92.83	91	94.62
STN-ResNet50	94.87	92	91.45	91.5
SRGAN-STN-ResNet50	96.58	94.62	94.59	94.59

Table 6. Performance Analysis of Models on the DDSM dataset

Models	Accuracy (%)	Sensitivity (%)	Precision (%)	F-measure (%)
ResNet50	88	86	87	86
STN	94	91.8	91	91
STN-ResNet50	94.27	93	94	93
SRGAN-STN-ResNet50	97.08	95.91	95.08	95.48

Table 7. The Performance Comparisons of Proposed and Existing Systems on MIAS and DDSM Datasets

S No.	Existing and Proposed Method	Accuracy on	
		MIAS	DDSM
1	Shrinivas D Desai *et al.* [7]	84	87
2	X. Yu [29]	95	94
3	Runyu Song *et al.* [31]	84	84
4	Yu *et al.* [27]	89	91
	Proposed (SRGAN-STN-ResNet-50)	96.58	97.08

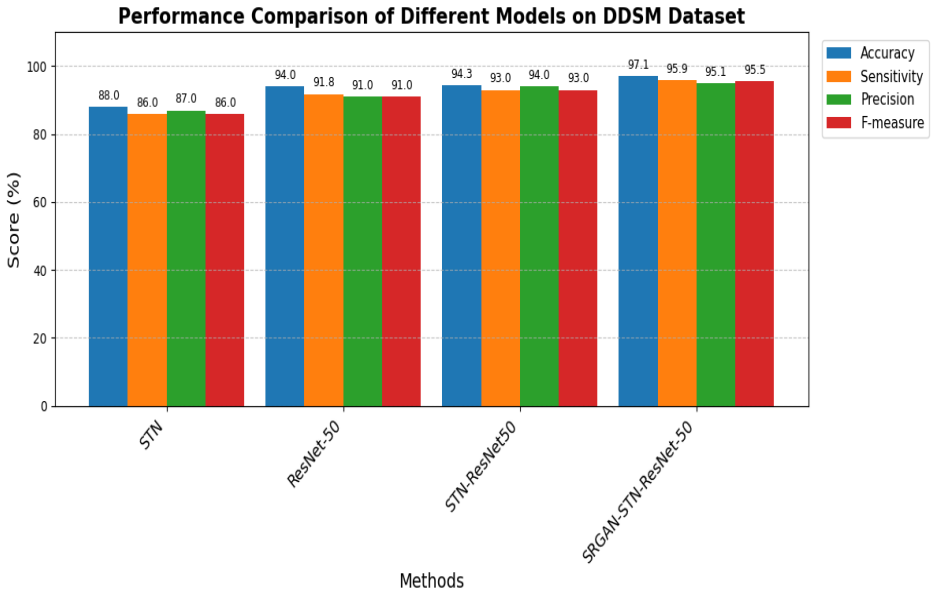

Figure 6. Performance Comparison of models on the MIAS dataset

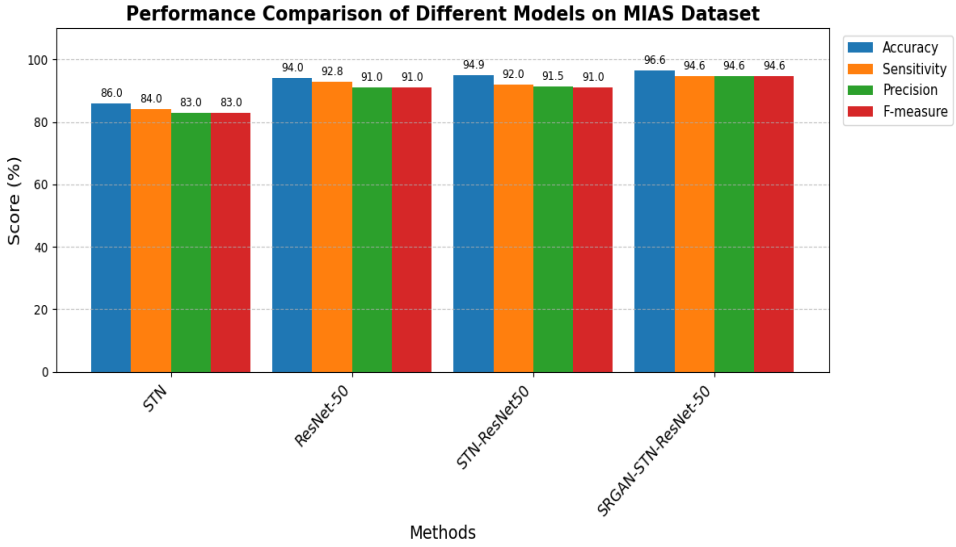

Figure 7. Performance Comparison of models on the DDSM dataset

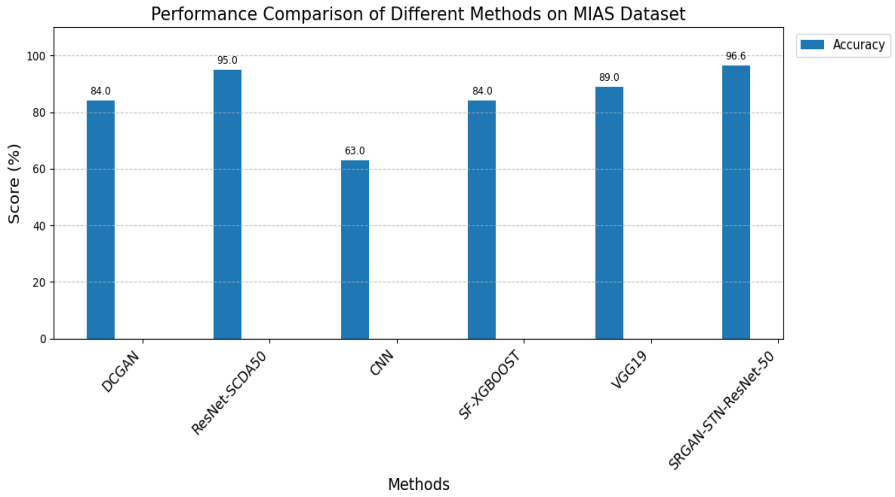

Figure 8. Comparative Performance Evaluation of Proposed and Existing Approaches on the MIAS dataset

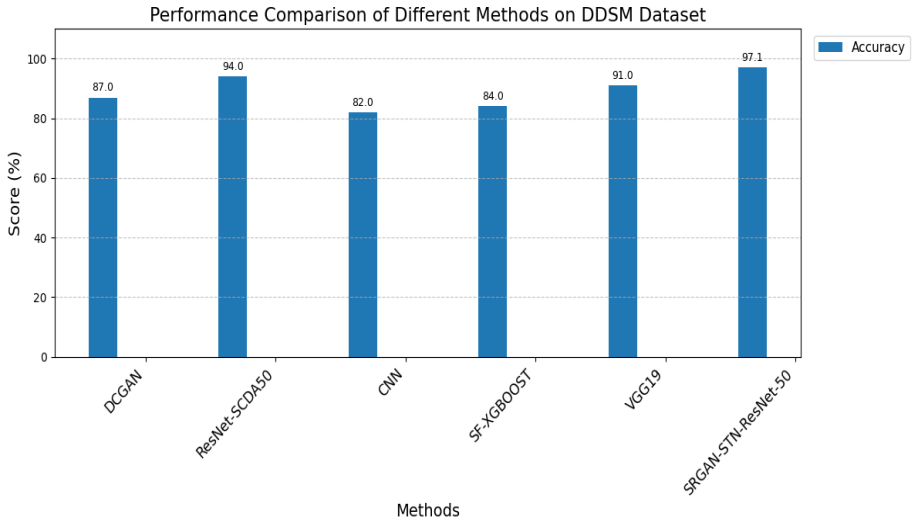

Figure 9. Comparative Performance Evaluation of Proposed and Existing Approaches on the DDSM dataset

Most of the existing works have provided accuracy, but have not provided all the performance metrics.

From Table 7, it is evident that the proposed method attained superior performance over the existing methods, and the reasons are stated below. Shrinivas D Desai *et al.* [7] classified the BC images by employing DCGAN and the reasons for its low performance are due to i) the DCGANs are more suitable for the generation of realistic images, and they are not powerful in capturing the fine details of mammogram images; ii) The DCGAN often fails in preserving fine edges and, they smooth the texture. The Runyu Song *et al*. [31] initially extracted texture features, and these parameters are fed into XGBoost. However, this method could not establish good accuracy results on BC images due to i) dependence on manually designed texture features; ii) XGBoost works on manually generated features, not on raw images.

5. Conclusion

For a better BC classification, this research finally integrated the advantages of SRGAN, STN transformations, and ResNet-50 to extract complex features. Experimental results reveal that generated images with fine quality can remarkably improve the performance of BC mammogram image classification. In terms of contribution, this research also proposes an embedded interpolation for resampling, which is beneficial for preserving anatomically relevant details. The experimental investigations showed that, when compared to other works, the SRGAN-ST-ResNet-50 achieved better performance. Thus, the proposed three approaches of the present research derived a precise analysis and addressed various disadvantages of breast cancer classification methods, and thus derived a good classification result.

References

[1] Y.C. Teh, G.H. Tan, N.A. Taib, K. Rahmat, C.J. Westerhout, F. Fadzli, M.H. See, S. Jamaris, C.H. Yip, Opportunistic Mammography Screening Provides Effective Detection Rates in a Limited Resource Health- Care System. BMC Cancer, 15(1), (2015) 405. https://doi.org/10.1186/s12885-015-1419-2

[2] S. Shams, R. Platania, J. Zhang, J. Kim, K. Lee, S.J. Park, Deep Generative Breast Cancer Screening and Diagnosis. Medical Image Computing and Computer Assisted Intervention – MICCAI 2018, (2018) 859–867. https://doi.org/10.1007/978-3-030-00934-2_95

[3] S. Destounis, A. Santacroce, Age to Begin and Intervals for Breast Cancer Screening: Balancing Benefits and Harms. American Journal of Roentgenology, 210(2), (2018) 279–284. https://doi.org/10.2214/AJR.17.18730

[4] R. Rawal, Breast Cancer Prediction using Machine Learning. Journal of Emerging Technologies and Innovative Research (JETIR), 13(24), (2020) 7.

[5] M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, R. Monczak, 'Computer-Aided Diagnosis of Breast Cancer based on Fine Needle Biopsy Microscopic Images. Computers in Biology and Medicine, 43(10), (2013) 1563–1572. https://doi.org/10.1016/j.compbiomed.2013.08.003

[6] M. Broeders, S. Moss, L. Nyström, S. Njor, H. Jonsson, E. Paap, N. Massat, S. Duffy, E. Lynge, and E. Paci, The Impact of Mammographic Screen- ing on Breast Cancer Mortality in Europe: A Review of Observational Studies. Journal of Medical Screening, 19(1), (2012) 14–25. https://doi.org/10.1258/jms.2012.012078

[7] S. D. Desai, S. Giraddi, N. Verma, P. Gupta, S. Ramya, (2020) Breast cancer detection using

GAN for limited labeled dataset 12th International Conference on Computational Intelligence and Communication Networks (CICN), IEEE, India. https://doi.org/10.1109/CICN49253.2020.9242551

[8] S. Ara, A. Das, A. Dey, (2021) Malignant and Benign Breast Cancer Classification using Machine Learning Algorithms. International Conference on Artificial Intelligence (ICAI), IEEE, Pakistan. https://doi.org/10.1109/ICAI52203.2021.9445249

[9] Lotter, W., Sorensen, G., & Cox, D. (2017, September). A multi-scale CNN and Curriculum Learning Strategy for Mammogram Classification. In International Workshop on Deep Learning in Medical Image Analysis, Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-67558-9_20

[10] J. Xie, R. Liu, J. Luttrell, C. Zhang, Deep learning based Analysis of Histopathological Images of breast cancer. Frontiers in genetics, 10, (2019) 80. https://doi.org/10.3389/fgene.2019.00080

[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative Adversarial Networks. Communications of the ACM, 63(11), (2020) 139-144. https://doi.org/10.1145/3422622

[12] E. Wu, K. Wu, D. Cox, and W. Lotter, (2018) Conditional Infilling GANs for Data Augmentation in Mammogram Classification. In International Workshop on Reconstruction and Analysis of Moving Body Organs, Springer International Publishing, Cham.

[13] J. M. Wolterink, A. M. Dinkla, M. H. F. Savenije, P. R. Seevinck, C.A.T. van den Berg, I. Išgum, Deep MR to CT Synthesis using Unpaired Data In International workshop on Simulation and Synthesis in Medical Imaging, Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-68127-6_2

[14] D. Nie, R. Trullo, J. Lian, C. Petitjean, S. Ruan, Q. Wang, D. Shen, (2017) Medical Image Synthesis with Context-Aware Generative Adversarial Networks. In International conference on medical image computing and computer-assisted intervention, Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-66179-7_48

[15] J.T. Guibas, T.S. Virdi, P.S. Li, (2017) Synthetic Medical Images from Dual Generative Adversarial Networks. arXiv preprint arXiv:1709.01872. https://doi.org/10.48550/arXiv.1709.01872

[16] H. Salehinejad, S. Valaee, T. Dowdell, E. Colak, J. Barfett, (2018) Generalization of Deep Neural Networks for Chest Pathology Classification in X-Rays using Generative Adversarial Networks. In 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, Canada. https://doi.org/10.1109/ICASSP.2018.8461430

[17] L. Hou, A. Agarwal, D. Samaras, T.M. Kurc, R.R. Gupta, J.H. Saltz, (2017) Unsupervised Histopathology Image Synthesis. arXiv preprint arXiv:1712.05021.

[18] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, H. Greenspan, Synthetic Data Augmentation using GAN for Improved Liver Lesion Classification. IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, USA. https://doi.org/10.1109/ISBI.2018.8363576

[19] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, W. Shi, (2017). Photo-Realistic Single Image Super-Resolution using a Generative Adversarial Network. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4681-4690.

[20] A.H. Abdulaal, M. Valizadeh, B.M. Albaker, R.R. Yassin, M.C. Amirani, A.S. Shah, Enhancing Breast Cancer Classification Using a Modified GoogLeNet Architecture with Attention Mechanism. Al-Iraqia Journal for Scientific Engineering Research, 3(1), (2024) 47-63. https://doi.org/10.58564/IJSER.3.1.2024.145

[21] M. Alauthman, A. Al-Qerem, B. Sowan, A. Alsarhan, M. Eshtay, A. Aldweesh, N. Aslam, Enhancing Small Medical Dataset Classification performance using GAN. Informatics, 10(1), (2023) 28. https://doi.org/10.3390/informatics10010028

[22] M. Abedi, L. Hempel, S. Sadeghi, T. Kirsten, GAN-based Approaches for Generating Structured Data in the Medical Domain. Applied Sciences, 12(14), (2022) 7075. https://doi.org/10.3390/app12147075

[23] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, (2014) Generative Adversarial Nets. Advances in neural information processing systems, 27.

[24] E. Wu, K. Wu, W. Lotter, (2020) Synthesizing Lesions using Contextual GANs Improves Breast Cancer Classification on Mammograms. arXiv preprint arXiv:2006.00086. https://doi.org/10.48550/arXiv.2006.00086

[25] J. A. Ferwerda, (2003) Three Varieties of Realism in Computer Graphics Human Vision and Electronic Imaging, United States. https://doi.org/10.1117/12.473899

[26] K. He, X. Zhang, S. Ren, and J. Sun, Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, USA https://doi.org/10.1109/CVPR.2016.90

[27] X. Yu, W. Pang, Q. Xu, M. Liang, Mammographic Image Classification with Deep Fusion Learning. Scientific Reports, 10(1), (2020) 14361. https://doi.org/10.1038/s41598-020-71431-x

[28] M.D. Halling-Brown, L.M. Warren, D. Ward, E. Lewis, A. Mackenzie, M.G. Wallis, L.S. Wilkinson, R.M. Given-Wilson, R. McAvinchey, K.C. Young, OPTIMAM Mammography Image Database: A Large-Scale Resource of Mammography Images and Clinical Data, 3(1) (2021) e200103. https://doi.org/10.1148/ryai.2020200103

[29] X. Yu, C. Kang, D. S. Guttery, S. Kadry, Y. Chen, Y.D. Zhang, ResNet-SCDA-50 for Breast Abnormality Classification. IEEE/ACM Transactions on Computational Biology and Bioinformatics, IEEE, 18(1), (2021) 94 – 102. https://doi.org/10.1109/TCBB.2020.2986544

[30] L. Sun, J. Wang, Z. Hu, Y. Xu, Z. Cui, Multi-View Convolutional Neural Networks for Mammographic Image Classification. IEEE Access, 7, (2019) 126273-126282. https://doi.org/10.1109/ACCESS.2019.2939167

[31] R. Song, T. Li, Y. Wang, Mammographic Classification Based on XGBoost and DCNN with Multi Feature. IEEE Access, 8, (2020) 75011-75021. https://doi.org/10.1109/ACCESS.2020.2986546

Authors Contribution Statement

Talari Swapna: Conceptualization, Methodology, Data Curation, Software, Writing – Original Draft. V. Vijaya Kumar: Supervision, Formal Analysis, Validation – Review & Editing. Both the authors have read and approved the final version of the manuscript.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Has this article screened for similarity?

Yes

About the License