IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 18 March 2024, accepted 12 June 2024, date of publication 21 June 2024, date of current version 11 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3417176

RESEARCH ARTICLE

Alleviation of Health Data Poverty for Skin Lesions Using ACGAN: Systematic Review

ASWATHY RAVIKUMAR[1], HARINI SRIRAMAN[1], (Member, IEEE), CHANDAN CHADHA[1], AND VIJAY KUMAR CHATTU[2,3,4]

[1]School of Computer Science and Engineering, Vellore Institute of Technology, Chennai 600127, India
[2]Centre for Global Health Research, Saveetha Medical College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai 602105, India
[3]RESTORE Laboratory, Department of OS and OT, Temerty Faculty of Medicine, University of Toronto, Toronto, ON M5G 1V7, Canada
[4]Department of Community Medicine, Faculty of Medicine, Datta Meghe Institute of Medical Sciences (DMIMS), Wardha 442107, India

Corresponding author: Harini Sriraman (harini.s@vit.ac.in)

This work was supported by the Vellore Institute of Technology.

ABSTRACT Skin-based infections are one of the primary causes of the global disease burden. Digital Health Technologies powered by data science models have the potential to revolutionize global health care. Health data poverty refers to the failure of individual people, teams, or communities to profit in research or development owing to a deficiency of representative data. Generative Adversarial Network-based synthetic images can be viable solutions to health data poverty since timely detection and frequent monitoring are extremely critical for the survival of the patients. This study aims to investigate the possibility of obtaining photo - realistic dermatoscopic images of Skin Lesions via Generative Adversarial Networks (GAN), followed by distributing the images to augment the existing dataset to further enhance the performance of a Convolutional Neural Network for the task of classification. The medical and technological publications in six databases: PubMed, Web of Science, IEEE Xplore, Science Direct, Scopus, and Google Scholar were investigated. A Deep Learning pipeline has been created and a set of deep learning models such as VGG16 (Visual Geometry Group 16), DenseNet, Xception, and Inception-ResNet v2 have been assembled. We have used condition-based generative adversarial networks (GANs) besides the traditional data augmentation approaches such as rotation and scaling. To highlight the image features that eventually lead to classification are highlighted using a Local Interpretable Model-Agnostic Explanation (LIME) strategy. It was inferred from the results of the classification that DenseNet-201 with GAN Augmentation was the best individual model, with an accuracy of around 82%, while models such as VGG-16 and SVM (Support Vector Machine) were unable to compete. It was also observed that starting with the pre-trained ImageNet weights sped up the convergence and prevented models from over fitting in the absence of the regularization effect of augmented data. However, the exploitation of the data was still not perfectly optimal, as over fitting with data augmentation and early stopping was observed, which can be used by more extensive data augmentation techniques. The GAN augmentation showed to reduce the data imbalance and increase the data percentage of the less representative classes. A data augmentation approach based on synthetic data that has been obtained from GAN helps us to classify images of lesions of the skin with high accuracy. We can also infer from the results obtained that, enriching the data with GAN-produced data samples results in a significant performance increase. In the field of medical imaging, where particularly large training datasets are not available, novel data augmentation and generation procedures can be beneficial.

INDEX TERMS Deep learning, health data poverty, machine learning, scalability, digital health, GAN.

I. INTRODUCTION

With the emergence of technology and continual advancement in the field of Deep Learning, it has been demonstrated

The associate editor coordinating the review of this manuscript and approving it for publication was Yu-Da Lin.

that digital health solutions have the potential to revolutionize health care [1]. If such tools and approaches could be implemented reliably and at scale, they would have the possibility of providing equal access to high-quality treatment - in Digital Health - for everybody, anywhere, and so reduce the global wellness and health disparity. However, it is quite probable that these technologies would hinder the existing equity in health care. With this perspective in mind, the challenge called ''Health Data Poverty'' for Skin Lesions has been discussed in this paper. Skin Diseases and particularly, Skin Cancer the most common type of cancer, as a result, it is often ignored by most people in its early stages. Of all the cancer-related cases that are reported, around 33% are related to Skin Cancer, in one form or the other. It was reported by dermatologists who conducted a study in 2013 that about 85 million Americans (27% of the population; more than 1 in 4 individuals) and more than 9,500 people in the U.S. are diagnosed with skin cancer every day [12]. The first step for the diagnosis of such a disease is usually through visual examination of the dermatoscopic images, which is quite inaccurate at times since the perception and the inferences drawn by the naked eye can be quite misleading. One such type of disease is Melanoma, which is life - threatening. The usual stages of diagnosis of Melanoma involve a visual inspection, followed by a biopsy. Hence, the precision of ocular inspection is essential due to the invasiveness of biopsy on patients. A study conducted in 2005 [31] proposed a popular rule for the screening of patients that were suspected of Melanoma, based on the geometric characterization of the Skin Lesions. The proposed rule was famously called the ''ABCDE'' rule, which stands for Asymmetric Shape, Border, Color, Diameter and Evolution. However, melanoma patients might not be aware of the severity of their condition without the aid of these clinical insights, missing the ideal opportunity for treatment, which is exactly what gives rise to the problem of ''Health Data Poverty''. Such cases are seldom seen in India, due to the presence of Eumelanin in India's people (who are relatively dark-skinned), which offers some protection against the development of skin cancer. Nevertheless, as per a study conducted in 2016 [21], 3.18% of all cancer patients in India had skin cancer. Of these, basal cell carcinomas made up 54.76 %, squamous cell carcinomas 36.91 %, and malignant melanoma only 8.3 %. Most patients (88%) were from rural areas, and most (92%) of them worked in agriculture. A significant amount of skin photos has been gathered recently due to the quick growth of computer hardware and software technologies, and sophisticated deep-learning-based models have been built to automatically analyze these skin images. These models' exceptional image analysis abilities make them highly promising as screening tools for skin conditions. Mao et al. for instance, suggested that the Google Inception v3 mode could be modified for melanoma diagnosis from skin images [27]. The model's classification performance was on par with that of human dermatologists after its parameters were tweaked using 130K skin picture samples. Alex et al. proposed a Fully Convolutional Neural Network (FCNN) model to extract multi-scale features from skin images and perform lesion classification and achieved an accuracy of 81.8% on a 10-class skin disease classification problem [28], where the best-reported accuracy for classification on the same dataset was about 67%. Similarly, a more contemporary approach was suggested by [33] who recommended using the deep Residual Network (ResNet), and they reached cutting-edge performance to accomplish all cases of Melanoma detection. Based on previous studies, prior efforts have also been made to deploy GAN to solve the problem of tackling skin lesions. A study conducted in 2018 made use of the deployment of GAN to create incredibly realistic representations of skin lesions [46]. These studies, however, have not shown how much (if any) performance advantage may be obtained by supplementing the training data with these artificial images. Instead, they have concentrated on the creation and synthesis of genuine images. Gains in performance on a task involving the classification of liver lesions have been made using GAN-generated data [48]. AlexNet was used in yet another previously published study that employed DCNN [53]. 200 images made up the data collection. However, 4400 images were created using image augmentation (rotating all the photographs). The AlexNet model was trained on ImageNet data for this study's transfer learning model, and the last layer—which is divided into melanoma, seborrheic keratosis, and nevus—was substituted with the softmax layer. They employed stochastic gradient descent (SGD) algorithmic software to alter the weights. They managed to obtain a 98% accuracy rate.

Despite these models' excellent quantitative performance, there are still several issues that need to be resolved, some of which are:

1. Robustness of the Training Model - Inception and ResNet, two well-known deep learning models trained on common computer vision tasks, are frequently adopted in existing research. It is challenging to ensure that one model will perform consistently well on various skin lesion images.

2. Limitation of the samples - the number of skin photos supplied in most of the research is insufficient to train the intricate deep learning model [58]. Due to the low number of skin photos, these models will likely need to be trained on other large-scale image data sets before being utilized to enable convergence.

3. Interpretation and Inference - in actual clinical practice, decision support would not be sufficient with just quantitative classification results. As mentioned before, dermatologists must use diagnostic standards to make the diagnosis.

Hence, the model that has been developed will address the following research questions related to GAN in digital health data poverty:

- Health Tasks - Diseases addressed in health data poverty using GAN
- Asserting the availability of sufficient digital health data for augmentation using GAN
- Obtaining effective real-time results.

II. MATERIALS AND METHODS

A. SEARCH STRATEGY

We carried out a thorough search utilizing a mix of topics and keywords to incorporate the topics of GAN and digital health. In addition, we ran referential retracing using the most current research to verify that all relevant papers were included. The medical and technological publications in six databases: PubMed, Web of Science, IEEE Xplore, Science Direct, Scopus, and Google Scholar were investigated. The study focused on recent publications from 2019 to 2022. During the filtering and review process, publications with flawed methods and results were removed.

B. SEARCH TERMS

The search was done using interventions and application keywords, including ''generative adversarial networks'', ''GANs'', ''digital data augmentation'' and ''digital data impoverishment.''

C. SEARCH ELIGIBILITY CRITERIA

This research focuses on the applicability of GANs in digital medical data augmentation. Only papers that presented GAN-based approaches for health data augmentation were considered. We did not evaluate studies that examined GAN-based approaches for other applications. Also omitted was any research that used deep learning, and machine learning techniques but did not include GANs. Also omitted were characteristics of students' GANs for non-imaging data. Purely peer-reviewed articles, conference proceedings and papers, and published works were included in this list of credible research. There were no limits on publishing country, research design, or results. Included are studies written in English and published between 2019 and 2022. In addition to removing duplicate data, unnecessary titles and contents were eliminated. Only those research that described direct and indirect GAN-based strategies for medical data enhancement were considered. The search and selection are clearly shown in Figure 1.

D. METHODS

The 2 baseline models that are used for the working of the pipeline are GAN and CNN.

Convolutional Neural Network (CNN) - CNN is a popular deep neural network whose architecture is inspired by the neural connectivity patterns in the animal visual cortex. It has been established through previous studies that CNN performs exceptionally well on computer vision tasks like image and video analysis. Convolution and pooling are two fundamental procedures used in CNN. A tiny filter that mimics the response of the receptive field (specified by the filter) centred on each pixel is applied to the entire image during convolution, and the resulting map is then passed through a nonlinear activation function to get the final response.

E. GAN

Following the breakthrough by Goodfellow et al. [1], GANs have demonstrated promising results for image creation in computer recognition in general. GANs produce realistic pictures while lacking a well-defined objective function and requiring rigorous training accompanied by oscillations and mode collapse, where the generator learns a minimal number of patterns. Two subnetworks comprise GAN: G-generator and D- discriminator. The generator network produces data with a particular distribution while training and the Discriminator network is used to detect if the data collected is genuine or fake. The ultimate objective of the procedure is to develop a G that can produce information with a distribution near the actual data distribution. When used for digital health, it can extend datasets with insufficient quantities of digital health data, allowing neural network approaches to be employed in conjunction with the big datasets.

The purpose of a Generative Adversarial Network in the context of the above-mentioned issue is to produce synthetic data while reducing patient identifiability, which may be characterized based on the likelihood of re-identification provided the combination of all supplied patient data. The framework is used to produce synthetic information that closely resembles the joint variation distribution in the selected dataset, giving a readily available, legally and morally sound alternative to facilitate more outstanding open data sharing and allowing the creation of AI-based solutions—comparing the ability to support to the synthetically created data to those fit to the actual data across datasets to assess the similarity in simulation results and to assist in identifying the original observations from the synthetic data. As a result, a data-driven GAN is anticipated to regularly outperform state-of-the-art approaches and exhibit dependability in joint distributions. Furthermore, this approach was created to produce datasets that may be made public while substantially reducing the danger of patient confidentiality violations.

In recent years, several GAN variations have been proposed. In this work, the GAN technique is used to generate more skin image samples, and those samples should not only be real but also be able to enhance the performance of lesion categorization. A conditional GAN was proposed by Sarra et al., which depends on the data classes for both data generation and discrimination probability [78]. By extending the conditional GAN [81], Frid-Adar et al. further proposed ''Auxiliary Classifier GANs'' (AC-GAN). In the AC-GAN framework, the discriminator evaluates the likelihood that the image is real or false along with the class probability, and the generator accepts a class label or condition along with the random noise vector. The GAN pipeline used in the proposed work is shown in Figure 2 and the basic GAN model is shown in Figure 3.

III. METHODOLOGY

The proposed workflow of the pipeline has 4 sections, which are as follows:

A. PRE-PROCESSING OF DATA

The dataset used (HAM10000, ISIC 2018 challenge) has around 10014 RGB images of skin lesions, with a pixel

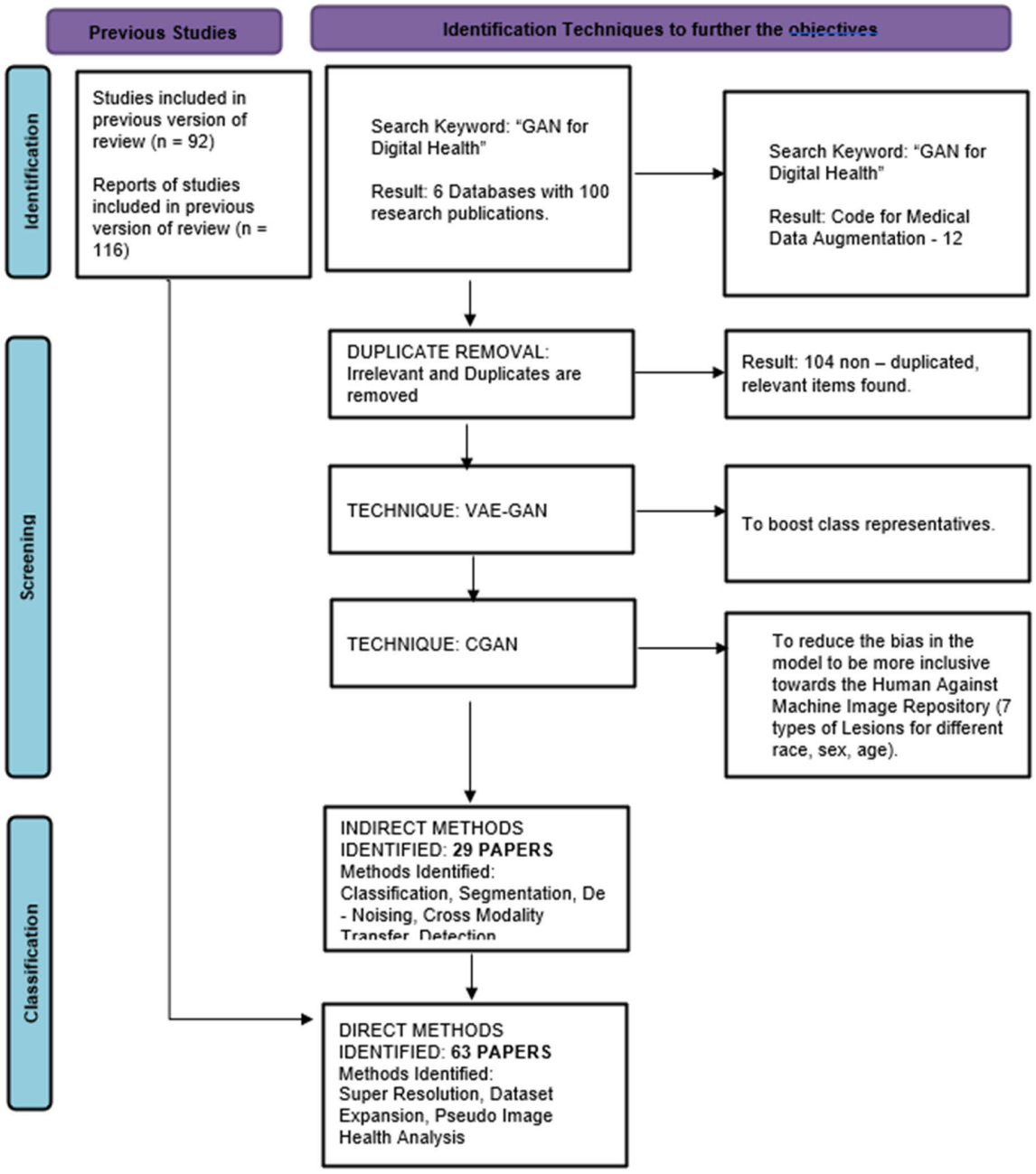

FIGURE 1. PRISMA: Preferred Reporting items for systematic reviews and meta-analyses for the review process.

resolution of 450 × 600 px. Since this dataset is heavily biased towards 3 of the classes namely, Melanoma, Nevus and Benign Keratosis, which constitute around 8900 of the total image samples, the model will initially exclude the remaining classes as shown in Table 1. This is done so that our GAN is fed with ample input data for it to learn the conditional distributions. 80% of the samples included were segregated under training data and the rest were treated as part of final testing/ validation data. Similarly, 20% of the training dataset was set aside when training the CNNs for internal validation. For greater attention on the regions of interest, the training dataset was preprocessed using the Gray World colour constancy technique, and the testing dataset was processed using a 90% square centre cut of the images. Then, all images were scaled to 224 × 224 pixels for training the CNNs and to 64 × 64 pixels for training the GAN (later sampled to match the CNN input size).

B. DATA AUGMENTATION

Each training sample had a set of random distortions in contrast and brightness applied to it, along with clipped zoom, rotations, and flips generating 4 samples per training sample.

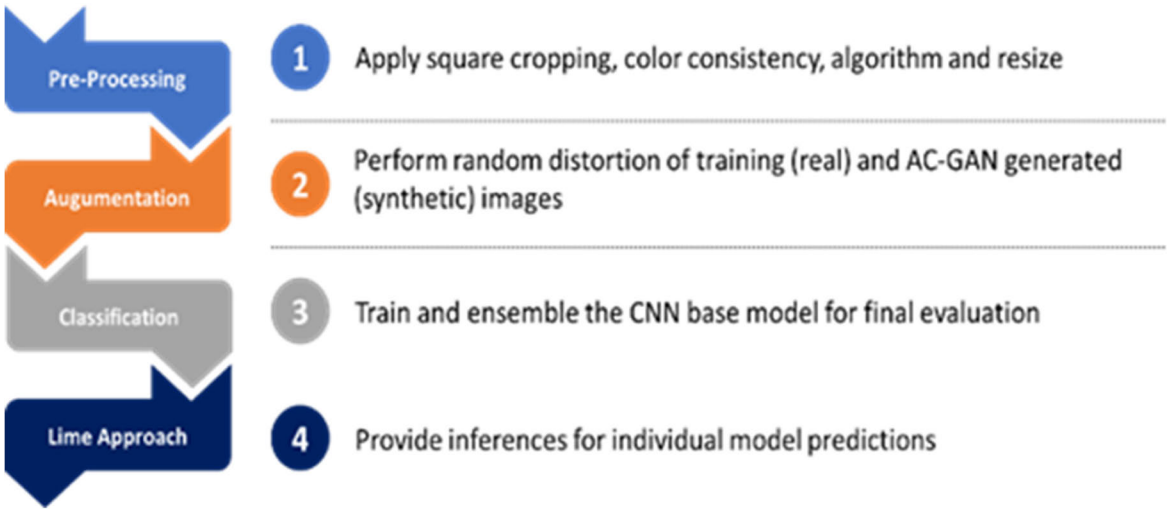

FIGURE 2. Proposed method.

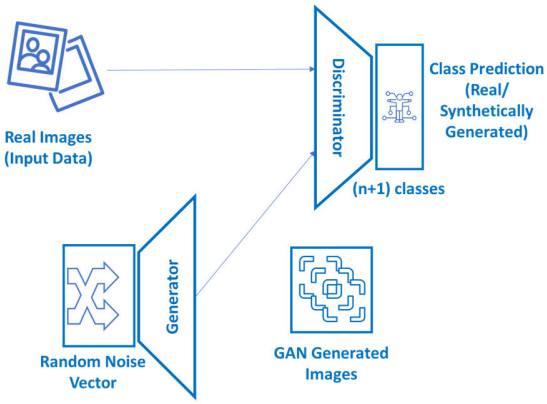

FIGURE 3. PRISMA: Basic model architecture for GAN-based classifier.

TABLE 1. HAM10000 dataset – lesion type vs. Image count.

Lesion Type	Image Count
Melanocytic Nevi	6705
Melanoma	1113
Benign Keratosis	1099
Basal Cell Carcinoma	514
Actinic Keratosis	325
Vascular Lesion	142
Dermatofibroma	115

Additional images were generated using the AC-GAN which follows the concept of a two-member mini max game. The generator function is used to accept the result of combining

TABLE 2. HAM10000 dataset – training and testing distribution for 3 most heavily biased classes.

	Melanocytic Nevus	Melanoma	Benign Keratosis
Training	890	5364	879
Testing	223	1341	220

a random noise vector with a class embedding, producing a picture with the dimensions 64 × 64 ×3 and pixel values within the range [−1, 1]. The discriminator accepted both real and false images and produced probabilities of the images' veracity using a sigmoid function as well as probabilities of the images' membership in a certain class using a soft max function. Both the generator and discriminator network made use of Leaky Rectified Linear Units (Leaky Re-LU) as the activation functions to batch normalization layers, which shall improve the training stability of the GANs [56], [57]. The training and testing data distribution is shown in Table 2.

C. TRAINING CLASSIFICATION MODEL

Keeping in mind the robustness of the model, an aggregation/ ensemble scheme of models has been adopted instead of just a single model. For base classification models, popular image classification architectures such as VGG-16, DenseNet (shown in Figure 4), Xception (shown in Figure 5), and Inception ResNet v2 are deployed [67], [68], [69]. While VGG only combines convolutional and pooling layers, it is incredibly inefficient and has a poor parameter-to-accuracy ratio. To increase performance, more modern architectures use microarchitectures like the Inception module, residual connections, and dense connections between layers, intending to enable deeper and more accurate networks. By stacking depth-wise convolutions with residual connections, Xception

TABLE 3. GAN implementation available online for digital health.

Methodology	Dataset	Code Source	Reference
GAN	brain MRI images	Git Hub	[1]
CGAN	COVID-19 CT	Kaggle	[2]
Cycle GAN	Multi-Contrast MR Images	Git Hub	[3]
GAN	brain tumour multi-contrast MRI	Git Hub	[4]
Hierarchical Amortized GAN	MRI	Git Hub	[5]
GAN	Liver Lesion	Git Hub	[6]
GAN	MRI	Git Hub	[7]
Patho-GAN:	Diabetic Retinopathy	Git Hub	[8]
GAN	mammography	Git Hub	[9]
GAN	XRay	Git Hub	[10]
Fila-sGAN	retinal fundus images	Private	[11]
GAN	melanoma	Kaggle	[12]

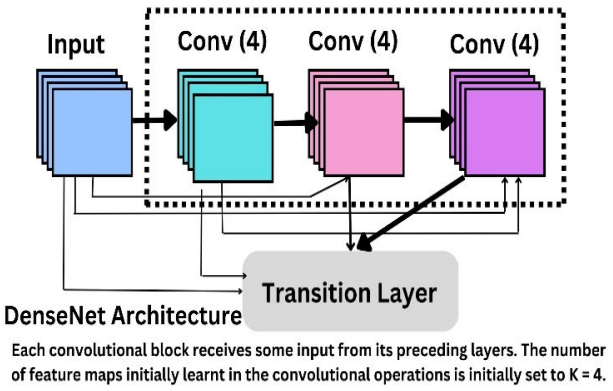

FIGURE 4. DenseNet model architecture.

alters the Inception v3 architecture and exhibits a more effective utilization of model parameters.

Inception ResNet v2 enhances Inception v3 with residual connections to quicken training and maybe increase accuracy. By implementing dense connections across layers, DenseNet promotes feature reuse and parameter sharing, resulting in improved performance with fewer parameters. The purpose of using an ensemble of designs is to minimize variance and increase accuracy because the methodologies used by various architectures allow for variation in the errors observed

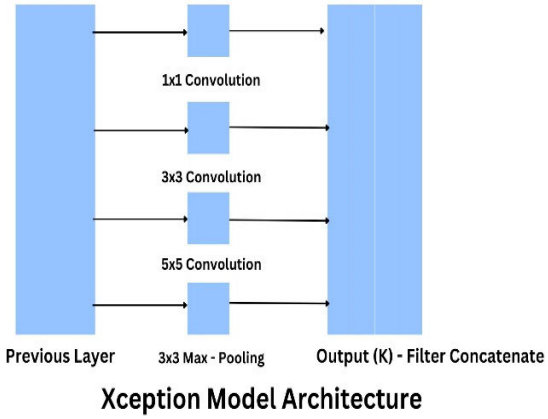

FIGURE 5. Xception model architecture (CNN).

during testing. The individual deep learning classifiers are trained using Categorical Cross Entropy (CCE) as the loss function. The size of different classes is taken while calculating the weight of the cross entropy of the input classes. The performance of each model has been calculated using Balanced Multiclass Accuracy (BACC). This is the sum of the probabilities of all classes in terms of their True Positive (TP) and False Positive (FP) predictions, divided by the total number of classes taken as input (n). Data Imputation for the ensemble strategies adopted makes use of flattening out the soft max layers of the CNNs to get the predicted class labels of the image, with a Support Vector Machine (SVM). Using the SVM ensemble approach, we evaluated the ensemble on the holdout test set after training the ensemble on concatenations of soft max layers from the validation data. The baseline comparison has been done by comparing the model to SIFT (Scale Invariant Feature Transform) with an SVM [59]. The RGB images received as input are changed to grayscale measure, followed by quantifying the image vector points as 128 length - vector clusters through K - Means clustering. The vector points are then taken as a reference to evaluate the RBF - kernel-based SVM [74].

D. INTERPRETATION OF MODEL USING LIME

Local Interpretable Model-Agnostic Explanations (LIME) tool that aims to provide classifier interpretations independent of the model. Being model agnostic means that LIME operates as a black box, treating a classifier's internal operations as irrelevant, and locally figuring out an input-to-output mapping. LIME additionally treats the CNNs as "black boxes," perturbing the picture it feeds to a particular CNN and estimating the CNN's decision function to comprehend the decisions made by our models. Using a sparse linear model centred on a single image, CNN's decision function is estimated. By mastering this decision function, LIME can identify image super pixels that, depending on their significance, influence the diagnoses made by CNNs.

The objective of the pipeline is to learn the generator network's distribution over the given data. The Generator

consists of a neural network that takes in the parameters Weight (W), input noise vector (z), and a prior distribution function (P(z)). As for the Discriminator network, it represents a CNN that indicates the probability of an input vector (x), being a part of the skin lesion classes, or from the generator distribution (P(g)). Hence, its output consists of a dimensional vector of "n + 1" classes while the initial probability distribution consists of only "n" classes. While the discriminator model is being trained, it attempts to maximize the chances of assigning the correct label to the input vector (x) by running multiple cycles in its effort to distinguish images that are generated synthetically using the generator sampling function (P(g)) from those that are part of the actual data samples. This happens parallelly along with the execution of the Generator function (G), which has been trained initially to deceive the Discriminator network by minimizing the value of [log (1 - D(G(z)))].

IV. RESULTS

A. MEASURES AND METRICS FOR DIGITAL HEALTH DATA AUGMENTATION

The quantitative assessment measures of synthetic images are based on the following criteria: synthetic image pixel level precision evaluation, generative dataset allocation overlaps assessment, subjective evaluation of radiologist evaluations, and informal evaluation of downstream task performance enhancement. The direct measures include pixel accuracy measures calculated based on the authentic images with mono modality to compare with the generated images. Evaluation of distribution overlap in the dataset generated to find the data's total fidelity. Evaluation of the ratings given by radiologists in which the generated images are quantified by the doctor based on the level of accuracy. The indirect measurement includes the task performance in which the usability of the data generated is increased using the model and improvement in its performance. The data augmentation of digital health develops data of different diversity, making the medical data reachable to all geographical locations.

B. GAN-DIRECT AND INDIRECT CONTRIBUTIONS TO DIGITAL HEALTH

Researchers find it harder to collect labelled medical pictures than unlabeled images for several reasons, including time, effort, and economic costs. However, there is currently a high demand for medical photographs among researchers. When dealing with such a vast quantity of data, the deep learning-based model may produce a higher performance in medical picture classification, segmentation, and augmentation than the hand-crafted components. However, the distribution of the pictures created by conventional augmentation techniques is comparable to that of the originals. Therefore, these techniques are unsuitable for creating more significant instances among patients. As a result, GAN is increasingly used in medical imaging analysis, particularly for data augmentation and multimodal image translations.

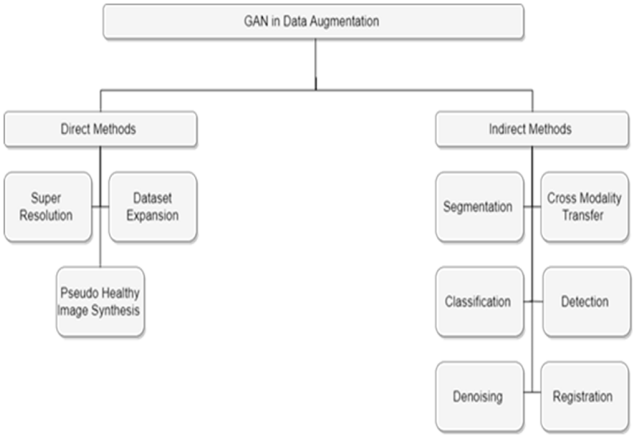

FIGURE 6. GAN in data augmentation.

Traditional data augmentation techniques can only provide data with a pattern that closely resembles the original data. GAN provides an answer to the shortage of information in medical imaging analysis. The advent of health digitization, hospital data management, IoT-based health platforms, wearable technology, and other systems have contributed to the exponential increase of patient records, including electronic healthcare records.

In Digital Health Domain, the most effective use of GAN is digital healthcare data generation and augmentation, which helps ease the issues of inadequate medical pictures or unbalanced data categories which come under the direct methods. GAN's Contribution to Digital Health Data Poverty is through Segmentation, Cross modality transfer, Classification, Detection, Registration, Denoising, and Reconstruction.

C. GAN TECHNIQUES FOR DIGITAL HEALTH DATA AUGMENTATION TO ALLEVIATE DIGITAL DATA POVERTY

The study focuses on how GAN aids digital health data poverty directly and indirectly and is shown in (Figure 6). GAN has been used to process digital health images for various tasks, including segmentation, modality transfer, disease classification, image reconstruction, and registration as shown in Table 4.

1) SEGMENTATION

GAN enhances the precision of digital health image segmentation due to its proven ability to create and capture data distribution. Implementing GAN and its extended variations may increase the segmentation results of digital health images [14], [15]. Gaining the acceptance of physicians and patients and eliminating the volatility, poor reproducibility, and un-interpretability of Generative Adversarial Networks might make this a significant future study path.

2) CROSS-MODALITY TRANSFER

The primary purpose of the Cross-Modal Generative Adversarial Network (CM - GAN) is to get a deeper understanding

TABLE 4. Data implementation available online for digital health.

Objects Analysed	Model	Task	Remarks	Reference
Healthcare Providers	GAN	Detection	Area Under Curve – 97.4%	[13]
Synthetic Datasets	GAN	Comparison	Accuracy - 79% (after 35000 epochs)	[14]
Multispectral Satellite Images	CNN	Prediction	the correlation coefficient of 0.66	[15]
Biomedicine	GAN	Segmentation	Applications of GAN in the medical domain and applications of GAN in medical informatics and bioinformatics	[16]
Human Brain	GAN	Un sampling	Accuracy - 83.88% (sample size = 20)	[17]
Health monitoring Sensors	Auto Encoders	Monitoring	Accuracy 99.86% (including latent losses)	[18]
Prostate Gland	GAN	Prediction	-	[19]
Human Brain	GAN	Detection	AUC- 93.72% (sample size = 6)	[20]
Private Hospital's Medical Data	GAN	Anomaly Detection	Max correlation = 0.19 (sample size 30389)	[21]
Wireless Devices	Selective GAN	Data Augmentation	-	[22]
High-Risk Pregnancies	DCGAN	Monitoring	Accuracy - 97%	[23]
Disease Image Bio - marking	GAN	Anomaly Detection	Sensitivity - 93.01%	[24]
Chest X-Ray Images	GAN	Generation	Accuracy 93.75% (+- 5.33%)	[25]
Chest X - Rays	GAN/ CNN	Segmentation	-	[26]
Pixel-wise analysis of the Human Body	Multi-Task cGAN	Prediction	Accuracy 96.46%	[27]
Super Resolution of Medical Images	Progressive GAN	Image Analysis	Accuracy 83%	[28]
Human Brain	Generative GAN	Lesion Detection	Sensitivity 93.1% (mean = 3.67, SD = 1.32)	[29]
Human Pancreas/ Prostate Gland	GAN	Blood Glucose Prediction	Accuracy 91.03%	[30]
Human Lungs	GAN	Segmentation	Intersection Over Union 94%	[31]

TABLE 4. *(Continued.)* **Data implementation available online for digital health.**

Air Quality	Conditional GAN	Imputation	Accuracy 67% (sample size = 3 polluting gases)	[32]
Glaucoma Clinical Trials	Cyclical GAN	Classification	Accuracy 96% (Mean Difference = 0.69 um)	[33]
Protein Function	Synthetic GAN	Prediction	Accuracy 99.99% (for Leave-One-Out Cross-Validation)	[34]
Human Brain – EEG Signals	GAN	Imputation	Accuracy 82.85% (SD = 5.89)	[35]
Accelerated Multi Contrast MRI Signals	GAN	Image Reconstruction	-	[36]

of the discriminative generic model for overcoming the heterogeneity barrier. It is offered to simulate the joint probability distribution across data of diverse modalities [16]; this is their primary contribution. The cross and intra-modality correlations in GAN models may be investigated concurrently, with both competing to enhance cross-modal correlation learning. In addition, cross-modal autoencoders featuring weight-sharing restrictions are presented as generative model components. This allows them to use cross-modal training for a generic model and to keep reconstruction data for capturing consistent experience in each modality [17].

3) CLASSIFICATION

To increase classification performance, GANs battle with generative and discriminative classifiers. The proposed system provides users of GANs with classification and detection challenges. Several GAN network architectures and training data set dimensions were tested with the discriminative network baseline and Bayes' classifiers [18]. As processor power, cache, and storage capacity increase, neural networks keep increasing the number of hidden layers and the number of elements in each layer to solve more problems [19]. Deep Learning's fundamental issues have not been solved by the most recent algorithms offered with a given amount of processing power. They could include performance versus a multitude of class labels, generalization – overfitting, and having adequate data to train the exponentially expanding number of distinct computing units.

4) DETECTION

In recent years, unsupervised methods have been adopted for detecting anomalies because GAN may draw aberrant conclusions via adversarial training of a sample representation. To give some motivation for the development of GAN-based anomaly detection, this study - analyses the idea of anomaly, provides some criteria for addressing the problem of anomaly detection and discusses the present issues associated with anomaly detection [20]. The paper concentrates on the conceptual and technical progress, theoretical foundation, practical tasks, and other practical relevance of GAN-based outlier detection for current works [21].

5) REGISTRATION

The traditional techniques that use any kind of Generative Adversarial Network attempt to solve the issue of image registration using methods that are primarily repetitive and time-consuming. Modern Deep Learning based registration approaches often extract deep characteristics for iterative usage. In [22], the research attempts to provide an end-to-end deep learning approach for recording patient data while ensuring the security and privacy of multimodal picture data. The method uses GAN to reduce the cost function to generate the correct dataset in less than one second. Various studies confirm its precision for multimodal pictures and data [23].

6) DE-NOISING

A compound Generative Adversarial Network (GAN) may also be used for de-noising. A GAN can be utilized to learn, generate noise, and generate paired-image data. This paired-image data can train a de-noising network, such as CNN [24], [25], [26]. GAN is capable enough to perform the above-mentioned process because a GAN model can be trained to learn sophisticated data in terms of real noise [27]. The trained and realistic noise model solves the problem of poor de - noising performance due to a lack of data [28]. This is where the proposed work aims to take the use of GAN in the sense of its de-noising capability to alleviate "Health Data

Poverty.'' Some of the significant works in recent times on indirect applications are listed in Table 4.

V. CASE STUDY

The ImageNet weight of all models was made available from Keras Applications repository [72]. The processing for the initial sample batch (size = 40) was done through Adam optimizer with the learning rate being close to (1/ 10000), after 20 cycles of training. This was stopped when the training accuracy did not show much improvement. Alongside, DenseNet-169 was trained with initial AC-GAN samples (shifted to model input size) along with data augmentation techniques. The balanced accuracies of each training model and their ensemble are mentioned below. SIFT with SVM, Xception, and DenseNet 121 are not a part of the ensemble models as they did not boost the accuracy of the final model considerably. While considering only 80% of the total data (input + generated) as the validating data (the rest 20% is assigned as independent training data), it can be seen that the pipeline can achieve a classification accuracy of more than 85% while considering the 3 most heavily biased classes i.e., Melanoma, Nevus, and Benign Keratosis. This is modelled in comparison to the traditional image classification models such as SVM trained on SIFT, which fetches an accuracy of a little over 53% [26], [27]. As it is evident from the 3 confusion matrices plotted for the data using DenseNet 169 (with 40% GAN augmentation), DenseNet 201 (with 50% GAN augmentation), and Ensemble SVM, the accuracy in predicting the correct class of disease across each row increases as we move from the DenseNet models towards the SVM ensemble approach. These confusion matrices show how individual model predictions can be balanced by an ensemble to increase classification accuracy. For instance, the accuracy of Melanoma classification in DenseNet 201 is relatively poor but is balanced out through DenseNet 169. Similarly, the accuracy of prediction of Benign Keratosis is relatively poor in DenseNet 169, which is balanced out through DenseNet 201 [59]. The confusion matrix for the Dense Net 169, DenseNet 201 and SVM ensemble model is shown in Figures 7, 8 and 9 respectively.

The Convoluted Neural Networks were trained using different classification models on GAN for different levels of accuracy of Data Augmentation. The generated images were iterated over DenseNet 169 and DenseNet 201 for no levels of GAN implementation, followed by 10%, 20%, 40% and 50% implementation in subsequent stages. It was observed and inferred through these cycles that the best accuracy for each model was achieved when DenseNet 169 was trained with 40% GAN augmentation (resulting in an accuracy of over 82%), while for DenseNet 201, the best accuracy (above 88%) was recorded with 50% GAN augmentation.

Tables 6 and 7 show that compared to DenseNet-169, a deeper and much more complex model like DenseNet-201 has a greater capacity for additional features offered by generated images.

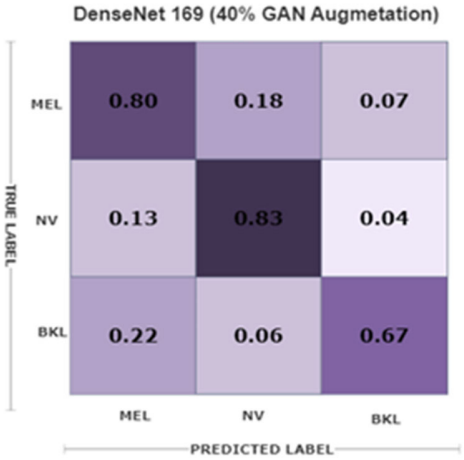

FIGURE 7. Confusion matrix using DensetNet 169 model (40% GAN augmentation).

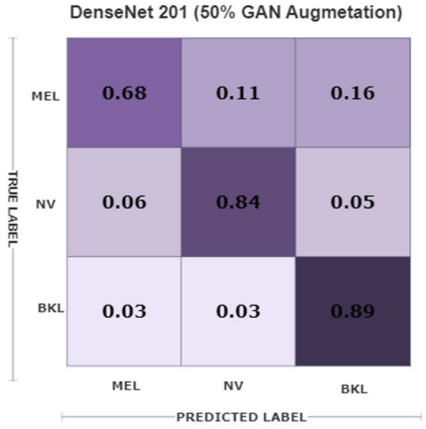

FIGURE 8. Confusion matrix using DensetNet 201 model (50% GAN augmentation).

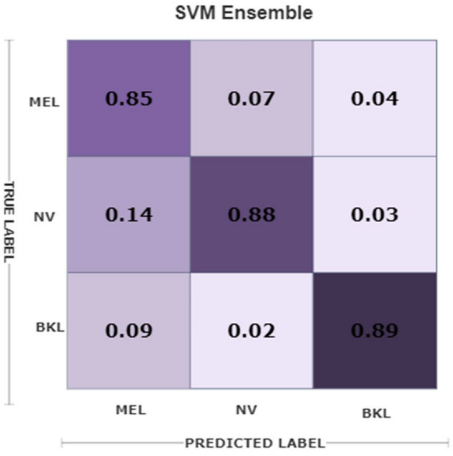

FIGURE 9. Confusion matrix using SVM ensemble model.

Since only 20% of the total data has been assigned as ''Independent Training Data'', we estimate the count of photo – realistic images for every class based on this criterion.

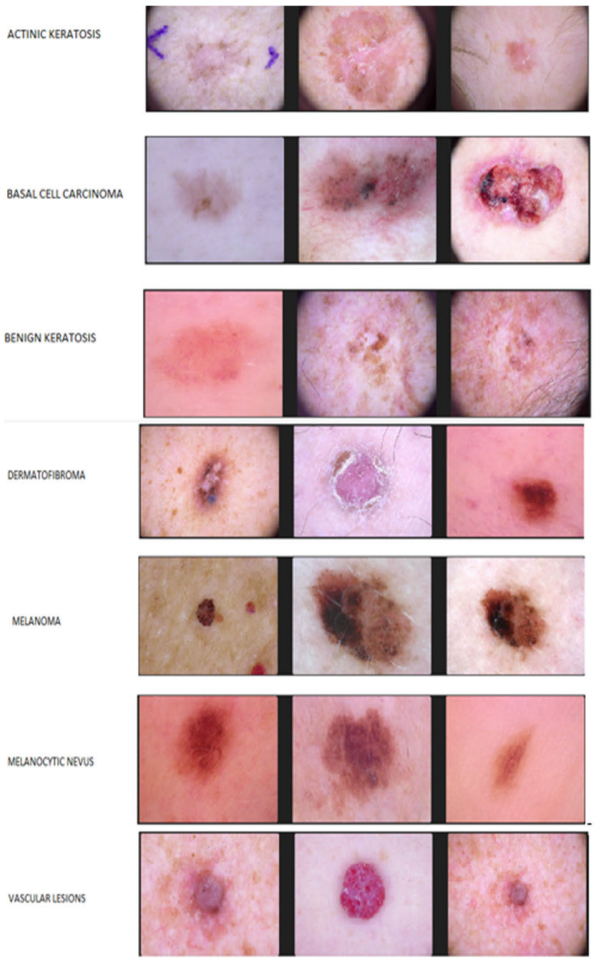

FIGURE 10. Sample images fed as Input (part of ISIC 2018 challenge – HAM10000 dataset)n data augmentation.

Hence, for each classification model, we can calculate the count of photo – realistic images (indistinguishable by the discriminator function) generated by the Auxiliary Classifier GAN (AC-GAN).

$$\begin{aligned} &\textit{Count of photo} - \textit{realistic image generated for a} \\ &\quad \textit{class}\,(\textit{Nearest Whole Number}) = (20/100) \\ &\quad * (\textit{class count in initial dataset}) \\ &\quad * (\textit{max. accuracy in Epoch}). \end{aligned} \tag{1}$$

The trend for both models shows that a performance enhancement at a moderate level (using GAN augmentation) is ideal. The Auxiliary Classifier Generative Adversarial Network (AC-GAN) plays its role to generate synthetic copies of the input sample data, keeping them juxtaposed with the original input data with the original class labels. The metric for placing the images side-by-side is through the calculation of the Euclidean Distance, between the 2 nearest neighbours (implemented through K - Means Clustering algorithm) in the training dataset. One striking inference made from the AC-GAN-generated synthetic images was that although the

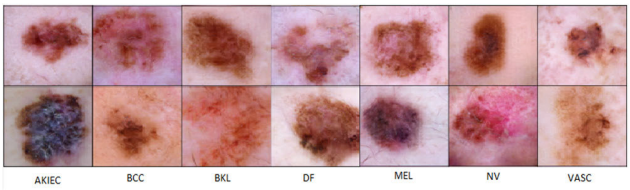

FIGURE 11. Sample auxiliary classifier generative.

images of Skin Lesions produced resemble those present in the initial training set (recorded using the parameters of size, form, and colour), the AC-GAN succeeded in creating its unique images of Skin Lesions. Among the seven classes four classes are underrepresented in terms of data and the inequality needs to be removed. The GAN data augmentation aids to reduce the class imbalance and the fewer representative classes are augmented and the data percentage is increasing as shown in Table 8.

These images were not generated by memorizing and recreating images of Skin Lesions from the training dataset. Furthermore, even though some samples within classes are remarkably similar, there is still some diversity in the skin lesions that were generated, both within each class and among the three classes, demonstrating the ability of AC-GAN to capture multiple modes of data distribution.

Adversarial Network (AC-GAN) generated synthetic images.To provide a demonstration of how a Convoluted Neural Network (CNN) works to make decisions based on classification, we have adopted a model-agnostic method, known as LIME (Local Interpretable Model-Agnostic Explanations) [42]. Here, we make use of the DenseNet 201 model to provide Positive diagnoses of the region of Skin esion that contains the top group of pixels for each prediction. It has its primary application in highlighting the meaningful regions of the input data images containing Skin Lesions. A True Positive diagnosis for the 3 most heavily biased classes of Skin Lesions (Melanoma, Nevus, and Benign Keratosis) was done from the test set, with DenseNet 201 as the classification model.

1) All the data collected for CNN and AC-GAN was processed and the results were analyzed using an NVIDIA GTX - 1050 GPU. To further boost the processing and to look for better results, we made use of the Google COLAB (P100) GPU. Figure 12 shows the percentage improvement in the less representative classes using GAN and overcoming class imbalance.

The method of early stopping of execution cycles (stopping after 5 Epochs), will increase the amount of training data that is to be fed into the model (around 80% of the total data is Training Data, and 20% is Testing Data), and performed Data Augmentation in the initial stages of the pipeline execution. From Table 10 it was found that the average F1 – Score: is 0.83, and the Lowest F1 – Score is obtained for Actinic Keratosis (AKIEC) which is equal to 0.71. The Highest F1 – Score is obtained for Melanocytic Nevus (NV) which

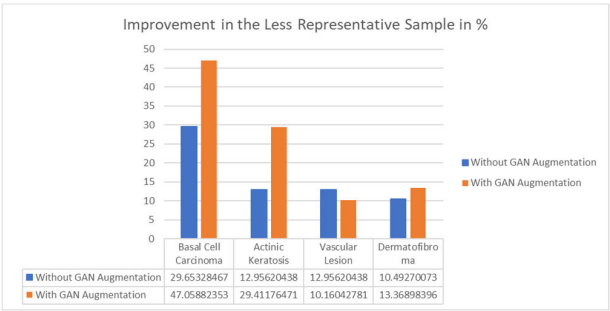

FIGURE 12. Sample improvement in less representative classes using GAN.

is equal to 0.94. This is because of the large number of Training Samples available for Melanocytic Nevus. The count of initial samples classified as Melanocytic Nevus, along with its discussion as one of the three most heavily biased classes has been mentioned in Table 2. The accuracy of Vascular Lesions remains constant after multiple cycles since it is one of the least represented/ biased classes. The count of photo-realistic images generated by the ACGAN for Vascular Lesions remains fairly constant, even after 5 epochs. Hence, to cater to the need for the overfitting of data, we have resorted to the early stopping of the cycles so that the problem of overfitting data is solved. Also, Data Augmentation has been performed on the data that involves generating new training data from the existing data by applying transformations like rotation, scaling, or flipping. This can increase the size of the training data and help prevent overfitting.

VI. CHALLENGES

A. CHALLENGES IN DIGITAL HEALTH DATA

1) DATA PRIVACY/SECURITY

The patient's permission is required to collect medical photographs for scientific study. It is unclear whether produced pictures or datasets derived from them should be deemed original or fresh data and consequently receive consent from the patient. In addition, the legal status of further data is undetermined. GAN uses domain conversion, which may endanger patients' privacy far more than the original photographs. To apply new technology, it is vital to assess not only its practicality but also its morality and legality.

2) DATA SHORTAGE

Although several datasets are available to the public, the majority were not created to be used with GAN but for other healthcare applications. Existing medical databases are of varying quality, with some being obsolete and fragmented. For specialized tasks, like the transition from MRI and CT scans, distinguishing images of a specific scale is not practical. The bulk of investigators receives them from hospitals.

3) IMAGE CONFIDENCE

The perception of a medical picture may influence the patient's life in digital health; hence, many strategies that are useful in other disciplines for equivalent applications may not be applicable in this medical sector. On occasion, clinicians may not have good trust in a standard medical image, necessitating multi-level identification. As a result, the generalization performance of an adequately trained model is substandard, or changes across certain data domains cannot be consistently executed. This issue merits consideration. However, it does not indicate that all GANs will be misdiagnosed.

4) INTERRELATIONSHIP BETWEEN HEALTH, SOCIO-ECONOMIC STATUS AND RACE/ETHNICITY

Diseases and illnesses, functional decline, disability, and mortality are just a few of the characteristics of health. Even though they are typically not clinically diagnosed until at least middle age, chronic diseases and disorders are influenced by lifelong factors that are associated with both socioeconomic position and race/ethnicity. Socioeconomic and racial/ethnic disparities are more pronounced in some aspects of health and from some sources than others. Even though there have been some studies done and data about the impact of socioeconomic status and race/ethnicity on many diseases' onsets are accessible, they are not all-inclusive. Furthermore, there aren't any data sets available for many developing and underdeveloped countries.

B. CHALLENGES IN GAN

Despite the usefulness of GAN, there remain obstacles to overcome before it can be used effectively in medical imaging research.

1) OPTIMIZING METHOD

The sophistication of the optimization problem is the most typical GAN problem. The Nash equilibrium represents the ideal state in a mini-max game where each participant must achieve optimum cost. However, achieving this balance is challenging.

2) MODE COLLAPSE

During the training phase, GAN is susceptible to mode collapse, which causes the creation of pictures with an odd look or just one kind. Several reasons for the collapse include instability and declining gradients, converging issues, and generator dominance. Generator domination indicates that GAN is learned in alternation with discriminator. However, the generator could be dominant. GAN lacks clear mathematical evidence for adversarial learning and structure change. Medical pictures with extensive dimensions and intricate patterns are prone to pattern collapse.

3) CROSS DOMAIN INTEGRATION

With the fast growth of computer hardware, fully convolutional training time will be drastically decreased. Instances in which GANs are merged with networks from these other fields. However, the incorporation and mixing of better models from other domains are currently restricted in medical image enhancement.

4) CLASS LEAKING

Another problem for GAN is class leaking, which occurs when an image created from one class contains attributes from another class. This topic is complex because it necessitates the definition of suitable metrics to restrict the created classes better and avoid property mixing.

5) METRICS FOR EVALUATION

The first obstacle is the absence of standards to assess the quality of synthetic pictures. Presently, medical image production has been proved to be aesthetically compelling. However, the quality of created pictures is still evaluated using conventional indices, such as Mean Square Error, peak signal-to-noise ratio, and structural similarity index measure. These assessment indices are insufficiently objective to assess the quality of health image production [104]. Future trends include GAN loss functions and assessment indices appropriate for medical imaging. Unlike authentic images, medical visuals often have a complex structure, and the particulars might transmit essential pathological data. Additionally, there are substantial distinctions between the different types of healthcare images.

6) PRACTICAL IMPLEMENTATION

Medical imaging uses of GAN are primarily in the lab setting and have not yet entered the clinical phase. The combination of GAN with surgical simulation is effective. Generated pictures may guide and imitate surgery, especially those in three dimensions. In this environment, there is a need for more collaboration and communication with physicians.

VII. DISCUSSION

Increasingly, deep learning approaches are being used in the medical field for various reasons, including disease classification and prognosis, monitoring patients, and systems that support clinical decisions. In addition, the increasing use of remote surveillance medical devices as part of the IoMT has eased the retrieval of health information by enabling constant monitoring and immediate data services by healthcare practitioners. However, the obtained data may not be adequate to construct correct algorithms due to potential problems in real-world environments, such as loss of connection, rough usage, abuse, or poor adherence to a surveillance system. Therefore, to train neural network models, data augmentation methods may be utilized to generate synthetic datasets of suitable size. A recent examination of AI models divided across racial subpopulations has shown disparities in how individuals are evaluated, treated, and billed for their healthcare expenses [101]. In deep learning for medical data, fairness issues have been observed, leading to the unjust distribution of scarce medical resources or high health risks for specific populations. Considering this, the healthcare sector has lately shown a growing interest in addressing issues of justice. However, the convergence between deep learning in the medical sector and fairness in deep learning has not yet been thoroughly investigated. In [102], have established the bridge by identifying fairness issues, outlining probable biases, classifying mitigation techniques, and highlighting future difficulties and possibilities.

Increasing the number of training examples is the objective of medical image improvement. After research and assessment, GAN-based augmented networks can generate new training samples for the two reasons outlined below. First, conventional, and improved models may be optimally trained on medium-sized datasets, including medical photos and metadata. Using the upgraded approach to expand the amount of the data set enables the model to be trained on the original foundation more effectively and achieve more robust performance. Second, GAN models will adapt to data with a high number of pictures but few labels, but it may be challenging to train traditional models to perform well on such datasets. For example, more endoscopic pictures of the digestive system without illness and fewer with tumours and accompanying labelling. A blend-GAN may utilize synthetic pictures and many authentic photos as positive and negative instances, accordingly, and may employ adversarial learning to enhance the texture consistency of pictures. GANs might be unable to produce sufficiently faultless synthetic pictures for datasets with particular challenges within the period. However, synthetic pictures share many characteristics with genuine photos and might thus be used to train future models. The performance of subsequent models may be boosted in the presence of augmented images, as opposed to their absence. For instance, researchers may filter synthetic images and choose higher-quality photos for further task training. The existing digital medical is sufficient for augmentation using GAN. The review showed that model accuracy was improved with the synthesized images compared to the existing dataset-based operations.

GAN's Contribution to Digital Health Data Poverty is through Segmentation, Cross modality transfer, Classification, Detection, Registration, Denoising, and Reconstruction.

In [103] conditional Generative Adversarial Networks are employed to produce fresh synthetic fair data with chosen attributes of the original data. A method for evaluating data biases is essential for determining the quantity and kind of synthetically collected and labelled data required for each demographic group. The experimental findings demonstrate that the suggested strategy may effectively minimize various forms of biases while simultaneously improving the predictive performance of the model. In this study, they explore the possibility of using AI to identify and correct biases by focusing on biased information, the fundamental source of bias in AI. In [104] a broad method, debiased-GAN, to solve this issue by intentionally supplementing an NLP dataset with unbalanced examples. The study was used for finding the race in tweets. They define bias as learnt connections between the race of the user and the conversational style of tweets. That, relative to tweets by African American individuals, tweets by white users are much more frequently connected with conversational tweets. Because biases frequently take the form of

explicit or implicit assumptions, the learnt connections in the study are comparable to those in real life. The synthetic data is created using a language model which creates realistic tweets that are ethnicity low vision. Debiased LMs using generative adversarial networks using reinforcement learning. A reward is supplied by a classifier separately trained to determine the ethnicity of Twitter users based on their tweets. The debias-GAN may increase the classifier's fairness measures by a factor of up to seven while retaining classification results. The GAN algorithm could be applied to medical data to limit the impact of sensitive variables such as race, gender, etc. Likewise, the algorithm will lessen the link between sensitive features and biased outcomes.

While addressing racial inequalities in digital health would be neither easy nor clear, a substantial amount of scientific evidence, considerably greater than most academics and practitioners realize, has indicated critical goals for holistic efforts to minimize and eventually eradicate health inequities [105], [106]. We have two options for achieving the suggested objective. One is to lessen the dominance of one type of group by reducing bias in the underlying model, and the other is to incorporate a sub-model to boost bias in favour of the weaker group. The first method is known as Controlled GAN (CGAN), and the second is known as VAE-GAN. By using Class-Conditional GAN, the diversity of a class can be improved.

A multi-class classifier [107] is employed as a baseline against which the algorithm's performance may be measured. Moreover, the GAN technique is used with the classifier to reduce its output bias. In this research, the method is judged valid if the p-percentage exceeds 80 per cent. The p-% of the classifier, which was initially 39% for race as well as 30% for gender, climbed to 76% for race and 82% for gender following the use of the GAN method developed in this work. Whereas the p-% for race also isn't considered fair, it has increased significantly since the GAN was implemented. In the future, a more varied dataset might be utilized with the GAN algorithm to improve the p-percentage to be more equitable. GAN algorithm may be used to résumé data to reduce the impact of sensitive characteristics, such as race and gender, on digital health. Similarly, the algorithm may lessen the association between sensitive characteristics and biased outcomes for a variety of additional applications.

By analyzing and manipulating the racial distribution of different learning datasets, the effects of different training distributions on produced picture quality as well as the racial distributions of the synthesized images are studied. The racial compositions of produced pictures are faithful to the training examples. Moreover, it was noticed that truncation, a method employed to create pictures of greater quality in interpretation, magnifies racial disparities in the information.

In the analyzed UTKFace dataset, the linked VAE has high fairness for gender and education levels but does not compare to the regular VAE for other variables, notably gender and race variables, as measured by the demographic's parity differential metric. It implies that the suggested linked VAE has been stated to be robust and accurate, but fairness is still an issue that must be addressed and enhanced. In the future, researchers may be able to improve the autoencoder gradient descent against biased attribute types in digital health data. Researchers could research several forms of fairness definitions, such as equality of opportunity and fair subgroups accuracy, as well as the fairness of the model.

Several works in modern medical image enhancement research do not employ GAN as the reference model. Instead, it may be disassembled by researchers and included in GANs. The present refinement and development of computer-assisted diagnostic imaging technology enable the acquisition of multimodal pictures of patients. Consequently, improved diagnostic and treatment procedures based on combined medical pictures have gradually attracted the interest of scientists, and the need for multimodal healthcare image enhancement has evolved appropriately. Picture generation has been employed for 4D medical image problems; however, the present work is not intended for dataset enhancement but gives a strong illustration of GAN in health image augmentation.

VIII. FUTURE DIRECTION

In the future, improved models would be advantageous for jobs like predicting 4D time series. Supported diagnostic and treatment strategies based on multimodal medical pictures have steadily drawn researchers' interest, increasing the demand for innovative multimodal healthcare for image analysis. Imaging creation has been employed for 4D medical image tasks, but this study mainly provides the impact of GAN-based health data augmentation, thus overcoming digital data poverty. It is envisaged that future uses of augmented models will include 4D time series prediction.

There are several efficient data augmentation techniques other than GAN (Generative Adversarial Networks) that can be used to increase the size of a dataset and reduce the data imbalance. Here are a few examples:

A. COLOUR JITTERING

This technique involves applying random colour transformations to images, such as changing brightness, contrast, and saturation. This can be used to make models more robust to lighting conditions.

B. ADDING NOISE

This technique involves adding random noise to images, such as Gaussian or salt-and-pepper noise. This can be used to simulate real-world conditions and improve model robustness to noise.

C. MIXUP

This technique involves linearly interpolating between two randomly selected images to create a new image. This can be used to create new images that are a combination of multiple images in the dataset, which can improve model performance.

TABLE 5. Direct methods for data generation and augmentation.

Organ	Dataset	Model	Remarks	Reference
Head /Neck -Brain	(BraTS) 2020 dataset, brain tumour dataset	AGGrGAN-(DCGAN +WGAN)	style transfer	[37]
	MRI-Based Brain Tumour	Vanilla GAN and DCGAN	Automatic image quality checking using CNN, MobileNetV2, and ResNet152V2	[38]
	Tumour Public dataset	cGAN	Detection and Classification. Accuracy measure Detection -99% Classification - 98%	[39]
	BRATS	GANs	Sensitivity 97.48%	[40]
	BRATS	CPGGANs	Accuracy improves by 0.64%	[41]
	Brain Axial MR Images	CPGGANs	Sensitivity 91%	[42]
	TCGA-GBM TCGA-LGG	GAN	Accuracy 88.82%	[43]
	Nanfang Hospital General Hospital MRI Brain Images during 2005-10	GAN	Accuracy 98.57%	[44]
	Figshare BRATS	MSG GAN	Accuracy 88.7%	[45]
	1133 private images	MAD GAN	Accuracy 92.1%	[46]
	ADNI database PET images	DCGAN	71.45% classification accuracy	[47]
	Private Dataset	GAN	Accuracy 93%	[48]
Abdomen - Breast	DDSM + CBIS and MIAS	GAN	image enhancement technique	[49]
	INBreast	blend-GAN	Detection 0.01 Test positivity Rate	[50]
	DDSM	NcGAN	Detection 0.013- Area under the curve	[51]
	DDSM	cGAN	Classification 0.009- Area under the curve	[52]
	INBreast	NcGAN	Classification 0.03- Area under the curve	[53]
	DDSM	GAN	Classification 0.002- Area under the curve	[54]
	Private dataset with 357 images	semi-supervised GAN	Classification accuracy 90.41%	[55]
	Private dataset	Conditional GAN	69.59% pixel-wise F1 score	[56]
	METABRIC (MB) dataset	WGAN	78% accuracy	[57]
	UKE dataset	Cycle GAN	Precision 82%	[58]
	ACDC	tsf-GAN	Lesion segmentation	[59]

TABLE 5. *(Continued.)* **Direct methods for data generation and augmentation.**

Abdomen - Heart	ACDC	tsf-GAN	Cardiac segmentation	[60]
	ACDC /SCD	tsf-GAN	Cardiac segmentation	[61]
	ACDC	cGAN	Cardiac segmentation	[62]
	CTG signal dataset	TSGAN	solves data imbalance. Increase in the Quality Index by 44% from other models.	[63]
	Private dataset	3D GAN	Segmentation quantitative analysis was done to show the effectiveness of GAN	[64]
	200 real FHR data	CCWGAN-GP	Accuracy improvement of 12%	[65]
	PTB Diagnostic ECG Database	SLC-GAN	Overcome data shortage and imbalance in ECG classification 99.06% accuracy	[66]
	Cleveland and Statlog datasets	GAN	99.3% accuracy	[67]
Abdomen	LiTS , IRCAD	tGAN	Vessel segmentation Differential scanning calorimetry 0.02 and 0.06	[68]
	Private dataset	NcGAN	Liver tumour classification Sensitivity improvement by 6.3%	[69]
	Private dataset	NcGAN	Liver tumour classification Sensitivity improvement by 4.4%	[70]
	Private dataset	tGAN	Liver registration	[71]
	Private dataset of endoscopic images of gastro	GAN	Lesion location detection	[72]
	CBIS-DDMS	TMP-GAN	lesion detection augmented dataset improves by the precision of 2.59%	[73]
	Private Dataset	GAN	synthesizes high-resolution virtual contrast CTs.	[74]
	434 patient's Private dataset	GAN	AUC – 83.2	[75]
Thorax	NIH	cGAN	Xray Lesion Detection- 0.041 Accuracy improvement	[76]
	RSNA	NcGAN	X-ray Disease Classification 2.02% Accuracy improvement	[77]
	CXR	NcGAN	X-ray Disease Classification	[78]
	Covid_1,2,3	cGAN	X-ray Disease Classification 10% Accuracy improvement	[79]

TABLE 5. *(Continued.)* **Direct methods for data generation and augmentation.**

	Private DataSet	NcGAN	X-ray Disease Classification 94.5% accuracy	[80]
	LIDC-IDRI	Blend GAN	location detection- Lung CT lesion average CPM improvement by 0.03.	[81]
	LIDC-IDRI	cGAN	Lung CT Lesion contour segmentation	[82]
	LIDC-IDRI	cGAN	Lung CT Lesion contour segmentation	[83]
	LUNA	cGAN	Classification Accuracy improvement of 1%	[84]
	Private Dataset	NcGAN	Classification Accuracy improvement by 5%	[85]
	Private Dataset	NcGAN	Classification Accuracy improvement by 6%	[86]
	LIDC-IDRI	cGAN	Classification Accuracy 1.17%	[87]
	Fujita Health University Hospital 133 lung images	StyleGAN	Obtained multiple images from a single sketch	[88]
	673 lung cancer CT	SLS-PriGAN	knowledge acquisition network accuracy 91%,	[89]
	COVID-19 CT segmentation dataset	cGAN	lung lesion segmentation Accuracy 99.87% loss function learns based on the entire image	[90]
	Dataset of NLM, USA	Cyclic GAN	Accuracy 97.19%	[91]
	Covidx public dataset	U-Net GAN	-	[92]

D. CUTOUT

This technique involves masking out a random portion of an image and replacing it with black pixels. This can be used to create new images and improve model robustness to occlusions.

These techniques can be used individually or in combination to create a larger and more diverse data set, which can improve the performance of the model. In future work, we will use other data augmentation techniques to solve the data poverty issue.

1) DIRECT METHODS FOR DATA GENERATION AND AUGMENTATION

Data Augmentation study is broadly classified based on different techniques in augmentation based on the dataset and diseases on which data is augmented. The prime goal of GAN deployment in digital health is to enhance the effectiveness of models, including classification and classification algorithms. However, insufficient dataset sizes and poor image quality diminish the performance of digital health deep learning models. Therefore, pseudo-healthy image creation may also be seen as data augmentation.

2) CLASSIFICATION BASED ON THE METHODS OF DATA AUGMENTATION

Super-resolution uper-resolution techniques generate higher images from limited images to get extra data. Using GAN technology, the super-resolution of health data primarily depends on the generator's capabilities. Combinations of low- and high-resolution images are often used for GAN-based

TABLE 6. Recorded observations for classification.

Classification Model	Validation/ Train Accuracy	Test Accuracy
SIFT with SVM	0.56	0.54
Xception	0.74	0.72
DenseNet 121	0.74	0.73
DenseNet 169 (40% GAN Augmentation)	0.80	0.80
DenseNet 201 (50% GAN Augmentation)	0.81	0.81
Mean Ensemble (Non-GAN trained CNN)	0.83	0.81
Mean Ensemble (GAN-trained CNN)	0.86	0.85
SVM Ensemble	0.86	0.86

TABLE 7. Results using SVM ensemble – maximum accuracy achieved for each epoch (Max. 5 Epochs).

Epoch Number	Iteration Range	Max. Accuracy Achieved
1	0 - 350	0.68
2	351 – 700	0.74
3	701 – 1050	0.78
4	1051 – 1400	0.83
5	1401 - 1750	0.85

medical image super-resolution. Almalioglu et al. [43] designed a hybrid loss function with excellent fidelity for a framework that combines an attention mechanism with Conditional GAN for endoscopic pictures.

3) DATASET EXPANSION

Like a data expansion technique, GAN can generate medical pictures that seem somewhat lifelike to the human eye [44]. This section analyses the effect of synthetic

TABLE 8. Improvement in less representative classes using GAN.

Classes	Data Percentage without GAN Augmentation	Data Percentage with GAN Augmentation
Basal Cell Carcinoma	29.65	47.05
Actinic Keratosis	12.95	29.41
Vascular Lesion	12.95	10.16
Dermatofibroma	10.49	13.36
Basal Cell Carcinoma	29.65	47.05

TABLE 9. Types of skin lesions.

S.No	Abbreviation	Lesion Type
1	AKIEC	Actinic Keratosis
2	BCC	Basal Cell Carcinoma
3	BKL	Benign Keratosis
4	DF	Dermatofibroma
5	MEL	Melanoma
6	NV	Melanocytic Nevi
7	VASC	Vascular Lesion

TABLE 10. Evaluation metrics for each lesion type.

S.No	Class/ Lesion Type	Precision	Recall	F1 - Score
1	AKIEC	0.71	0.70	0.7050
2	BCC	0.82	0.91	0.8627
3	BKL	0.78	0.83	0.8042
4	DF	0.86	0.77	0.8125
5	MEL	0.8	0.78	0.7899
6	NV	0.93	0.96	0.9448
7	VASC	0.87	0.94	0.9036

pictures on the precision of deep learning models. Typically, there are a few medical photos with annotations. To augment datasets, GAN may create pictures with specified labels. Diaz-Pinto et al. [45] provided a precise technique for assessing glaucoma using GAN and semi-supervised learning. Not only can the system produce pictures synthetically, but it can also classify them automatically. Because of privacy concerns and medical organizations' reluctance to share data, health datasets tend to be unbalanced. Salehinejad et al. [46] used Deep CNNs, which were trained to recognize the genesis of five kinds of CXR pictures using a mix of actual and fake data. The findings demonstrate that these systems outperformed comparable ones trained solely with genuine pictures. It is unlikely that every synthetic picture is acceptable for the

training procedure; hence it is also required to implement effective selection methods for synthetic images.

4) PSEUDO-HEALTHY IMAGE SYNTHESIS

The goal of pseudo-healthy data synthesis is to produce a healthy picture from an unhealthy one. Discovering irregularities and grasping disease-related changes is possible using healthy synthetic images. The method proposed in [47] proposes a straightforward and effective limitation to enhance the mapping of abnormal to normal by setting stricter limits on the generator. Typically, domain transformation alters the whole appearance of a picture, but in this manner, the generator can just modify the aberrant portion. Consequently, it must be susceptible to the features of aberrant portions. Some of the significant works in recent times based on organs are listed in Table 5.

IX. CONCLUSION

In this work, several augmentation techniques using artificial data derived from generative adversarial networks (GANs) were examined. These ensemble techniques help us accurately characterize images of skin lesions, along with generating photo - realistic synthetic images of the input data. Results also reveal that adding GAN-produced image samples to the training data does significantly improve performance when compared to the conventional way of optimizing the current deep neural network architectures. In application domains like medical imaging, where sizable training datasets are often not easily available, these and other unique data generation and augmentation techniques can be very helpful. Further, scaling up the resolution of the images produced by the AC-GAN while maintaining significant aspects of the training data could enhance the training of the CNNs, albeit in our testing, higher-resolution GANs were quite unstable. Different GANs' generation abilities can be assessed in subsequent attempts at high-resolution image generation by their contribution to increasing the classification task's accuracy.

REFERENCES

[1] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014, *arXiv:1406.2661*.

[2] A. P. Lozano. (Sep. 30, 2022). *Apolanco3225/Data-Augmentation-and-Segmentation-with-GANs-for-Medical-Images*. Accessed: Oct. 31, 2022. [Online]. Available: https://github.com/apolanco3225/Data-Augmentation-and-Segmentation-with-GANs-for-Medical-Images/blob/1864b79160ff513a8b6a3d918348478919219178/startGANs.ipynb

[3] (2022). *COVID-19 Chest CT Image Augmentation GAN Dataset*. Accessed: Oct. 31, 2022. [Online]. Available: https://www.kaggle.com/datasets/mloey1/covid19-chest-ct-image-augmentation-gan-dataset

[4] S. Karlsson. (Oct. 27, 2022). *Simontomaskarlsson/GAN-MRI*. Accessed: Oct. 31, 2022. [Online]. Available: https://github.com/simontomaskarlsson/GAN-MRI

[5] KSH0660. (Oct. 28, 2021). *KSH0660/BrainTumor*. Accessed: Oct. 31, 2022. [Online]. Available: https://github.com/KSH0660/BrainTumor

[6] BatmanLab. (Oct. 21, 2022). *Hierarchical Amortized GAN (HA-GAN)*. Accessed: Oct. 31, 2022. [Online]. Available: https://github.com/batmanlab/HA-GAN

[7] N. Essipova. (Oct. 13, 2022). *GAN Synthetic Medical Image Augmentation*. Accessed: Oct. 31, 2022. [Online]. Available: https://github.com/NicoEssi/GAN_Synthetic_Medical_Image_Augmentation

[8] Nikhil. (Oct. 14, 2022). *Nikkkkhil/GAN MRI Data Augmentation*. Accessed: Oct. 31, 2022. [Online]. Available: https://github.com/nikkkkhil/GAN-MRI-data-augmentation

[9] Y. Niu. (Oct. 26, 2022). *Patho-GAN: Interpretation + Medical Data Augmentation*. Accessed: Oct. 31, 2022. [Online]. Available: https://github.com/zzdyyy/Patho-GAN

[10] M. Craig. (May 16, 2022). *GAN-Mammo*. Accessed: Oct. 31, 2022. [Online]. Available: https://github.com/craigmichaelm/GAN-Mammo

[11] S. Rossetti. (Sep. 29, 2022). *GAN-based Synthetic Medical Image Augmentation for Increased CNN Performance in COVID-19 Classification*. Accessed: Oct. 31, 2022. [Online]. Available: https://github.com/rossettisimone/AUGMENTATION_GAN

[12] (2022). *Synthesizing Images With GANs*. Accessed: Oct. 31, 2022. [Online]. Available: https://web.bii.a-star.edu.sg/archive/machine_learning/Projects/filaStructObjs/Synthesis/downloads.html

[13] (2022). *Melanoma and GAN*. Accessed: Oct. 31, 2022. [Online]. Available: https://kaggle.com/code/sandhiwangiyana/melanoma-and-gan

[14] M. Robin, J. John, and A. Ravikumar, "Breast tumor segmentation using U-NET," in *Proc. 5th Int. Conf. Comput. Methodologies Commun. (ICCMC)*, Apr. 2021, pp. 1164–1167, doi: 10.1109/ICCMC51019.2021.9418447.

[15] I. A. Corley and Y. Huang, "Deep EEG super-resolution: Upsampling EEG spatial resolution with generative adversarial networks," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Mar. 2018, pp. 100–103.

[16] Y. Murakami, T. Magome, K. Matsumoto, T. Sato, Y. Yoshioka, and M. Oguchi, "Fully automated dose prediction using generative adversarial networks in prostate cancer patients," *PLoS ONE*, vol. 15, no. 5, May 2020, Art. no. e0232697, doi: 10.1371/journal.pone.0232697.

[17] T. Uemura, J. J. Näppi, C. Watari, T. Hironaka, T. Kamiya, and H. Yoshida, "Weakly unsupervised conditional generative adversarial network for image-based prognostic prediction for COVID-19 patients based on chest CT," *Med. Image Anal.*, vol. 73, Oct. 2021, Art. no. 102159, doi: 10.1016/j.media.2021.102159.

[18] T. Schlegl, P. Seeböck, S. M. Waldstein, G. Langs, and U. Schmidt-Erfurth, "F-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks," *Med. Image Anal.*, vol. 54, pp. 30–44, May 2019, doi: 10.1016/j.media.2019.01.010.

[19] L. Lan, L. You, Z. Zhang, Z. Fan, W. Zhao, N. Zeng, Y. Chen, and X. Zhou, "Generative adversarial networks and its applications in biomedical informatics," *Frontiers Public Health*, vol. 8, pp. 1–14, May 2020. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpubh.2020.00164

[20] W. Njima, A. Bazzi, and M. Chafii, "DNN-based indoor localization under limited dataset using GANs and semi-supervised learning," *IEEE Access*, vol. 10, pp. 69896–69909, 2022, doi: 10.1109/ACCESS.2022.3187837.

[21] C. Chen, Q. Dou, H. Chen, and P.-A. Heng, "Semantic-aware generative adversarial nets for unsupervised domain adaptation in chest X-ray segmentation," 2018, *arXiv:1806.00600*.

[22] D. Mahapatra, B. Bozorgtabar, and R. Garnavi, "Image super-resolution using progressive generative adversarial networks for medical image analysis," *Computerized Med. Imag. Graph.*, vol. 71, pp. 30–39, Jan. 2019, doi: 10.1016/j.compmedimag.2018.10.005.

[23] X. Zhou, X. Liu, G. Lan, and J. Wu, "Federated conditional generative adversarial nets imputation method for air quality missing data," *Knowl.-Based Syst.*, vol. 228, Sep. 2021, Art. no. 107261, doi: 10.1016/j.knosys.2021.107261.

[24] S. You, B. H. Cho, S. Yook, J. Y. Kim, Y.-M. Shon, D.-W. Seo, and I. Y. Kim, "Unsupervised automatic seizure detection for focal-onset seizures recorded with behind-the-ear EEG using an anomaly-detecting generative adversarial network," *Comput. Methods Programs Biomed.*, vol. 193, Sep. 2020, Art. no. 105472, doi: 10.1016/j.cmpb.2020.105472.

[25] S. Harini and A. Ravikumar, "Effect of parallel workload on dynamic voltage frequency scaling for dark silicon ameliorating," in *Proc. Int. Conf. Smart Electron. Commun. (ICOSEC)*, Sep. 2020, pp. 1012–1017, doi: 10.1109/icosec49089.2020.9215262.

[26] J. John, A. Ravikumar, and B. Abraham, "Prostate cancer prediction from multiple pretrained computer vision model," *Health Technol.*, vol. 11, no. 5, pp. 1003–1011, Sep. 2021, doi: 10.1007/s12553-021-00586-y.

[27] J. Mao, H. Wang, and B. F. Spencer, "Toward data anomaly detection for automated structural health monitoring: Exploiting generative adversarial nets and autoencoders," *Structural Health Monitor.*, vol. 20, no. 4, pp. 1609–1626, Jul. 2021, doi: 10.1177/1475921720924601.

[28] V. Alex, K. P. M. Safwan, S. S. Chennamsetty, and G. Krishnamurthi, "Generative adversarial networks for brain lesion detection," *Proc. SPIE*, vol. 10133, p. 101330, Feb. 2017, doi: 10.1117/12.2254487.

[29] K. Naidoo and V. Marivate, "Unsupervised anomaly detection of healthcare providers using generative adversarial networks," in *Responsible Design, Implementation and Use of Information and Communication Technology* (Lecture Notes in Computer Science), M. Hattingh, M. Matthee, H. Smuts, I. Pappas, Y. K. Dwivedi, and M. Mäntymäki, Eds. Cham, Switzerland: Springer, 2020, pp. 419–430, doi: 10.1007/978-3-030-44999-5_35.

[30] I. Vaccari, V. Orani, A. Paglialonga, E. Cambiaso, and M. Mongelli, "A generative adversarial network (GAN) technique for Internet of Medical Things data," *Sensors*, vol. 21, no. 11, p. 3726, May 2021, doi: 10.3390/s21113726.

[31] A. Perez, S. Ganguli, S. Ermon, G. Azzari, M. Burke, and D. Lobell, "Semi-supervised multitask learning on multispectral satellite images using Wasserstein generative adversarial networks (GANs) for predicting poverty," 2019, *arXiv:1902.11110*.

[32] J. Yoon, L. N. Drumright, and M. van der Schaar, "Anonymization through data synthesis using generative adversarial networks (ADS-GAN)," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 8, pp. 2378–2388, Aug. 2020, doi: 10.1109/JBHI.2020.2980262.

[33] (2022). *Indoor Localization Using Data Augmentation via Selective Generative Adversarial Networks—Princeton University*. Accessed: Sep. 27, 2022. [Online]. Available: https://collaborate.princeton.edu/en/publications/indoor-localization-using-data-augmentation-via-selective-generat

[34] S. Venkatasubramanian, "Ambulatory monitoring of maternal and fetal using deep convolution generative adversarial network for smart health care IoT system," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 1, 2022, doi: 10.14569/ijacsa.2022.0130126.

[35] D. I. Morís, J. de Moura, J. Novo, and M. Ortega, "Portable chest X-ray synthetic image generation for the COVID-19 screening," *Eng. Proc.*, vol. 7, no. 1, pp. 1–9, 2021, Art. no. 6, doi: 10.3390/engproc2021007006.

[36] Q. Wang, W. Xue, X. Zhang, F. Jin, and J. Hahn, "Pixel-wise body composition prediction with a multi-task conditional generative adversarial network," *J. Biomed. Informat.*, vol. 120, Aug. 2021, Art. no. 103866, doi: 10.1016/j.jbi.2021.103866.

[37] T. Zhu, X. Yao, K. Li, P. Herrero, and P. Georgiou, "Blood glucose prediction for type 1 diabetes using generative adversarial networks," in *CEUR Workshop Proc.*, 2020, pp. 1–5.

[38] F. Munawar, S. Azmat, T. Iqbal, C. Grönlund, and H. Ali, "Segmentation of lungs in chest X-ray image using generative adversarial networks," *IEEE Access*, vol. 8, pp. 153535–153545, 2020, doi: 10.1109/ACCESS.2020.3017915.

[39] G. Lazaridis, M. Lorenzi, S. Ourselin, and D. Garway-Heath, "Improving statistical power of glaucoma clinical trials using an ensemble of cyclical generative adversarial networks," *Med. Image Anal.*, vol. 68, Feb. 2021, Art. no. 101906, doi: 10.1016/j.media.2020.101906.

[40] C. Wan and D. T. Jones, "Protein function prediction is improved by creating synthetic feature samples with generative adversarial networks," *Nature Mach. Intell.*, vol. 2, no. 9, pp. 540–550, Aug. 2020.

[41] W. Lee, J. Lee, and Y. Kim, "Contextual imputation with missing sequence of EEG signals using generative adversarial networks," *IEEE Access*, vol. 9, pp. 151753–151765, 2021, doi: 10.1109/ACCESS.2021.3126345.

[42] S. U. H. Dar, M. Yurt, M. Shahdloo, M. E. Ildiz, B. Tinaz, and T. Çukur, "Prior-guided image reconstruction for accelerated multi-contrast MRI via generative adversarial networks," *IEEE J. Sel. Topics Signal Process.*, vol. 14, no. 6, pp. 1072–1087, Oct. 2020, doi: 10.1109/JSTSP.2020.3001737.

[43] Y. Almalioglu, K. B. Ozyoruk, A. Gokce, K. Incetan, G. Irem Gokceler, M. Ali Simsek, K. Ararat, R. J. Chen, N. J. Durr, F. Mahmood, and M. Turan, "EndoL2H: Deep super-resolution for capsule endoscopy," *IEEE Trans. Med. Imag.*, vol. 39, no. 12, pp. 4297–4309, Dec. 2020, doi: 10.1109/TMI.2020.3016744.

[44] Y. Jiang, S. Yin, and O. Kaynak, "Data-driven monitoring and safety control of industrial cyber-physical systems: Basics and beyond," *IEEE Access*, vol. 6, pp. 47374–47384, 2018, doi: 10.1109/ACCESS.2018.2866403.

[45] A. Diaz-Pinto, A. Colomer, V. Naranjo, S. Morales, Y. Xu, and A. F. Frangi, "Retinal image synthesis and semi-supervised learning for glaucoma assessment," *IEEE Trans. Med. Imag.*, vol. 38, no. 9, pp. 2211–2218, Sep. 2019, doi: 10.1109/TMI.2019.2903434.

[46] H. Salehinejad, E. Colak, T. Dowdell, J. Barfett, and S. Valaee, "Synthesizing chest X-ray pathology for training deep convolutional neural networks," *IEEE Trans. Med. Imag.*, vol. 38, no. 5, pp. 1197–1206, May 2019, doi: 10.1109/TMI.2018.2881415.

[47] X. Chen and E. Konukoglu, "Unsupervised detection of lesions in brain MRI using constrained adversarial auto-encoders," 2018, *arXiv:1806.04972*.

[48] D. Mukherkjee, P. Saha, D. Kaplun, A. Sinitca, and R. Sarkar, "Brain tumor image generation using an aggregation of GAN models with style transfer," *Sci. Rep.*, vol. 12, no. 1, Jun. 2022, Art. no. 1, doi: 10.1038/s41598-022-12646-y.

[49] H. H. N. Alrashedy, A. F. Almansour, D. M. Ibrahim, and M. A. A. Hammoudeh, "BrainGAN: Brain MRI image generation and classification framework using GAN architectures and CNN models," *Sensors*, vol. 22, no. 11, p. 4297, Jun. 2022, doi: 10.3390/s22114297.

[50] R. K. Gupta, S. Bharti, N. Kunhare, Y. Sahu, and N. Pathik, "Brain tumor detection and classification using cycle generative adversarial networks," *Interdiscipl. Sci., Comput. Life Sci.*, vol. 14, no. 2, pp. 485–502, Jun. 2022, doi: 10.1007/s12539-022-00502-6.

[51] C. Han, L. Rundo, R. Araki, Y. Nagano, Y. Furukawa, G. Mauri, H. Nakayama, and H. Hayashi, "Combining noise-to-image and image-to-image GANs: Brain MR image augmentation for tumor detection," *IEEE Access*, vol. 7, pp. 156966–156977, 2019, doi: 10.1109/ACCESS.2019.2947606.

[52] C. Han, L. Rundo, R. Araki, Y. Furukawa, G. Mauri, H. Nakayama, and H. Hayashi, "Infinite brain MR images: PGGAN-based data augmentation for tumor detection," 2019, *arXiv:1903.12564*.

[53] C. Han, K. Murao, T. Noguchi, Y. Kawata, F. Uchiyama, L. Rundo, H. Nakayama, and S. Satoh, "Learning more with less: Conditional PGGAN-based data augmentation for brain metastases detection using highly-rough annotation on MR images," in *Proc. 28th ACM Int. Conf. Inf. Knowl. Manage.*, Nov. 2019, pp. 119–127, doi: 10.1145/3357384.3357890.

[54] C. Ge, I. Y. Gu, A. S. Jakola, and J. Yang, "Enlarged training dataset by pairwise GANs for molecular-based brain tumor classification," *IEEE Access*, vol. 8, pp. 22560–22570, 2020, doi: 10.1109/ACCESS.2020.2969805.

[55] N. Ghassemi, A. Shoeibi, and M. Rouhani, "Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images," *Biomed. Signal Process. Control*, vol. 57, Mar. 2020, Art. no. 101678, doi: 10.1016/j.bspc.2019.101678.

[56] S. Deepak and P. M. Ameer, "MSG-GAN based synthesis of brain MRI with meningioma for data augmentation," in *Proc. IEEE Int. Conf. Electron., Comput. Commun. Technol. (CONECCT)*, Jul. 2020, pp. 1–6.

[57] C. Han, L. Rundo, K. Murao, T. Noguchi, Y. Shimahara, Z. Á. Milacski, S. Koshino, E. Sala, H. Nakayama, and S. Satoh, "MADGAN: Unsupervised medical anomaly detection GAN using multiple adjacent brain MRI slice reconstruction," *BMC Bioinf.*, vol. 22, no. S2, p. 31, Apr. 2021, doi: 10.1186/s12859-020-03936-1.

[58] J. Islam and Y. Zhang, "GAN-based synthetic brain PET image generation," *Brain Informat.*, vol. 7, no. 1, p. 3, Mar. 2020, doi: 10.1186/s40708-020-00104-2.

[59] M. Waters, M. Inkman, N. Andruska, R. Brenneman, S. S. Markovina, J. K. Schwarz, and J. Zhang, "Generative adversarial neural networks augment marker and pathway analysis of treatment resistant HPV+ head and neck squamous cell carcinoma.," *J. Clin. Oncol.*, vol. 40, no. 16, Jun. 2022, Art. no. e18059, doi: 10.1200/jco.2022.40.16suppl.e18059.

[60] O. N. Oyelade and A. E. Ezugwu, "A novel wavelet decomposition and transformation convolutional neural network with data augmentation for breast cancer detection using digital mammogram," *Sci. Rep.*, vol. 12, no. 1, Apr. 2022, Art. no. 1, doi: 10.1038/s41598-022-09905-3.

[61] C. Lin, R. Tang, D. D. Lin, L. Liu, J. Lu, Y. Chen, D. Gao, and J. Zhou, "Breast mass detection in mammograms via blending adversarial learning," in *Simulation and Synthesis in Medical Imaging* (Lecture Notes in Computer Science), vol. 11827. Cham, Switzerland: Springer, 2019, pp. 52–61, doi: 10.1007/978-3-030-32778-1_6.

[62] T. Shen, K. Hao, C. Gou, and F.-Y. Wang, "Mass image synthesis in mammogram with contextual information based on GANs," *Comput. Methods Programs Biomed.*, vol. 202, Apr. 2021, Art. no. 106019.

[63] E. Wu, K. Wu, D. Cox, and W. Lotter, "Conditional infilling GANs for data augmentation in mammogram classification," 2018, *arXiv:1807.08093*.

[64] H. Li, D. Chen, W. H. Nailon, M. E. Davies, and D. I. Laurenson, "Signed Laplacian deep learning with adversarial augmentation for improved mammography diagnosis," in *Proc. 22nd Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, Shenzhen, China. Berlin, Germany: Springer-Verlag, Oct. 2019, pp. 486–494, doi: 10.1007/978-3-030-32226-7_54.

[65] C. Muramatsu, M. Nishio, T. Goto, M. Oiwa, T. Morita, M. Yakami, T. Kubo, K. Togashi, and H. Fujita, "Improving breast mass classification by shared data with domain transformation using a generative adversarial network," *Comput. Biol. Med.*, vol. 119, Apr. 2020, Art. no. 103698, doi: 10.1016/j.compbiomed.2020.103698.

[66] T. Pang, J. H. D. Wong, W. L. Ng, and C. S. Chan, "Semi-supervised GAN-based radiomics model for data augmentation in breast ultrasound mass classification," *Comput. Methods Programs Biomed.*, vol. 203, May 2021, Art. no. 106018, doi: 10.1016/j.cmpb.2021.106018.

[67] W. Preedanan, K. Suzuki, T. Kondo, M. Kobayashi, H. Tanaka, J. Ishioka, Y. Matsuoka, Y. Fujii, and I. Kumazawa, "Improvement of urinary stone segmentation using GAN-based urinary stones inpainting augmentation," *IEEE Access*, vol. 10, pp. 115131–115142, 2022, doi: 10.1109/ACCESS.2022.3218444.

[68] C. Guttá, C. Morhard, and M. Rehm, "Applying GAN-based data augmentation to improve transcriptome-based prognostication in breast cancer," *MedRxiv*, vol. 19, pp. 1–29, Oct. 2022, doi: 10.1101/2022.10.07.22280776.

[69] A. El-Ghoussani, D. Rodríguez-Salas, M. Seuret, and A. Maier, "GAN-based augmentation of mammograms to improve breast lesion detection," in *Bildverarbeitung Für Die Medizin* (Informatik aktuell), K. Maier-Hein, T. M. Deserno, H. Handels, A. Maier, C. Palm, and T. Tolxdorff, Eds. Wiesbaden, Germany: Springer, 2022, pp. 321–326, doi: 10.1007/978-3-658-36932-3_66.

[70] C. Chen, C. Qin, H. Qiu, C. Ouyang, S. Wang, L. Chen, G. Tarroni, W. Bai, and D. Rueckert, "Realistic adversarial data augmentation for MR image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, Lima, Peru. Berlin, Germany: Springer-Verlag, Oct. 2020, pp. 667–677, doi: 10.1007/978-3-030-59710-8_65.

[71] K. Chaitanya, N. Karani, C. F. Baumgartner, E. Erdil, A. Becker, O. Donati, and E. Konukoglu, "Semi-supervised task-driven data augmentation for medical image segmentation," *Med. Image Anal.*, vol. 68, Feb. 2021, Art. no. 101934, doi: 10.1016/j.media.2020.101934.

[72] Y. Skandarani, N. Painchaud, P.-M. Jodoin, and A. Lalande, "On the effectiveness of GAN generated cardiac MRIs for segmentation," 2020, *arXiv:2005.09026*.

[73] S. Amirrajab, S. Abbasi-Sureshjani, Y. A. Khalil, C. Lorenz, J. Weese, J. Pluim, and M. Breeuwer, "XCAT-GAN for synthesizing 3D consistent labeled cardiac MR images on anatomically variable XCAT phantoms," in *Proc. 23rd Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, Lima, Peru. Berlin, Germany: Springer-Verlag, Oct. 2020, pp. 128–137, doi: 10.1007/978-3-030-59719-1_13.

[74] R. D. I. Puspitasari, M. A. Ma'sum, M. R. Alhamidi, Kurnianingsih, and W. Jatmiko, "Generative adversarial networks for unbalanced fetal heart rate signal classification," *ICT Exp.*, vol. 8, no. 2, pp. 239–243, Jun. 2022, doi: 10.1016/j.icte.2021.06.007.

[75] C. Tiago, A. Gilbert, A. S. Beela, S. A. Aase, S. R. Snare, J. Šprem, and K. McLeod, "A data augmentation pipeline to generate synthetic labeled datasets of 3D echocardiography images using a GAN," *IEEE Access*, vol. 10, pp. 98803–98815, 2022, doi: 10.1109/ACCESS.2022.3207177.

[76] Y. Zhang, Z. Zhao, Y. Deng, and X. Zhang, "FHRGAN: Generative adversarial networks for synthetic fetal heart rate signal generation in low-resource settings," *Inf. Sci.*, vol. 594, pp. 136–150, May 2022, doi: 10.1016/j.ins.2022.01.070.

[77] W. Li, Y. M. Tang, K. M. Yu, and S. To, "SLC-GAN: An automated myocardial infarction detection model based on generative adversarial networks and convolutional neural networks with single-lead electrocardiogram synthesis," *Inf. Sci.*, vol. 589, pp. 738–750, Apr. 2022, doi: 10.1016/j.ins.2021.12.083.

[78] R. R. Sarra, A. M. Dinar, M. A. Mohammed, M. K. A. Ghani, and M. A. Albahar, "A robust framework for data generative and heart disease prediction based on efficient deep learning models," *Diagnostics*, vol. 12, no. 12, p. 2899, Nov. 2022, doi: 10.3390/diagnostics12122899.

[79] T. Russ, S. Goerttler, A.-K. Schnurr, D. F. Bauer, S. Hatamikia, L. R. Schad, F. G. Zöllner, and K. Chung, "Synthesis of CT images from digital body phantoms using CycleGAN," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 10, pp. 1741–1750, Oct. 2019, doi: 10.1007/s11548-019-02042-9.

[80] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Synthetic data augmentation using GAN for improved liver lesion classification," 2018, *arXiv:1801.02385*.

[81] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, Dec. 2018, doi: 10.1016/j.neucom.2018.09.013.

[82] D. F. Bauer, T. Russ, B. I. Waldkirch, C. Tönnes, W. P. Segars, L. R. Schad, F. G. Zöllner, and A.-K. Golla, "Generation of annotated multimodal ground truth datasets for abdominal medical image registration," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 16, no. 8, pp. 1277–1285, Aug. 2021, doi: 10.1007/s11548-021-02372-7.

[83] T. Kanayama, Y. Kurose, K. Tanaka, K. Aida, S. Satoh, M. Kitsuregawa, and T. Harada, "Gastric cancer detection from endoscopic images using synthesis by GAN," in *Proc. 22nd Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, Shenzhen, China. Berlin, Germany: Springer-Verlag, Oct. 2019, pp. 530–538, doi: 10.1007/978-3-030-32254-0_59.

[84] Q. Guan, Y. Chen, Z. Wei, A. A. Heidari, H. Hu, X.-H. Yang, J. Zheng, Q. Zhou, H. Chen, and F. Chen, "Medical image augmentation for lesion detection using a texture-constrained multichannel progressive GAN," *Comput. Biol. Med.*, vol. 145, Jun. 2022, Art. no. 105444, doi: 10.1016/j.compbiomed.2022.105444.

[85] J. Liu, Y. Tian, C. Duzgol, O. Akin, A. M. Ağıldere, K. M. Haberal, and M. Coşkun, "Virtual contrast enhancement for CT scans of abdomen and pelvis," *Computerized Med. Imag. Graph.*, vol. 100, Sep. 2022, Art. no. 102094, doi: 10.1016/j.compmedimag.2022.102094.

[86] Y.-Y. Duan, J. Qin, W.-Q. Qiu, S.-Y. Li, C. Li, A.-S. Liu, X. Chen, and C.-X. Zhang, "Performance of a generative adversarial network using ultrasound images to stage liver fibrosis and predict cirrhosis based on a deep-learning radiomics nomogram," *Clin. Radiol.*, vol. 77, no. 10, pp. e723–e731, Oct. 2022, doi: 10.1016/j.crad.2022.06.003.

[87] Y. Xing, Z. Ge, R. Zeng, D. Mahapatra, J. Seah, M. Law, and T. Drummond, "Adversarial pulmonary pathology translation for pairwise chest X-ray data augmentation," in *Proc. 22nd Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*. Shenzhen, China: Springer, 2019, pp. 757–765, doi: 10.1007/978-3-030-32226-7_84.

[88] P. Ganesan, S. Rajaraman, R. Long, B. Ghoraani, and S. Antani, "Assessment of data augmentation strategies toward performance improvement of abnormality classification in chest radiographs," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 841–844, doi: 10.1109/EMBC.2019.8857516.

[89] V. Bhagat and S. Bhaumik, "Data augmentation using generative adversarial networks for pneumonia classification in chest xrays," in *Proc. 5th Int. Conf. Image Inf. Process. (ICIIP)*, Nov. 2019, pp. 574–579, doi: 10.1109/ICIIP47207.2019.8985892.

[90] A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman, and P. R. Pinheiro, "CovidGAN: Data augmentation using auxiliary classifier GAN for improved COVID-19 detection," *IEEE Access*, vol. 8, pp. 91916–91923, 2020, doi: 10.1109/ACCESS.2020.2994762.

[91] D. Srivastav, A. Bajpai, and P. Srivastava, "Improved classification for pneumonia detection using transfer learning with GAN based synthetic image augmentation," in *Proc. 11th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence)*, Jan. 2021, pp. 433–437.

[92] C. Han, Y. Kitamura, A. Kudo, A. Ichinose, L. Rundo, Y. Furukawa, K. Umemoto, Y. Li, and H. Nakayama, "Synthesizing diverse lung nodules wherever massively: 3D multi-conditional GAN-based CT image augmentation for object detection," in *Proc. Int. Conf. 3D Vis. (3DV)*. Québec City, QC, Canada: IEEE, Sep. 2019, pp. 729–737, doi: 10.1109/3DV.2019.00085.

[93] D. Jin, Z. Xu, Y. Tang, A. P. Harrison, and D. J. Mollura, "CT-realistic lung nodule simulation from 3D conditional generative adversarial networks for robust lung segmentation," 2018, *arXiv:1806.04051*.

[94] H. Shi, J. Lu, and Q. Zhou, "A novel data augmentation method using style-based GAN for robust pulmonary nodule segmentation," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Aug. 2020, pp. 2486–2491.

[95] M. Nishio, C. Muramatsu, S. Noguchi, H. Nakai, K. Fujimoto, R. Sakamoto, and H. Fujita, "Attribute-guided image generation of three-dimensional computed tomography images of lung nodules using a generative adversarial network," *Comput. Biol. Med.*, vol. 126, Nov. 2020, Art. no. 104032, doi: 10.1016/j.compbiomed.2020.104032.

[96] Y. Onishi, A. Teramoto, M. Tsujimoto, T. Tsukamoto, K. Saito, H. Toyama, K. Imaizumi, and H. Fujita, "Multiplanar analysis for pulmonary nodule classification in CT images using deep convolutional neural network and generative adversarial networks," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 1, pp. 173–178, Jan. 2020, doi: 10.1007/s11548-019-02092-z.

[97] Y. Onishi, A. Teramoto, M. Tsujimoto, T. Tsukamoto, K. Saito, H. Toyama, K. Imaizumi, and H. Fujita, "Investigation of pulmonary nodule classification using multi-scale residual network enhanced with 3DGAN-synthesized volumes," *Radiological Phys. Technol.*, vol. 13, no. 2, pp. 160–169, Jun. 2020, doi: 10.1007/s12194-020-00564-5.

[98] D. Wang, L. Tian, C. Shi, Y.-X. Wei, H. Wang, T.-T. Liu, M. Gong, Y.-W. Zhang, R.-G. Yu, and X.-H. Wu, "Network pharmacology-based prediction of the active ingredients and mechanism of Shen Gui capsule for application to coronary heart disease," *Comput. Biol. Med.*, vol. 122, Jul. 2020, Art. no. 103825, doi: 10.1016/j.compbiomed.2020.103825.

[99] R. Toda, A. Teramoto, M. Kondo, K. Imaizumi, K. Saito, and H. Fujita, "Lung cancer CT image generation from a free-form sketch using style-based pix2pix for data augmentation," *Sci. Rep.*, vol. 12, no. 1, p. 12867, Jul. 2022, doi: 10.1038/s41598-022-16861-5.

[100] X. Wang, Z. Yu, L. Wang, and P. Zheng, "An enhanced priori knowledge GAN for CT images generation of early lung nodules with small-size labelled samples," *Oxidative Med. Cellular Longevity*, vol. 2022, pp. 1–9, Jun. 2022, doi: 10.1155/2022/2129303.

[101] Z. Khan, A. I. Umar, S. H. Shirazi, A. Rasheed, W. Yousaf, M. Assam, I. Hassan, and A. Mohamed, "Lung's segmentation using context-aware regressive conditional GAN," *Appl. Sci.*, vol. 12, no. 12, p. 5768, Jun. 2022, doi: 10.3390/app12125768.

[102] S. Buragadda, K. S. Rani, S. V. Vasantha, and M. K. Chakravarthi, "HCU-GAN: Hybrid cyclic UNET GAN for generating augmented synthetic images of chest X-ray images for multi classification of lung diseases," *Int. J. Eng. Trends Technol.*, vol. 70, no. 2, pp. 229–238, Feb. 2022, doi: 10.14445/22315381/ijett-v70i2p227.

[103] R. Gulakala, B. Markert, and M. Stoffel, "Generative adversarial network based data augmentation for CNN based detection of covid-19," *Sci. Rep.*, vol. 12, no. 1, Nov. 2022, Art. no. 1, doi: 10.1038/s41598-022-23692-x.

[104] J. Paul Cohen, M. Luck, and S. Honari, "Distribution matching losses can hallucinate features in medical image translation," 2018, *arXiv:1805.08841*.

[105] C. Chen, K. Hammernik, C. Ouyang, C. Qin, W. Bai, and D. Rueckert, "Cooperative training and latent space data augmentation for robust medical image segmentation," in *Proc. 24th Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, Strasbourg, France. Berlin, Germany: Springer-Verlag, Sep. 2021, pp. 149–159, doi: 10.1007/978-3-030-87199-4_14.

[106] E. Brion, J. Léger, A. M. Barragán-Montero, N. Meert, J. A. Lee, and B. Macq, "Domain adversarial networks and intensity-based data augmentation for male pelvic organ segmentation in cone beam CT," *Comput. Biol. Med.*, vol. 131, Apr. 2021, Art. no. 104269, doi: 10.1016/j.compbiomed.2021.104269.

[107] M. Loecher, L. E. Perotti, and D. B. Ennis, "Using synthetic data generation to train a cardiac motion tag tracking neural network," *Med. Image Anal.*, vol. 74, Dec. 2021, Art. no. 102223, doi: 10.1016/j.media.2021.102223.

ASWATHY RAVIKUMAR received the B.Tech. and M.Tech. degrees from the University of Kerala. She is currently pursuing the Ph.D. degree with Vellore Institute of Technology, Chennai, India. Since 2020, she has been a Research Associate with Vellore Institute of Technology. She has nine years of teaching experience at graduate and postgraduate levels. Her research interests include machine learning, deep learning, cloud computing, and high-performance computing. She is a member of ISTE and CSI. She was a recipient of the University of Kerala's First Rank Holder for M.Tech. degree, in 2013.

HARINI SRIRAMAN (Member, IEEE) was born in Chennai, India. She received the Bachelor of Engineering degree in computer science from Madras University, in 2003, the Master of Engineering degree in computer science from the College of Engineering, Guindy, in 2009, and the Ph.D. degree in computer science engineering from Vellore Institute of Technology, Chennai, India, in 2018. She is currently an Associate Professor with Vellore Institute of Technology. She has 12 years of teaching experience and industrial experience of one year. She has published more than 20 papers in reputed international journals, conferences, and book series. Her research interests include hardware architectures for accelerated computing, distributed deep learning, and parallel and distributed systems.

CHANDAN CHADHA is currently pursuing the bachelor's degree in computer science with Vellore Institute of Technology (VIT). He has gained industry experience working with various technologies, such as Python, React.js, Node.js, Spring Boot, MySQL, PostgreSQL, MongoDB, HTML5, CSS, Java, and Azure.

VIJAY KUMAR CHATTU is currently a Medical Doctor (M.B.B.S.) specializing in Community Medicine (M.D.), Occupational Medicine (Fellowship), Epidemiology Fellowship (WHO, UCLA), Health Policy (M.Ph.), Global Health Governance (M.Phil.), Global Health Diplomacy (Diploma), International Law (Diploma), and International Relations (Ph.D.) Besides, he holds various Fellowships and Certifications in occupational medicine, sleep medicine, refugee mental health, global mental health, climate change and health, cancer epidemiology, and HIV/AIDS epidemiology. Have vast international experience working in South Asia, Middle East, Caribbean, and Americas. He was with diverse employers, such as the Ministries of Health and Higher Education, WHO, World Bank, UNAIDS-TSF, and INGOs. Listed in World's Top 2% of Scientists in Public Health and Medicine by Stanford University Rankings and Elsevier BV, in 2023, 2022, and 2021 (In both Lifetime and Single year categories). He has ranked no. 5 among the First Ten Best Scientists in Medicine and Health Sciences/Epidemiology and Public Health with the University of Toronto (AD Scientific Index 2024). He has also ranked no. 3 among the Highly cited Public Health Researchers in India (Stanford University and Elsevier Rankings 2023).