Self-Supervised Augmentation of Quality Data Based on Classification-Reinforced GAN

Seunghwan Kim and Sukhan Lee

November 30, 2022

Self-Supervised Augmentation of Quality Data Based on Classification-Reinforced GAN

SeungHwan Kim
School of Artificial Intelligence
Sungkyunkwan University
Korea, Suwon
shkim218@skku.edu

Sukhan Lee*
School of Artificial Intelligence
Sungkyunkwan University
Korea,Suwon
lsh1@skku.edu

Abstract— **In deep learning, the quality of ground truth training data is crucial for the resulting performance. However, depending on applications, collecting a sufficient amount of quality data from a realistic setting is problematic. In this case, data augmentation can play an important role as long as augmentation ensures data quality and diversity for training, preferably in an unsupervised way. Recently, a number of GAN variants have been emerged for improved quality in data augmentation. Although successful, further improvement is necessary for enhancing diversity in addition to quality in data augmentation. In this paper, we propose a GAN-based approach to self-supervised augmentation of quality data based on Classification-Reinforced GAN referred to here as CLS-R GAN, to extending diversity as well as quality in data augmentation. In CLS-R GAN, a discriminator-independent classifier additionally self-trains the generator by classifying the fake data, as well as augmenting the real data in an unsupervised way. Extensive experiments were conducted, including an application to augmenting liver ultrasonic image data, to verify the effectiveness of CLS-R GAN based on standard evaluation metrics. The results indicate the effectiveness of CLS-R GAN for improved quality and diversity in augmented data.**

Keywords*— *GAN, Unsupervised Data Augmentation, Generator, Self-Training

I. INTRODUCTION

GAN frameworks provide an effective means for self-supervised data augmentation with the fake data generated by their generators. To date, several GAN variants have been proposed for the sake of maximization of generated fake data's qualitie and diversities. To this end, many GAN variants incorporate such data attributes as clusters and classes as well as self-supervised data augmentation into the adversarial framework, often, with a classifier attached to the discriminator with shared weights. For example, ClusterGAN [1] and its derivatives [2,3,4,5] exploit the clustering structure embedded in fake data to enhance the quality in data generation with additional training of the generator based on an attached encoder. ACGAN [6], InfoGAN [8] and their derivatives [7,9,10] exploit the class structure and data attributes, instead, by using a classifier attached to the discriminator with shared weights. They achieve quality improvement in data generation by imposing the generator to follow class and attribute configurations embedded in the real and fake data. On the other hand, DAGAN [11], SSGAN [12] and their derivatives [13,14,15] introduce an augmentation of real and fake data in GAN frameworks by applying self- supervised data transformation, especially, for emphasizing data diversity. Note that, by "quality," we mean that the generated data are as real as the real data, while, by "diversity," we mean that the generated data genuinely cover the distribution of real data without bias. The state-of-the-art approaches introduced above have contributed, on their own sake, to the advancement of GAN-based data generation in terms of data quality as well as diversity. However, they are also subject to certain performance limitations due to the issues associated with goal inconsistency [12] and constraint in data augmentation [13]. A necessity to break through the current performance limit exists by overcoming the above issues. Goal inconsistency takes place in GAN frameworks due to additional learning of class, cluster and attribute distributions with a classifier shared its weights with the discriminator, which disturbs the adversarial framework to converge to real data distribution. To mitigate the issue of goal inconsistency, attempts [21,26] were made, with a limited success, to have adversarial competition applied to both data and class/attribute distributions at some expense of stability in convergence. Nonetheless, goal inconsistency remains as an issue to overcome for further advancement in GAN-based data generation. On the other hand, constraint in data augmentation takes place as the portfolio of data transformation should be restricted to avoid the augmented data altering the distribution of real data [13]. Such constraint in data augmentation serves as a limiting factor for self-supervised augmentation of real and fake data to improve diversity in GAN-based data generation. All in all, achieving both high quality and high diversity in GAN-based data generation remains as a challenge.

In this paper, we propose an approach to GAN-based data augmentation that offers both high quality and high diversity in data generation and clustering. The proposed GAN framework is referred to here as "CLS-R GAN." In CLS-R GAN, an attached yet discriminator-independent classifier corrects the fake data qualified for self-training the generator as well as for self-augmenting the real data. An overview of the proposed approach is presented in Section III.

*Corresponding author : Sukhan Lee(lsh1@skku.edu)
This work was supported, in part, by "5G Edge Brain Based Intelligent Manufacturing" project IITP-2022-0-00067, in part, by AI Graduate School Program, Grant No.2019-0-00421, and by ICT Consilience Program, IITP-2020-0-01821, sponsored by the Korean Ministry of Science and Information Technology (MSIT), and, in part, by "Deep Cross-Sensory Transfer for Visually Impaired" project, 2020R1A2C200956811, of NRF sponsored by Korea Ministry of Science and ICT.

II. RELATED WORK

GAN [23] was introduced by pairing a discriminator and a generator competing adversarially, in which the generator is trained to generate fake data by closely following the distribution of real data while the discriminator is trained to refine the distribution of real data by discriminating fake data from real data. Such capability of GAN in adversarial generation of fake data makes GAN a natural candidate for self-supervsed augmentation of real data. However, further improvement of original GAN has been needed for generation of quality fake data in data augmentation, apart from neccesity to solve the instability or mode collapse problem in training. Many improved variants of GAN have emerged to deal such need to address those problems. As the first case, CGAN [16] was proposed to generate fake data in target categories on the basis of specific conditional vectors. Meanwhile, ACGAN [6] is its improved version, allowing the discriminator to check that generated fake data are in target classes on the basis of attached clasifier with shared-weights. In the variant of InfoGAN [8], the discriminator checks the classes of generated fake data applying mechanism of information maximization based on class attributes. Besides, WGAN [17,18] was also proposed with the aim of stability enhancement in the original GAN, thorugh application of gradient clipping and penalty for efficient solution in the mode collapse problem.

For generation of quality data, apart from structuring fake data with class labels, also clustering can also be directly applied to fake data generation [20]. In this respect, ClusterGAN [1] was the first where generated fake-data were clustered by attaching an encoder, and where cyclic loss associated with clustering consistency was minimized. To improve clustering functions, SEMI-ClusterGAN [19] applies information maximization to clustering. Recently, an extensive incorporation of clustering has been proposed, with top-priority of diversity improvement for the generated fake data [22]. And for this, feature embedding clusters have been used in CGAN to augment real-data used for generating GAN-based fake data.

As far as self-supervised data augmentation is concerned, traditionally, a predefined set of data modification, such as image rotation, flipping, cropping, as well as corruption with noise, has been used. On the other hand, GAN and its variants can generate fake data which will be used for data augmentation. For instance, in the recently proposed SSGAN and DAGAN, we have shown remarkable results using methods, according to which generator and discriminator can be learnt by rotating or augmenting images, stabilizing the learning of GAN. According to [21], there have been studies showing that the GAN can be better learnt when using generated data from GAN learning, compared to when using only real data. And in [15], upgrading generator can be performed by augmenting both real and fake data. In the study of [13], it shows that using data augmentation excluding rotation and flip while training GAN can affect the generator negatively in learning real data distribution.

III. CLS-R GAN: AN OVERVIEW

The proposed CLS-R GAN consists of the following three phases: 1) The classifier pre-training phase, the attached classifier is pre-trained with the ground truth data independently of the discriminator training. 2) The adversarial training phase, the generator and discriminator play a min-max game so as to generate fake data in line with real data distribution. 3) The generator self-training phase , further tranining is made with the qualified fake data, which were chosen by the attached independent classifier. Fig. 1 illustrates the architecture of the proposed CLS-R GAN, where phase (1) and (3) are represented.

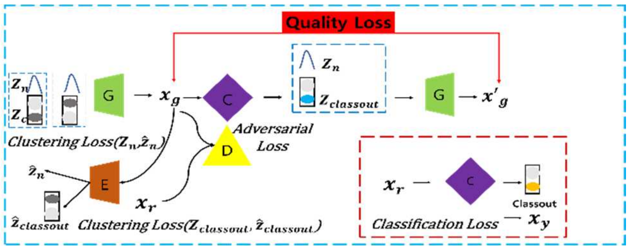


Fig. 1 Overview of Architecture of CLS-R GAN.

The proposed generator self-training is done in conjunction with adversarial training by minimizing additional quality loss, by which the generator can generate fake data more in consistent with their classes. In addition, the pre-trained classifier is used for selecting those fake data qualified for data augmentation. Next, we present more details on the proposed generator self-training phase.

IV. GENERATOR SELF-TRAINING PHASE

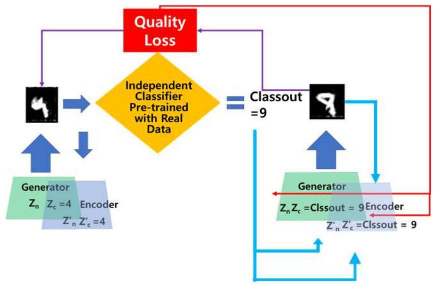
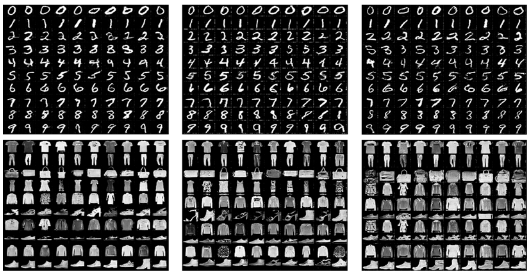


Fig. 2 Flows associated with the clustering and self-improvement loop (respectively marked in blue and red/purple) in the generator self-training phase.

Fig. 2 illustrates the proposed self-training process which consists of clustering and self-improvement loops (respectively marked in blue and red/purple).. And we found that it provides extra gradients in weight training for both generator and encoder, obviously more than original adversarial training. As such, it turned out to serve better for generated fake-data's quality improvement. In the self-improvement loop, the fake data were classified by the pre-trained classifier so that the generator would be trained based on the generated fake data with the class label provided by the classifier as ground truth data. In the clustering loop, the generator and encoder are trained in such a way that the generated fake data are well-clustered by class-label assigned to generator input.

Fig. 3 illustrates typical examples of fake data generated by the generator after the generator self-training is completed. Notice the high quality of generated fake data in terms of clustering and classification properties.



(a) (b) (c)

Fig. 3 Typical examples of fake data generated by the generator after the generator self-training is completed. (a) without bias MNIST(top) and Fashion-MNIST(bottom) (b) with size bias MNIST(top) and Fashion-MNIST(bottom) (c) with contents bias MNIST(top) and Fashion-MNIST(bottom).

Fig. 4 illustrates some of the highly deformed instances of qualified fake data associated with MNIST(a) and Fashion-MNIST(b) data that are selected by the independent classifier. Fig. 4 indicates that the generator self-training phase is especially effective for expanding the diversity in data augmentation.

(a) (b)

Fig. 4 Highly deformed examples of qualified fake data for MNIST(a) and Fashion-MNIST(b) that are selected by the independent classifier.

For data augmentation, we proposed a criteria for selecting quality fake data eligible for augmentation, for which we used the consistency of class labels between the generator input and the classifier output as well as the classification probability as the selection criteria. Note that, by adjusting the classification probability in the selection criteria, we can control the quality and diversity in data augmentation.

V. EXPERIMENTS

A. Evaluation metrics

To evaluate the performance of the proposed CLS-R GAN, we adopted the same standard metrics for quality assessment as used by other approaches under comparison. They include NMI, ARI, IS and FID. Besides, we also evaluated the generated fake data based on the accuracy and diversity measure, especially, for the analysis of the generated fake data in terms of diversity. Here, we define "Diversity" metric (Div.) in terms of the number of different species in individual classes composed of the fake dataset used in augmentation, similar to the way diversity is defined in biology. We determine the number of different species in a class based on within-class clustering of fake data by assigning inter-distances between fake data with the feature similarity index measure (FSIM). Then, diversity is measured by applying Gini-Simpson index to the number of species in classes. Note that NMI and ARI metrics are for evaluating the quality of clustering based on externally provided class labels such that the higher the value, the better the clustering quality is. FID captures the similarity of the generated fake data to the real data such that the lower the value, the better the quality of generated fake data is, while IS score becomes higher when the generated fake data are with higher probability in classification yet with more equally distributed class labels. On the other hand, the proposed diversity of measure is intended to indicate the variety of different species within individual classes such that the higher the diversity of measure, the higher the variety is. Notice that we introduced an, externally provided, high accuracy classifier solely for the evaluation of accuracy in class labels. For the high accuracy classifier, we adopted an architecture specialized for the classification of MNIST and Fashion-MNIST datasets that we used in experiments. To make the high accuracy classifier effective for classifying diversified fake data, we augmented the original 60K MNIST and 60K Fashion-MNIST training data to 96K for each by adding 36K MNIST and 36K Fashion-MNIST fake data that are manually annotated. The trained high accuracy classifier shows the accuracies of 99% and 97% for MNIST and Fashion-MNIST datasets, respectively, when tested with the original 10K MNIST and 10K Fashion-MNIST testing data plus 9K MNIST and 9K Fashion-MNIST testing data randomly selected from the manually annotated fake data.

B. Experimental setup

For comparative performance evaluation of the proposed CLS-R GAN, we used MNIST and Fashion-MNIST datasets for training and testing. For fair comparison, we trained and tested the CLS-R GAN and other GAN variants under comparison based on the same datasets. Also, in order to investigate the effect of data bias on performance, we prepared for three types of training datasets: 1) The standard MNIST and Fashion-MNIST datasets provided originally by publicly available databases . 2) The MNIST and Fashion-MNIST datasets with size bias, in which, for a chosen class, only a small number of class data are randomly selected and repeated to implement the data bias in size. 3) The MNIST and Fashion-MNIST datasets with content bias, in which, for a chosen class, only a single species of data in terms of feature similarity is selected and repeated to implement the data bias in content. More specifically, the standard MNIST and Fashion-MNIST datasets consist of 60K training and 10K testing data for each that amount to about 6K training and 1K testing data assigned to 10 classes. In the case of size biased MNIST and Fashion-MNIST datasets, we randomly selected a small number of data, say 20 data, from the class 5 datasets and repeated the randomly selected data to make up the original number. On the other hand, in the case of content biased MNIST and Fashion-MNIST datasets, we randomly selected a data from the class 5 and collected those data in the class 5 that are similar to the selected data with Structure Similarity Index Measure (SSIM) greater than 0.7, while removing out all the rest. Then, we repeated the collected similar data to make up the original number.

C. Evaluation of CLS-R GAN

We evaluated the performance of CLS-R GAN such that we can assess the generator self-training phase. The evaluation is done in comparison with the state-of-the-art GAN variants, including ACGAN, ClusterGAN. The reason why we have selected these variants for ACGAN and ClusterGAN is that ACGAN is well known as generate diverse image and ClusterGAN is well known as cluster the fake image well, respectively. The metrics used in evaluation include NMI, ARI, FID, IS, Accuracy (ACC.) and Diversity (Div.), as described in Section V.A. For fair comparison, we trained and tested the above GAN variants under comparison based on the same respective training and testing datasets.

Tables 1 and 2 show the performance evaluated based on the without bias MNIST and Fashion-MNIST datasets, respectively, whereas Tables 3 and 4 show the performance evaluated based on the respective MNIST and Fashion-MNIST datasets modified with size and content biases.

Table 1. Performance of CLS-R GAN compared with the state-of-the-art GAN variants: MNIST dataset without data bias.

Metrics	NMI	ARI	FID	IS	Acc.	Div.
ACGAN	0.839	0.727	148.0	1.86	92.05	0.544
ClusterGAN	0.916	0.865	94.64	1.85	0.10	0.535
CLS-R GAN	**0.980**	**0.980**	**60.46**	**1.98**	**99.84**	**0.720**

Table 2. Performance of CLS-R GAN compared with the state-of-the-art GAN variants: Fashion-MNIST dataset without data bias.

Metrics	NMI	ARI	FID	IS	Acc.	Div.
ACGAN	0.736	0.684	161.31	2.10	80.47	0.618
ClusterGAN	0.625	0.474	471.77	2.29	11.59	0.661
CLS-R GAN	**0.966**	**0.98**	**119.89**	**2.73**	**99.96**	**0.79**

Tables 1 and 2 indicate that the proposed CLS-R GAN provides superior performance to the state-of-the-art GAN variants. Especially, the superiority in performance becomes more evident when the augmentation criteria, $\mathbf{z_c}$==$\mathbf{z_{classout}}$, is incorporated into data augmentation.

Table 3. Performance of CLS-R GAN compared with the state-of-the-art GAN variants: MNIST dataset with data bias(size/content).

Metrics	NMI	ARI	FID	IS	Acc.	Div.
ACGAN	0.765 /0.768	0.73 /0.82	124.54 /143.96	1.76 /1.97	87.67 /82.66	0.57 /0.51
ClusterGAN	0.883 /0.854	0.84 /0.81	111.41 /133.24	1.78 /1.80	9.40 /0.59	0.59 /0.51
CLS-R GAN	**0.988 /0.976**	**0.98 /0.97**	**103.95 /105.9**	**1.98 /1.95**	**99.75 /100**	**0.70 /0.69**

Table 4. Performance of CLS-R GAN compared with the state-of-the-art GAN variants: Fashion-MNIST dataset with data bias(size/content).

Metrics	NMI	ARI	FID	IS	Acc.	Div.
ACGAN	0.724 /0.641	0.86 /0.70	150.53 /324.58	2.01 /1.13	80.24 /72.56	0.48 /0.22
ClusterGAN	0.642 /0.725	0.49 /0.62	498.32 /158.80	2.11 /1.96	16.86 /1.60	0.65 /0.53
CLS-R GAN	**0.95 /0.98**	**0.99 /0.98**	**131.85 /138.74**	**2.71 /2.56**	**99.99/ 98.3**	**0.76/ 0.59**

Tables 3 and 4 indicate that, similar to the comparative evaluation based on the original MNIST and Fashion-MNIST datasets, the proposed CLS-R GAN with the generator self-training phase shows top-tier performance for the datasets including size and content biases.

VI. Ablation Study

In this section, we examine effects of the application of quality loss in the network, in order to clarify: (a) whether quality loss allows the generator to have more diverse fake data, and (b) whether the encoder leads the generator to cluster fake data more in efficient way. In this study, we have artificially synthesized CLS-R GAN variants by either removing quality-loss or encoder, referred to as 'CLS-R GAN without quality loss' or 'CLS-R GAN without encoder', respectively. Doing so, further evaluation of the proposed qualiyt loss and encoder effect can be fully made. As described in Section V.A., the used metrics include NMI, ARI, IS, and Div. We have trained and evaluated those variants based on the same respective training and testing datasets for a fair comparison.

A. With quality loss CLS-R GAN vs Without quality loss CLS-R GAN.

We compared the diversity in each case with and without using image quality loss by evaluating diversity metrics for class diversity measure(Div.) and IS.

Tables 5 and 6 show the performance with and without quality loss evaluated based on the biased and unbiased MNIST and biased and unbiased Fashion-MNIST datasets, respectively.

Table 5. Comparison between CLS-R GAN with and without quality loss, based on MNIST datasets with or without bias.

Metrics	IS	Div.	IS	Div.	IS	Div.
Dataset	MNIST		MNIST size bias		MNIST content bias	
CLS-R GAN with Quality loss	**1.98**	**0.72**	**1.98**	**0.70**	**1.95**	**0.69**
CLS-R GAN without Quality loss	1.35	0.49	1.17	0.40	1.02	0.36

Table 6. Comparison between CLS-R GAN with quality loss and without it, based on biased and unbiased Fashion-MNIST datasets.

Metrics	IS	Div.	IS	Div.	IS	Div.
Dataset	Fashion-MNIST		Fashion-MNIST size bias		Fashion-MNIST content bias	
CLS-R GAN with Quality loss	**2.73**	**0.79**	**2.71**	**0.76**	**2.56**	**0.59**
CLS-R GAN without Quality loss	2.09	0.58	1.88	0.55	1.73	0.43

Table 5 shows comparative evaluation based on biased and unbiased MNIST datasets. Meanwhile, in Table 6 the comparision is based on biased and unbiased Fashion-MNIST datasets with and without the quality loss. As a result, we identified that the proposed CLS-R GAN with quality loss does indicate superior performance in terms of diversity than CLS-R GAN without quality loss.

B. CLS-R GAN with encoder vs CLS-R GAN without encoder.

Subsequently, to examine effectiveness of encoder for clustering, we compared performance of CLS-R GAN variants between with and without encoder, and also evaluated clustering metrics for NMI and ARI.

Table 7 and 8 show comparison for CLS-R GAN between with and without encoder.

Table 7. Comparison of with CLS-R GAN with and without encoder based on biased and unbiased of MNIST datasets.

Metrics	NMI	ARI	NMI	ARI	NMI	ARI
Dataset	MNIST		MNIST size bias		MNIST content bias	
CLS-R GAN with Encoder	**0.98**	**0.93**	**0.98**	**0.98**	**0.97**	**0.97**
CLS-R GAN without Encoder	0.33	0.44	0.23	0.21	0.20	0.15

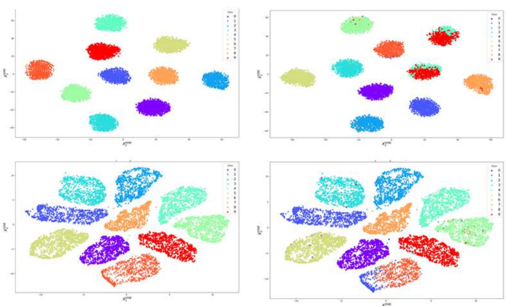
Table 8. Comparison of with CLS-R GAN with and without encoder based on biased and unbiased of Fashion-MNIST datasets.

Metrics	NMI	ARI	NMI	ARI	NMI	ARI
Dataset	Fashion-MNIST		Fashion-MNIST size bias		Fashion-MNIST content bias	
CLS-R GAN with Encoder	**0.96**	**0.98**	**0.95**	**0.99**	**0.98**	**0.98**
CLS-R GAN without Encoder	0.32	0.30	0.27	0.32	0.14	0.25

In the Table 7, the comparison is based on the biased and unbiased MNIST dataset, while in Table 8 it is based on biased and unbiased of Fashion-MNIST dataset. And those findings point out that the proposed CLS-R GAN variant with encoder performs with higher superiority in clustering, than the one without encoder.

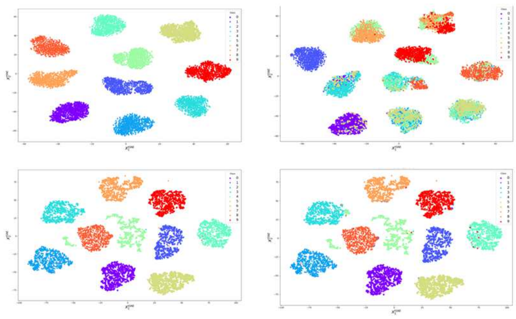
C. *Analysis performance of CLS-R GAN using t-SNE.*

For further exploration on improved performance of CLS-R GAN, we also conducted experiment to compare the t-SNE visualization-results of ClusterGAN and that of CLS-R GAN. Figures 5 to 7 illustrate the t-SNE Visualization clustering results based on biased and unbiased MNIST dataset. Figure 8, 9 and 10 illustrate the t-SNE visualization clustering results, based on biased and unbiased Fashion-MNIST dataset, respectively from the viewpoint of generator and that of high-accuracy classifier.



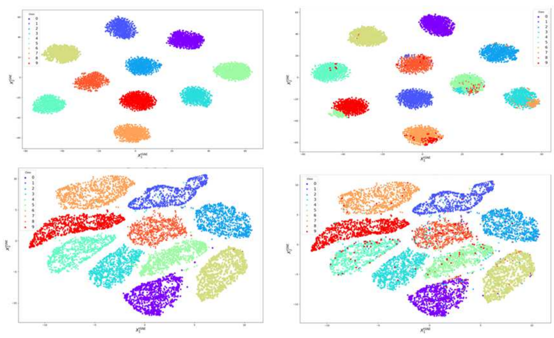
(a) (b)

Fig 5. Visualization of t-SNE result of fake data clustering generated by the generator based on without bias MNIST dataset from the perspective of generator and high accuracy classifier, respectively: (a) the result of generator view point of ClusterGAN t-SNE (top) and CLS-R GAN t-SNE (bottom) (b) the result of high accuracy classifier view point of ClusterGAN t-SNE(top) and CLS-R GAN t-SNE (bottom).



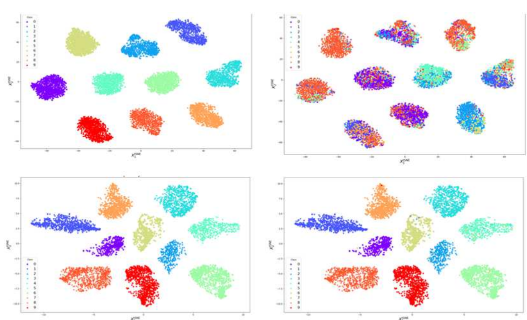
(a) (b)

Fig 6. Visualization of t-SNE result of fake data clustering generated by the generator based on with size bias MNIST dataset from the perspective of generator and high accuracy classifier, respectively: (a) the result of generator view point of ClusterGAN t-SNE (top) and CLS-R GAN t-SNE (bottom) (b) the result of high accuracy classifier view point of ClusterGAN t-SNE(top) and CLS-R GAN t-SNE (bottom).



(a) (b)

Fig 7. Visualization of t-SNE result of fake data clustering generated by the generator based on with contents bias MNIST dataset from the perspective of generator and high accuracy classifier, respectively: (a) the result of generator view point of ClusterGAN t-SNE (top) and CLS-R GAN t-SNE (bottom) (b) the result of high accuracy classifier view point of ClusterGAN t-SNE(top) and CLS-R GAN t-SNE (bottom).



(a) (b)

Fig 8. Visualization of t-SNE result of fake data clustering generated by the generator based on without bias Fashion-MNIST dataset from the perspective of generator and high accuracy classifier, respectively: (a) the result of generator view point of ClusterGAN t-SNE (top) and CLS-R GAN t-SNE (bottom) (b) the result of high accuracy classifier view point of ClusterGAN t-SNE(top) and CLS-R GAN t-SNE (bottom).

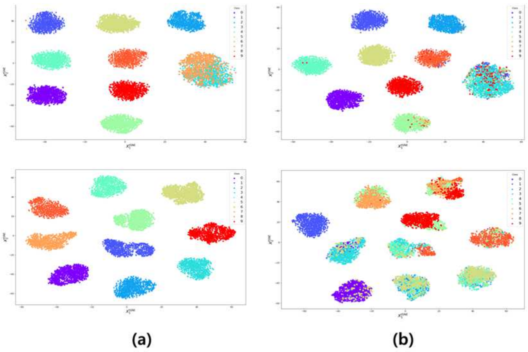


Fig 9. Visualization of t-SNE result of fake data clustering generated by the generator based on with size bias Fashion-MNIST dataset from the perspective of generator and high accuracy classifier, respectively: (a) the result of generator view point of ClusterGAN t-SNE (top) and CLS-R GAN t-SNE (bottom) (b) the result of high accuracy classifier view point of ClusterGAN t-SNE(top) and CLS-R GAN t-SNE (bottom).

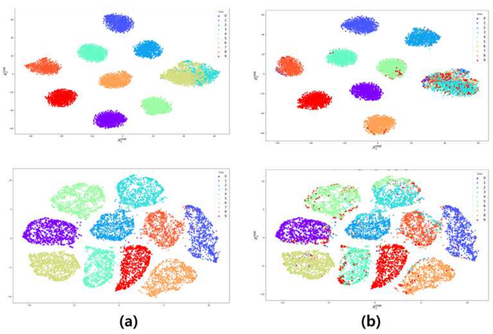


Fig 10. Visualization of t-SNE result of fake data clustering generated by the generator based on with contents bias Fashion-MNIST dataset from the perspective of generator and high accuracy classifier, respectively: (a) the result of generator view point of ClusterGAN t-SNE (top) and CLS-R GAN t-SNE (bottom) (b) the result of high accuracy classifier view point of ClusterGAN t-SNE(top) and CLS-R GAN t-SNE (bottom).

As figures 5 to 10 show, the performance of CLS-R GAN is obviously superior at clustering, compared to t-SNE Visualization ClusterGAN and CLS-R GAN, which were trained on the basis of both MNIST and Fashion-MNIST dataset, regardless of either biased or unbiased. Furthermore, the clustering performance of CLS-R GAN was better even for biased dataset, as well as specifying cluster labels more accurately.

VII. Discussion

Regarding data collection for medical deep-learning, it requires a lot of endeavor and expenses. Given that situation, GAN can be a cost-effective solution. CycleGAN [24], for instance, is used to convert ultrasonic images into CT images, while SRGAN [25] is used to convert low resolution MRI images in 3T, into high resolution MRI images in 7T. However, although the data capacity problem can be solved using GAN, there is no class label for the data generated with GAN. Therefore, in more recent studies, GAN variants such as ACGAN are used to solve these problems. As shown in Section V, ACGAN relatively cannot generate various images whereas CLS-R GAN relatively can generate more diverse augmented images with correct labeled classes. Fig.11 illustrates typical examples of fake data generated by the generator that completed self-training in ultrasonic dataset.

Fig. 11 Typical examples of fake data generated by the generator after self-training completion. Ultrasonic dataset.

Fig. 12 shows some highly deformed instances of qualified fake data in association with ultrasonic data which were chosen by independent classifier for training the generator.

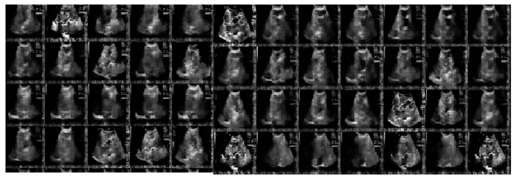


Fig. 12 Highly deformed examples of qualified fake data for ultrasonic dataset that are selected by the independent classifier for training the generator.

VIII. Conclusion And Future Work

In this paper, we presented an approach to GAN-based generation of high quality and diversity of data by introducing a novel self-training framework to GAN, named CLS-R GAN. Extensive experiments, including ablation studies, were conducted to verify the effectiveness of CLS-R GAN based on standard evaluation metrics. The results indicate that CLS-R GAN can provide improved quality and diversity in data augmentation, owing to the novel generator self-training approach based on self-improvement and clustering loops. Future research includes the applications of CLS-R GAN to a larger variety of datasets, including CIFAR10, STL10, ImageNet and medical datasets. In addition, a more thorough investigation on the optimal settings of system parameters and thresholds, including the effect of the proposed fake data selection criteria on data augmentation.

References

[1] Mukherjee, Sudipto, et al. "Clustergan: Latent space clustering in generative adversarial networks." Proceedings of the AAAI conference on artificial intelligence. Vol. 33. No. 01. 2019.

[2] Dam, Tanmoy, Sreenatha G. Anavatti, and Hussein A. Abbass. "Improving ClusterGAN Using Self-AugmentedInformation Maximization of Disentangling LatentSpaces." arXiv preprint arXiv:2107.12706 (2021).

[3] Kim, Yunji, and Jung-Woo Ha. "Contrastive Fine-grained Class Clustering via Generative Adversarial Networks." arXiv preprint arXiv:2112.14971 (2021).

[4] Liang, Jie, et al. "Sub-GAN: An unsupervised generative model via subspaces." Proceedings of the European Conference on Computer Vision (ECCV). 2018.

[5] Andreeva, Olga, et al. "Catalysis clustering with gan by incorporating domain knowledge." Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining. 2020.

[6] Odena, Augustus, Christopher Olah, and Jonathon Shlens. "Conditional image synthesis with auxiliary classifier gans." International conference on machine learning. PMLR, 2017.

[7] Odena, Augustus. "Semi-supervised learning with generative adversarial networks." arXiv preprint arXiv:1606.01583 (2016).

[8] Chen, Xi, et al. "Infogan: Interpretable representation learning by information maximizing generative adversarial nets." Advances in neural information processing systems 29 (2016).

[9] Liu, Zhiyue, Jiahai Wang, and Zhiwei Liang. "Catgan: Category-aware generative adversarial networks with hierarchical evolutionary learning for category text generation." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. No. 05. 2020.

[10] Lin, Min. "Softmax gan." arXiv preprint arXiv:1704.06191 (2017).

[11] Antoniou, Antreas, Amos Storkey, and Harrison Edwards. "Data augmentation generative adversarial networks." arXiv preprint arXiv:1711.04340 (2017).

[12] Hou, Liang, et al. "Self-supervised gans with label augmentation." Advances in Neural Information Processing Systems 34 (2021): 13019-13031.

[13] Tran, Ngoc-Trung, et al. "On data augmentation for gan training." IEEE Transactions on Image Processing 30 (2021): 1882-1897.

[14] Zhang, Dan, and Anna Khoreva. "PA-GAN: Improving GAN training by progressive augmentation." (2018).

[15] Zhao, Shengyu, et al. "Differentiable augmentation for data-efficient gan training." Advances in Neural Information Processing Systems 33 (2020): 7559-7570.

[16] Mirza, Mehdi, and Simon Osindero. "Conditional generative adversarial nets." arXiv preprint arXiv:1411.1784 (2014).

[17] Gulrajani, Ishaan, et al. "Improved training of wasserstein gans." Advances in neural information processing systems 30 (2017)

[18] Arjovsky, Martin, Soumith Chintala, and Léon Bottou. "Wasserstein generative adversarial networks." *International conference on machine learning*. PMLR, 2017.

[19] Dam, Tanmoy, Sreenatha G. Anavatti, and Hussein A. Abbass. "Improving ClusterGAN Using Self-AugmentedInformation Maximization of Disentangling LatentSpaces." arXiv preprint arXiv:2107.12706 (2021).

[20] Mishra, Deepak, Aravind Jayendran, and A. P. Prathosh. "Effect of the latent structure on clustering with gans." IEEE Signal Processing Letters 27 (2020): 900-904.

[21] Tang, Shichang. "Lessons learned from the training of gans on artificial datasets." IEEE Access 8 (2020): 165044-165055.

[22] Cao, Wenming, Zhiwen Yu, and Hau-San Wong. "GAN-based clustering solution generation and fusion of diffusion." Systems Science & Control Engineering 10.1 (2022): 24-42.

[23] Goodfellow, Ian, et al. "Generative adversarial networks." Communications of the ACM 63.11 (2020): 139-144.

[24] Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." Proceedings of the IEEE international conference on computer vision. 2017.

[25] Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

Jiang, Liming, et al. "Deceive D: Adaptive pseudo augmentation for gan training with limited data." Advances in Neural Information Processing Systems 34 (2021): 21655-21667.