Enhanced Conditions Based Deep Convolutional Generative Adversarial Networks

1st KaiXu Chen
Kanazawa University
KaKuma-machi 920-1192
Kanazawa, Japan
xchen@csl.ec.t.kanazawa-u.ac.jp

2nd ZhengXuan Zhao
Kanazawa University
KaKuma-machi 920-1192
Kanazawa, Japan
xzhao@csl.ec.t.kanazawa-u.ac.jp

3th Satoshi Yamane
Kanazawa University
Institute of Science and Engineering
Kakuma-machi 920-1192
Kanazawa, Japan
syamane@is.t.kanazawa-u.ac.jp

***Abstract*—Generative adversarial network (GAN) is a prevalent generative model. While it is effective, it has been shown to be very hard to train in practice. This work demonstrates how an improvement to the GAN framework can be used in a stable training, and in a *conditional* manner able to restrict their generation according to some alternate information such as a class label. Additionally, we explore different GAN structures, showing stable training method between images and attributes.**

***Index Terms*—Generative adversarial Networks(GANs), convolutional nerual networks(CNN), Wasserstein GAN, generative model**

I. Introduction

Generative adversarial network (GAN) [1] is a prevalent generative model. Deep convolutional generative adversarial network (DCGAN) [3], based on traditional generative adversarial networks, introduces convolutional neural networks (CNN) into the training for unsupervised learning to improve the effect of generative networks. Conditional generative adversarial network (CGAN) [2] is a conditional model which adds condition extension into GAN.

The recently proposed Wasserstein GAN (WGAN) [4] makes progress toward stable training of GANs. And the Improved Wasserstein GAN [5] with Gradient Penalty provides more stability while training and overall higher image quality.

In this paper, we present a new generative model called conditional-DCGAN (C-DCGAN), is a combination of DCGAN and CGAN which integrates the feature extraction of convolutional networks and condition auxiliary generative sample for image generation. Then, we use the Wasserstein distance and also use Gradient Penalty like [5].

We evaluate the proposed enhanced conditional-DCGAN model across several datasets and obtained enlightening results.

Our contributions can be summarized as follows:

- We propose a new conditional-DCGAN (C-DCGAN) model.
- We improve the original GAN structures, and use Wasserstein loss function instead of original GAN's loss function.
- We compared the performance of different network structures on different datasets.

II. Related Works

A. Generative Adversarial Networks

Generative models have been showing great success since the advent of GANs [6]. But the issue facing this approach is that there is no control on modes of the data being generated. And the other issue is the Mode collapse and diminished gradient, which make the model difficult to train.

By conditioning the model on additional information it is possible to direct the data generation process. Such conditioning could be based on class labels, on some part of data for inpainting like Conditional GAN [2].

Since the advent of Alexnet, most of the focus has been put on CNNs as a discriminative model. Only until recently have generative models been more focus. The Deep Convolutional GANs (DCGAN) [3] propose and evaluate a set of constraints on the architectural topology that make them stable to train in most settings.

B. Wasserstein Distance

Wasserstein GANs [4] has been shown that even in very simple scenarios the JS divergence does not supply useful gradients for the generator. On the other hand, the EM distance does not suffer from these problems of vanishing gradients.

WGAN-GP [5] uses the Improved Wasserstein GAN with Gradient Penalty, and it provides more stability while training and overall higher image quality.

III. Proposed Method

A. Conditional Deep Convolutional Generative Adversarial Networks

In this study, we present a new generative model called conditional-DCGAN (C-DCGAN), is a combination of DCGAN [3] and CGAN [2], which integrates the feature extraction of convolutional networks and condition auxiliary generative sample for image recognition. The result of simulation experiments shows that this model can improve the accuracy of image generation.

The structure is shown in Fig. 1.

It should be noted that, unlike the original CGAN generator model, which only connects label data when input, in order to enhance the guiding role of label data in training, we connect

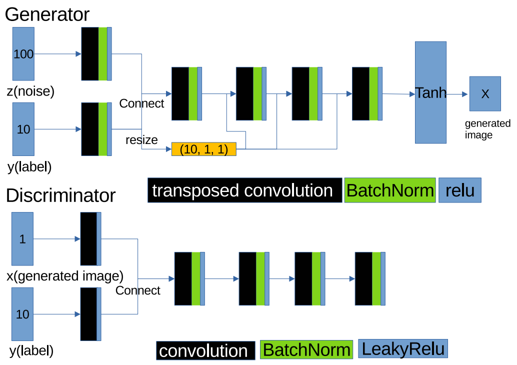

Fig. 1. conditation-DCGAN

the input data z and the input label y of each layer in the model.

B. Algorithm for Conditional DCGAN

And then, we focus on the Wasserstein distance [4], using the gradient penality for ensuring the 1-Lipschitz contsraint of the discriminator [5].

We show our algorithm in Algorithm 1.

Algorithm 1 Enhanced Conditional Deep Convolutional GAN

Require: The gradient penalty coefficient λ, the batch size m, the real data x, the label y.

In each training iteration:

1 Sample m examples $\{x^1, x^2, x^3 \dots x^m\}$ from real data $P_{data}(x)$

2 Sample m noise examples $\{z^1, z^2, z^3 \dots z^m\}$ from the prior $p_{prior}(z)$

3 Obtaining generated data $\{\widetilde{x}^1, \widetilde{x}^2, \widetilde{x}^3 \dots \widetilde{x}^m\}, \widetilde{x}^i = G(z^i)$

4 Update discriminator parameters θ^d to maximize

$\widehat{x} \leftarrow \varepsilon x + (1-\varepsilon)\widetilde{x}$

// $\widehat{x}$ sampling from x and $\widetilde{x}$

$\widetilde{v} = \frac{1}{m}\sum_{i=1}^{m} D_w(x^i|y) - \frac{1}{m}\sum_{i=1}^{m} D_w(\widetilde{x}^i|y) + \lambda(\| \bigtriangledown \widehat{x} Dw(\widehat{x}|y)\|_2 - 1)^2$

$\theta^d \leftarrow \theta^d + \eta \bigtriangledown \widetilde{v}(\theta^d)$

End discriminator update

5 Sample another m noise samples $\{z^1, z^2, z^3 \dots z^m\}$ from the prior $p_{prior}(z)$

$\widetilde{v} = -\frac{1}{m}\sum_{i=1}^{m} D_w(G(z^i|y))$

End generator update

IV. EXPERIMENTS AND RESULTS

A. Datasets

We first experiment on the MNIST dataset [7], which is a collection of handwritten digits. We use the official train test splits (including the validation in train).

Then, we experiment on the Fashion-MNIST dataset [8], which is a dataset of Zalando's article images—consisting of a training set of 60,000 examples and a test set of 10,000 examples.

B. Results

The results with different dataset showed in $Fig.2$. and $Fig.3$. We can see that, each column is conditioned on one label and each row is a different generated sample. Compared with vanilla GAN, it's able to control image generation conditioned on attributes, as well as improve the training speed. Meanwhile we greatly stabilize training with the use of Wasserstein distance and the gradient penality. The result of simulation experiments shows that this model can control image generation and have stable training.

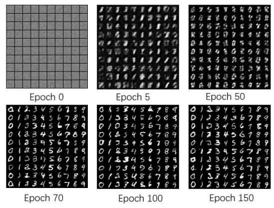

Fig. 2. MNIST

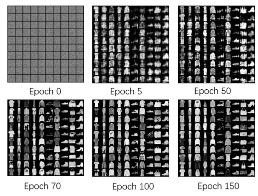

Fig. 3. Fashion-MNIST

V. CONCLUSION

We show a stable generative method for image generation conditioned on attributes. Apart from many other methods, we show experiments that use continuous attributes. In comparison to the original GAN loss, WGAN-GP offers a more stable training environment as well as higher quality samples.

As a result, the proposed model has a higher performance. Despite the longer training time, we believe these improvements are very advantageous in other applications.

REFERENCES

[1] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. arXiv preprint arXiv:1406.2661, 2014.
[2] Mirza M, Osindero S. Conditional generative adversarial nets. arXiv 2014[J]. arXiv preprint arXiv:1411.1784, 2014.
[3] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv preprint arXiv:1511.06434, 2015.
[4] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks[C]//International conference on machine learning. PMLR, 2017: 214-223.
[5] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of wasserstein gans[J]. arXiv preprint arXiv:1704.00028, 2017.
[6] Goodfellow, Ian J., Pouget-Abadie, Jean, Mirza, Mehdi, Xu, Bing, Warde-Farley, David, Ozair,Sherjil, Courville, Aaron C., and Bengio, Yoshua. Generative adversarial nets. NIPS, 2014.
[7] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 86(11):2278-2324, November 1998.
[8] Xiao H, Rasul K, Vollgraf R. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms[J]. arXiv preprint arXiv:1708.07747, 2017.