VOL. 18, NO. 22, NOVEMBER 2023 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

SEMI-SUPERVISED GAN FOR MEDICAL IMAGE SEGMENTATION

Pallavi Adke[1], Gaurav Adke[2], Shweta Patil[1], Darshana Bhavsar[1] and Aishwarya Mane[1]
Department of Electronics and Telecommunications, PCET's Pimpri Chinchwad College of Engineering and Research, Ravet Michelin Tyres, Pune
E-Mail: pallavi.adke@pccoer.in

ABSTRACT

Echocardiography is a popular ultrasound imaging method used for the diagnosis of heart conditions. With the advent of numerous image processing algorithms, echocardiographic image segmentation has become more significant. This is a crucial stage since it offers a framework for evaluating numerous cardiac parameters, including LV volume and heart wall, valve motion, ejection fraction, thickness, etc. All these factors are crucial for determining a heart's health. The task of manual segmentation requires skilled operators and takes a lot of time. By requiring the discriminator network to output class labels, we extend Generative Adversarial Networks to the semi-supervised type. This paper examines image segmentation techniques for echocardiography to find the borders of the left ventricle. In this paper, we introduce a new convolution neural network model for the auto-segmentation of the left ventricle in echo images. The division of a picture into regions is known as image segmentation. Segments, that computer vision can use to automatically understand. This method makes it easier to simultaneously evaluate and diagnose echo pictures. The segmentation of echocardiographic images can be utilized to measure cardiac characteristics like heart wall thickness.

Keywords: echocardiography, image segmentation, semi-supervised GAN, left ventricle, convolutional neural network.

Manuscript Received 31 May 2023; Revised 11 January 2024; Published 30 January 2024

INTRODUCTION

Using a technique of echocardiographic image segmentation, one can extract numerous different segments drawn from a given image, with each fragment's pixels being identical. These echocardiographic images are obtained by process of echocardiography. Using ultrasonic waves, the technique of echocardiography produces real-time images of the heart's architecture. same features and characteristics. Echocardiographic images are segmented using various techniques to identify the various cardiac structures, such as the right ventricle, left ventricle (LV), right atria, heart walls, left atria, etc. Due to its significance in cardiac diagnostics, LV segmentation is one of the chambers that has been analysed most and is an active topic of research. For LV segmentation, many techniques have been developed. This paper uses a semi-supervised GAN for auto segmentation purpose. In which we can auto-segment echocardiographic images by training the GAN model by supervised as well as unsupervised datasets. Compared to time-consuming subjective manual segmentation, this approach can deliver reliable, trustworthy, and consistent results.

RELATED RESEARCH WORK

Olaf Ranneberger *et al.* [1] build on more sophisticated 'fully convolutional network' architecture ". They change and develop this design to be able to do well with a select few training images and more precise segmentations. The main idea is adding a continuous layer to the regular contract network, where an unsampled operator replaces the task of the pooling operator. Consequently, these layers enhance the output's resolution. Ian J. Goodfellow *et al*. [2] this provides a framework which can provide training algorithms specific to many types of models and the optimization algorithms. This paper examines the unique situation where the discriminative model is also a multi-layer perceptron and the generative model creates samples by running random noise over it. This unique instance is known as an adversarial net. In this situation, only the highly effective backpropagation as well as dropout algorithms can be used to train both these of the models, and only forward propagation can be used to sample from the resulting models. Yuntan *et al*. [3] have employed a framework for adversarial learning for segmentation network semi-supervised training. With consistent outcomes, the model's performance after being trained on only half of the labelled data was comparable to that of the fully supervised trained model. Pallavi Kulkarni, Deepa Madathil [4] discussed nonlinear filters in detail. Section I and Section II discuss diffusion processes and advanced techniques for diffusion processes. Commonly used wavelet denoising methods are discussed in Section III, also various thresholding techniques are discussed in Section IV. Now section V describes filters for fractional arithmetic. Ms. Pallavi Kulkarni and Deepa Madathil [5] discussed the application of image segmentation methods of echocardiography.

Pallavi Kulkarni & Deepa Madathil [6] implemented a new adaptive threshold method for echocardiographic images. A new approach is based on wavelet transformation. On the patient's echocardiographic images, this method's performance is evaluated. Pallavi Kulkarni & Deepa Madathil [7] presents the use of deep learning for fully automated echo LV segmentation. The CNN model used in the current work is developed in relation to the U-NET architecture. This model produces an output image that shows the possibility that each pixel belongs to a specific class. (LV or background). This architecture uses an unsupervised

learning approach, namely AE for feature extraction from echo images. Joyce Henry *et al.* [8] presents a data set consisting of the New York City satellite image and the associated Google Maps page. The translation of satellite images to Google Maps format and vice versa as well as the conversion of Google Maps images to satellite images are all problems with image conversion. Manoj Bodu *et al.* [9] showed how to use a dataset of celebrities to create fake human images. By integrating a clean celebrity dataset into her GAN, we cancreate a model that can generate realistic portraits of people that do not exist in the real world.

MP Vaishnave *et al*. [10] describe some gaps in the literature to make it possible for the scientific community to create new data-driven algorithms. Last but not least, this publication offers cutting-edge quantitative measurements including precision, accuracy, F1 score and recall. To evaluate scene classification in satellite imagery. Ying Zhang *et al*. [11] these deduced semantics help the translation to do a region-like task, thus reducing a lot of per-pixel noise. It is simple to include the suggested semantic estimation and GPS integration into a variety of backbone GAN designs. P. Devabalan [12] provides remote sensing data poses the challenge of ideas about how to quickly organize and analyze the data. Phillip Isola [13] the results presentedin this article demonstrate that conditional adversarial networks, especially for frame-to-frame translation problems with extremely organized graphic output, are a promising method. Recommends that such networks can be deployed in a wide variety of environments as they learn loss tailored to the task and data at hand. Kumarapu Laxman [14] this work suggested a high- resolution frame-to-frame transformation method based on a multi-scale gradient-based U-Net (MSG U-Net). The proposed model shows better training by using gradients at different scales.

Swetava Ganguli [15] presents a workable model for creating maps (his GAN without an encoder, with reconstruction, and transmission loss in style). Since there were no hyper-parametric search and only 15 epochs of training, the results indicate that extended training times would yield more realistic results.

From a widely used different segmentation method, some require contour initialization to start recognizing boundaries. Depending on the user-specified starting circumstances, this may result in varying results. The main manual demarcation used by professionals for comparison in segmentation methods can result in discrepancies between observers. The robustness and reliability of segmentation methods in echocardiography depend on the combined use of the image's physical information, prior anatomy education, time restraints on frame fixation, and strength and shape constraints. It can be improved.

MATERIAL AND METHOD

This section outlines the implementation of the proposed method. Figure-1 illustrates the technique for LV segmentation of the heart from echocardiographic images using a semi-supervised GAN Echocardiographic images are subjected to pre-processing operations before semi-supervised GAN training Using the semi-supervised learning strategy, we train our model using an insignificant sample set from the labelled data from a vast subset of unlabelled data to segment left ventricle of heart from echocardiographic images.

The step-by-step process is explained in Figure-1.

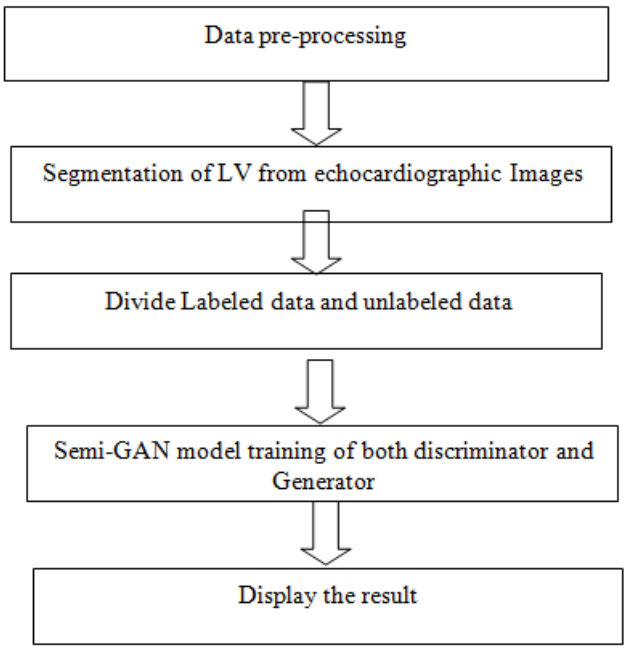

Figure-1. Flowchart of image segmentation usingsemi-supervised GAN.

Data Pre-Processing

Echocardiographic images are subjected to pre-processing operations before semi-supervised GAN training.

a) **Echocardiographic Images:** Echocardiography is a technique that creates internal images of the heart using sound waves. These images are known as echocardiographic images, which show four chambers of the heart i.e. right ventricle, left ventricle, right atria, left atria.

b) **Denoising of Images:** Image denoising's primary goal is to estimate the original image by removing noise from a version of the image that has been tainted by noise. One of the fundamental issues in the field of image processing is this one. Speckle noise has an impact on ultrasound images used in healthcare. Speckle, which can degrade the clarity of ultrasound images, is generally a granular noise. Because of the surrounding noise, the impact of tissues in the surroundings, the roughness of the tissues being imaged, the individual's movement of breathing

cycles, and other factors, speckles might appear during ultrasonic imaging. Multiple methods have been developed to decrease speckle noise in the image as many studies have shown that the presence of speckle significantly lowers the accuracy and reliability of automatic image processing algorithms and their interpretations. During the image pre-processing stage, the described speckle noise problem can be resolved.

Segmentation of Left Ventricle (LV) from Echo Images using Semi-Supervised GAN

Using the semi-supervised learning strategy, we train our model using a small set of labelled data from a vast set of unlabelled data to segment the left ventricle of the heart from echocardiographic images. Figure-2 shows block diagram of semi supervised GAN which includes the Generator and Discriminator with inputs of unlabelled and labelled images. As shown in Figure-2.

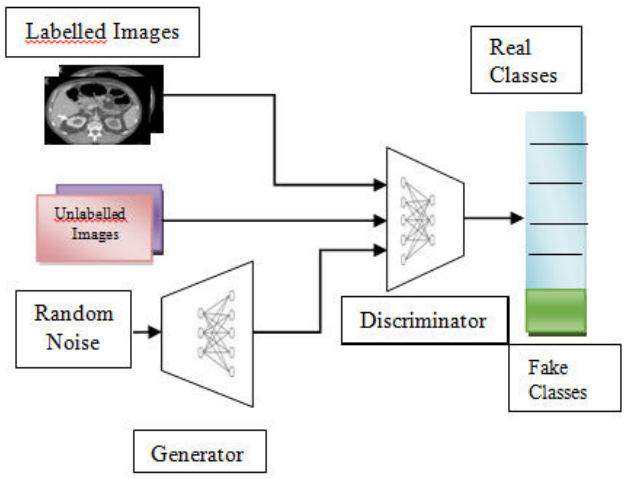

Figure-2. Block diagram of semi-supervised.

a) **Datasets Used:** Two Datasets are as follows:

i. **Supervised Discriminator:** A small partof the entire training set will be used as the dataset which will be labelled dataset.

ii. **Unsupervised Discriminator:** This data will be a complete training set i.e., unlabelled data set.

b) **Training semi-supervised:** Begin the training procedure by directing the supervised discriminator on a set of restarts through instructing the trainees, the supervised discriminator on a set of real labelled images, followed by the unsupervised discriminator on real unlabelled photos (labelled as "1") and fake images (labelled as "0") produced by t h e generator.

The generator is trained as follows:

A number of fake images are produced by the generator. These created images are sent to adiscriminator for classification, and the discriminator is given the goal label of "1" for real. Therefore, the loss would be very large if the photos did not look realistic,so our goal is to generate realistic-looking images to reduce the loss. As a result, as the generator is trained on more images, it begins to produce better images, and the discriminator does as well because it does not want its loss to increase while differentiating real from fake images. As a result, the discriminator and generator are playing a game against one another because as one grows better, the other must also improve.

c) **Generator Network:** Our generator model is built using the Keras Sequential API Given the simplicity of our goal; it is a simple generator model.

d) **Discriminator Network:** Discriminator that receives an image as input and produces '1' or '0' depending on if the image is real or not. These are values that can also be negative because the output in this case is from a dense layer that has not undergone any activation. Three different types of images are used to train the discriminator: labelled training images, unlabelled training images, and fake images which are produced by the generator. Its responsibility includes categorizing the labelled Training Images into the appropriate classes in addition to identifying real from Fake Images.

The Discriminator produces two distinct results:

- Softmax activation is used to sort labelled data into the appropriate classes, hence it is referred as a supervised discriminator, as shown in equation (1).

$$\sigma(Z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad (1)$$

Z = input vector
σ = softmax
e^n = input vector's standard exponential
K = multi-class classifier's number of classes
e^{zj} = Output vector's standard exponential

- Custom Activation to determine whether something is real or fake. Formula for it is given in equation (2).

$$\frac{\sum_i \exp(z_i)}{\sum_i \exp(z_i)+1} \quad (2)$$

Here, k is 10 in our application and y ranges from 0 to 1, and z is the output of the dense layer without any activation.

The primary advantage of the discriminator i s that it has been trained on both labelled and vast amounts

of unlabelled data because it must distinguish between real and fake images. To achieve this goal, the discriminator must learn to extract features for identifying whether an image is real or fake. By learning to distinguish between real and fake images, which it would not ordinarily be able to do on the basis of a short sample of labelled data, the discriminator will be better able to categorize the labelled data. The Implementation of the discriminator is shown in Figure-3 as follows.

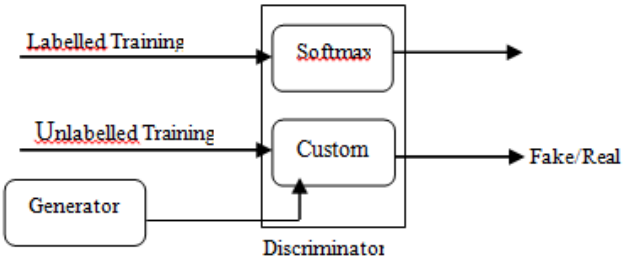

Figure-3. Block diagram of discriminator.

a) **Supervised Discriminator:** Uses softmax to divide our input image into one of the classes and accepts the discriminator model as input. The activation function essentially adds over all classes and scales output between [0,1] to determine the probability of the image being real or false because the chance of the image having real, or fake is the sum of probabilities over all classes.

The difference between the supervised and unsupervised discriminators' is their output nodes and their use of weights will be same.

b) **Unsupervised Discriminator:** This will use the supervised discriminator as input which we have previously developed, and it will apply a custom activation to the discriminator's output. Since we had kept the discriminator's output as the outcome of a dense layer, which provides values without any use of an activation function. Thus, the custom activation performsas given in equation (3).

$$v = \frac{\sum_i \exp(z_i)}{\sum_i \exp(z_i)+1} \quad (3)$$

Here, k is 10 in our application and y ranges from 0 to 1,
and z is the output of the dense layer without the use of any activation.

The activation function essentially adds through all classes and scales output between [0, 1] to determine the probability of the image getting real or fake because the chance of the image getting real or fake will correspond to the sum of probabilities through all classes. The use of a sigmoid was far less successful than his method.

The only difference between the unsupervised and the supervised discriminators is their output nodes but their use of weights is as it is.

Building GAN

Integrates the input from the discriminator and generator to create a model. The generator network is now being trained by doing this. We'll refer this as the composite model. As GAN is only used for unlabelled images, the discriminator is indeed the unsupervised discriminator. This model is used to segment the images as shown in Figure-4.

Figure-4. Image segmentation.

DATASET

In this research, we have used two types of datasets one is CAMUS dataset, and another is manually created by us.

Manual Dataset

The video recordings of 82 patients' cardiac cycles in DICOM format (57 males and 25 women) served as the dataset for our investigation. Underneath the supervision of Dr. D. Ramesh, Department of Cardiology head, this information was gathered from the Whitefield campus of Bangalore, India's Vydehi Institute of Medical Sciences and Research Centre. The iE33 Ultrasound System from Philips is the device used to collect data. Cardiac cycle recordings are gathered for both the two chamber and four chamber viewpoints of each subject. There are 40-45 picture frames in a recording of a cardiac cycle. Our goal was to choose pictures from each cardiac cycle at the systole and diastole stages. At the observable end states any four to five frames of the cardiac cycle are manually selected. In order to prevent missing specific frames at the systole and diastole stages, this is done. Out of the photographs that are accessible, 5% are used as the test set and 5% as the validation set. Data augmentation with random brightness shifts and horizontal and vertical shifts is thought to have a greater generalizing effect. The size of the original image is 1,024 x 768 pixels. These photos were cropped for training purposes to a size of 658 x 658 pixels in order to the exclude extraneous text information as well as solely include ultrasound output. When making predictions, similar cropping is done. These are the some echocardiographic images of manual dataset as shown inFigure-5.

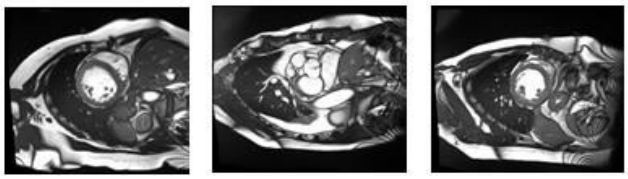

Figure-5. Raw input images.

ACDC Dataset

The University Hospital of Dijon collected actual clinical evaluations for the creation of the ACDC dataset. The collected data were completely anonymized and managed by the guidelines established by the hospital's regional ethical council in Dijon, France. Our dataset includes enough examples for both (1) a thorough training of machine learning algorithms and (2) a clear evaluation of the fluctuations in the key physiological parameters collected from cine-MRI (particularly diastolic volume and ejection fraction). The dataset consists of 150 exams from various patients, dispersed into 5 equally distributed subgroups (4 pathological and 1 group of healthy subjects) [17].

150 patients make up the study's target population, which is divided into the following 5 subgroups:

a) Normal Subjects 30- NOR.

b) 30 individuals had dilated cardiomyopathy (left ventricular ejection fraction less than 40% and diastolic left ventricular volume greater than 100 mL/m2). - DCM.

c) 30 patients had a history of myocardial infarction, with a left ventricular ejection fraction of less than 40% and many abnormally contracting myocardial segments. - MINF

d) 30 patients with abnormal right ventricle (volume of the right ventricular cavity higher than 110 mL/m^2 or ejection fraction of the right ventricle lower than 40%) -RV

e) With a left ventricular cardiac mass greater than 110 g/m2, numerous myocardial segments thicker than 15 mm in diastole, and a normal ejection fraction, 30 people have hypertrophic cardiomyopathy (HCM).

Physiological parameters such as left or right diastolic volume or ejection fraction, LV mass, local LV contraction and maximum myocardium thickness were used to clearly characterize each group.

RESULT ANALYSIS

The performance of the model can be evaluated by using the following parameters:

a) SNR: The root of the photon's quantity in the image's region of brightest is used to calculate the signal-to-noise ratio (S/N) of a digital microscopic image. The ideal range is from 20 to 40.

b) PSNR: This ratio is often used to compare the original and compressed images' quality. The ideal range is 30–50db.

c) SSIM: SSIM calculates how well movies and images are viewed. Compares two images-the original image and the rebuilt image-to determine how similar they are.

d) Jaccard index: A statistic used to assess the similarities and diversity of samples is the Jaccard index. It is commonly referred to as the Jaccard similarity coefficient.

These parameters are calculated by using these segmented and raw images as shown in Figure-6.

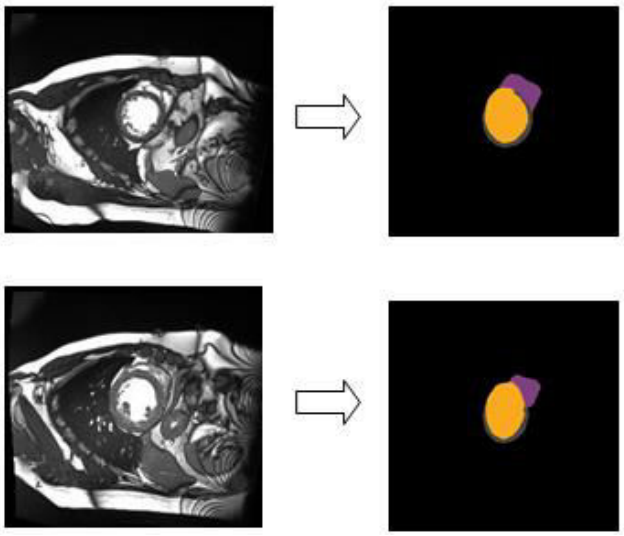

Figure-6. Input - Output images.

The analysis of the segmented Images from the manual dataset is given in the following Table-1:

Table-1. Result analysis

Images	Jaccard's Index	PSNR	SNR
Image Pair-1	0.565613	40.14	19.64
Image Pair-2	0.79633	39.63	19.06
Image Pair-3	0.58369	40.03	17.72
Image Pair-4	0.48795	40.82	18.34
Image Pair-5	0.09	39.70	18.41

The validation set is solely used to assess the model's performance; the training set is utilized to train the model.

The training accuracy vs validation accuracy graph is shown in the following Figure-7.

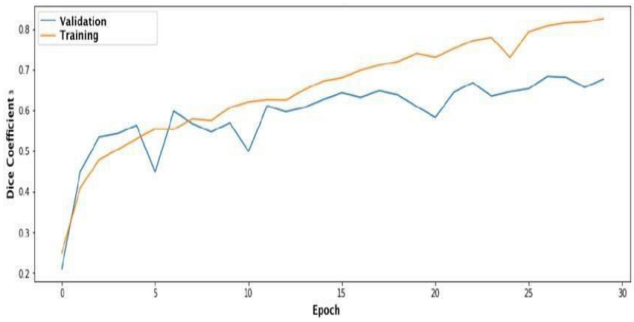

Figure-7. Training vs Validation accuracy graph.

Table-2. Output analysis.

Epoch	DisLoss	UnlabGenLoss	LabGenloss
0	9.63E-01	3.04E+00	5.54E+00
1	2.53E-01	2.25E+00	3.52E+00
2	3.48e-01	1.96e+00	1.42e+00
.....			
258	3.82e-01	1.79e+00	3.86e-01
259	4.55E-01	1.80E+00	3.62E-01
260	2.88E-01	1.97E+00	3.96E-01

The graph of this table is given below:

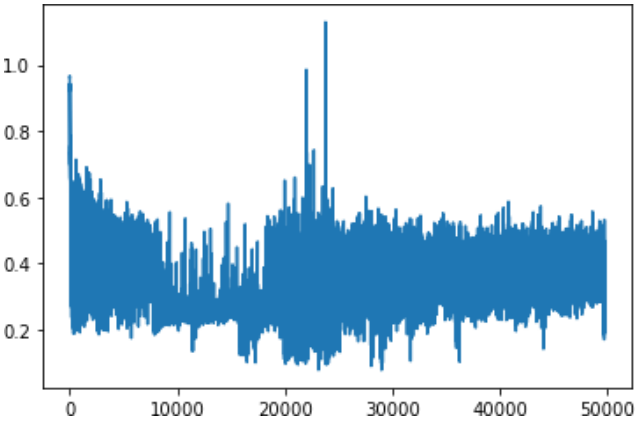

Figure-8. DisLoss graph.

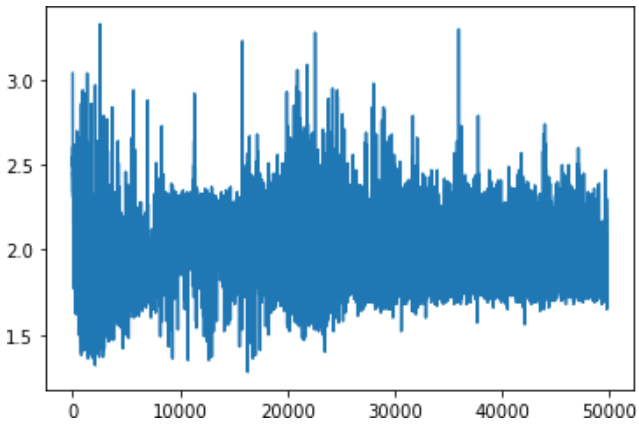

Figure-9. UnlabGenLoss graph.

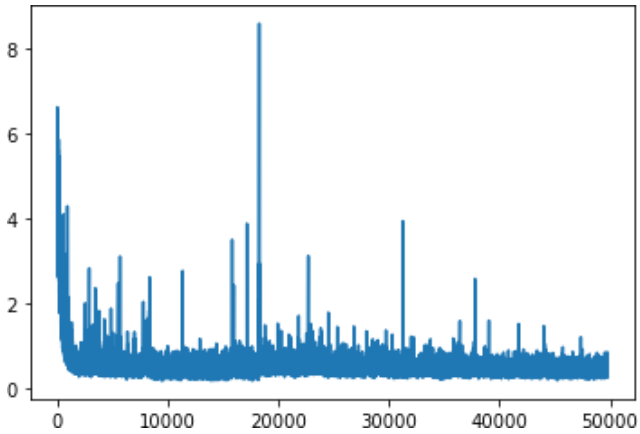

Figure-10. Lab gen loss graph.

Table-3. Intersection over Union.

S. No	IOU
1	0.351921857
2	0.405906651
.....	
4	0.706551391
5	0.696794532
6	0.613656524
7	0.571581892
8	0.710584224
9	0.60566298

Graph based on this intersect over union is as given below:

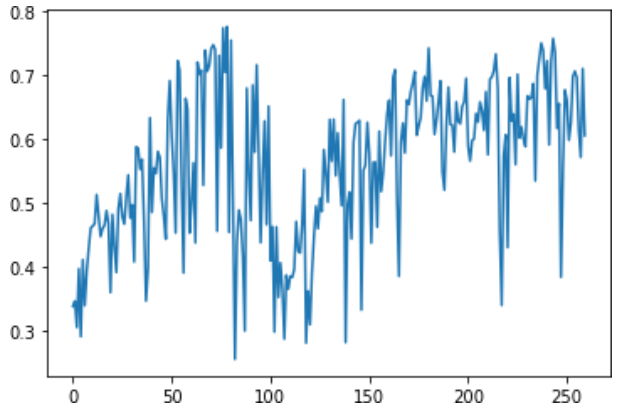

Figure-11. IOU graph.

ADVANTAGES

The advantages of this proposed system can be summarized below.

a) GANs go into the details of the data and can be easily interpreted into different versions. This is useful when working with machine learning.

b) It can improve model generalization and performance.

c) It can quickly identify trees, roads, bikers, people, parked automobiles, and even compute the distance between various objects using GAN and machine learning.

GANs generate data similar to the originals you feed a GAN an image; it creates a new version of the image that resembles the original image. Similarly, different versions of text, video, and audio can be generated.

CONCLUSIONS

The use of image segmentation methods within the field of echocardiography has been discussed. It introduces several significant segmentation research. The existence of noise, the lack of boundaries, and the inability to clearly distinguish between picture regions limit the applicability of the intensity gradient technique to echocardiography, even though it is numerically less expensive. The segmentation of echocardiographic images using novel techniques such as active shape models, active appearance models, machine learning, and fuzzy method techniques are gaining more and more attention from researchers. The input image is noise-free because to the suggested method's use of less raucous approximation wavelet coefficients for segmentation, which is another benefit. With various performance characteristics, the suggested technique is contrasted with FCN, DRSLE, ACE, and extended UNET. The suggested method, like existing level set methods and active contour, depends on user input to locate the left ventricle (LV) on echocardiographic pictures. In terms of distance-based and similarity-based performance parameters, it is seen that the suggested method outperforms existing non-deep learning approaches to echocardiographic image segmentation.

FUTURE SCOPE

Automated echo interpretation has the power to revolutionize clinical practice by delivering a quick, affordable, portable, and precise evaluation of heart anatomy and function. The end-to-end system for reliable real-time echo search, segmentation, and quantification was proposed in this work. The suggested approach, which has been tested on echo datasets, collects high-quality echoes from a single view, separates heart chambers with intricate overlaps, and outputs a highly correlated cardiac image with a professional evaluation. The system extracted measures and indications with great performance. The computational efficiency of the suggested system is also significantly higher than that of the prior art. Our framework's excellent effectiveness and performance make it easy to use and produce in real-world bedside echo and point-of-care ultrasound applications.

ACKNOWLEDGMENT

The research work is done under Core Research Grant of Science and Engineering Research Board File Number-CRG/2022/004046.

REFERENCES

[1] O. Ronneberger, P. Fischer and T. Brox. 2015. U-Net: Convolutional networks for biomedical image segmentation. In Proc. Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham. pp. 234-241.

[2] Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley *et al*. 2014. Generative adversarial nets. Advances in Neural Information Processing Systems. 27: 2672-2680.

[3] Yun Tan, Ling Tan *et al*. 2022. Semi - supervised medical image segmentation based on generative adversarial networks. Journal of New Media Tech Science Press. 4: 155-164.

[4] Pallavi Kulkarni, Deepa Madathil. 2018. A review on echocardiographic image speckle reduction filters. An International Journal of Medical Sciences. 29(12): 2582-2589.

[5] Pallavi Kulkarni, Deepa Madathil. 2018. A review of echocardiographic image segmentation techniques for left ventricular study. ARPN Journal of Engineering and Applied Sciences. 13: 3536-3541.

[6] Pallavi Kulkarni, Deepa Madathil. 2022. Adaptive Thresholding Method for Speckle Reduction of Echocardiographic Images. IETE Journal of Research. 68(2): 1034-1042.

[7] Pallavi Kulkarni, Deepa Madathil. 2022. Fully automatic segmentation of LV from echocardiography images and calculation of ejection fraction using deep learning. Int. J. Biomedical Engineering and Technology. 40(3): 241-261.

[8] Joyce Henry, Terry Natalie *et al.* 2021. Pix2Pix GAN for Image-to-Image Translation. Research Gate Publication. pp. 1-5.

[9] Boddu Manoj, Boda Bhagya Rishiroop. 2020. Image to Image Translation Using Generative Adversarial Network. International Journal of Scientific & Technology Research. pp. 105-108.

[10] M. P. Vaishanave, K. Suganya Devi, P. Srinivasan. 2019. A Study on Deep Learning Models for Satellite Imagery. International Journal of Applied Engineering Research ISSN 0973-4562, 14: 881- 887.

[11] Ying Zhang, Yifang Yin *et al*. 2020. An Enhanced GAN Model for Automatic Satellite-to-Map Image Conversion. IEEE Access, 8: 176704-176716.

[12] P. Devabalan. 2014. Satellite Image Processing on A Grid Based Computing Environment. International Journal of Computer Science and Mobile Computing. 3(3): 1039-1044

[13] Phillip Isola, Jun-Yan Zhu *et al*. 2017. Image-to-Image Translation with Conditional Adversarial Networks. IEEE Xplore. pp. 1125-1134.

[14] Kumarapu Laxman. 2021. Efficient High-Resolution Image-to-Image Translation using Multi-Scale Gradient U-Net. IEEE/CVF Conference on Computer VisionImage to Image translation.

[15] Swetava Ganguli. GeoGAN: A Conditional GAN with Reconstruction and Style Loss to Generate Standard Layer of Maps from Satellite Images. arXiv: 1902.05611.

[16] Pallavi Kulkarni, Deepa Madathil. 2021. Echocardiography Image Segmentation Using Semiautomatic Numerical Optimization Method Based On Wavelet Decomposition Thresholding.

[17] O. Bernard, A. Lalande, C. Zotti, F. Cervenansky, *et al.* 2018. Deep Learning Techniques for Automatic MRI Cardiac Multi-structures Segmentation and Diagnosis: Is the Problem Solved? in IEEE Transactions on Medical Imaging. 37(11): 2514-2525.