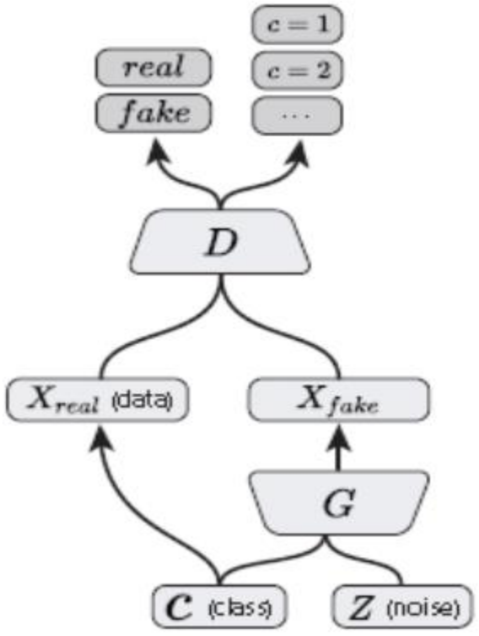
Active Generative Adversarial Network for Image Classification

Quan Kong,[*†] Bin Tong,[*†] Martin Klinkigt,[†]
Yuki Watanabe,[‡] Naoto Akira,[†] Tomokazu Murakami[†]
[†]Hitachi, Ltd. R&D Group, Japan
[‡]Hitachi America, Ltd. USA
{quan.kong.xz, bin.tong.hh, martin.klinkigt.ut}@hitachi.com
{naoto.akira.vu, tomokazu.murakami.xr}@hitachi.com
yuki.watanabe@hal.hitachi.com

AAAI 2019

Background

AC-GAN（Auxiliary Classifier GAN）



AC-GAN
(Present Work)

$$L^D_{\text{AC-GAN}} = \mathbb{E}[\log P(\text{real}|\mathbf{x}_i)] + \mathbb{E}[\log P(\text{fake}|\widehat{\mathbf{x}}_i)] + \mathbb{E}[\log P(y_i|\mathbf{x}_i)] + \mathbb{E}[\log P(y_i|\widehat{\mathbf{x}}_i)]$$

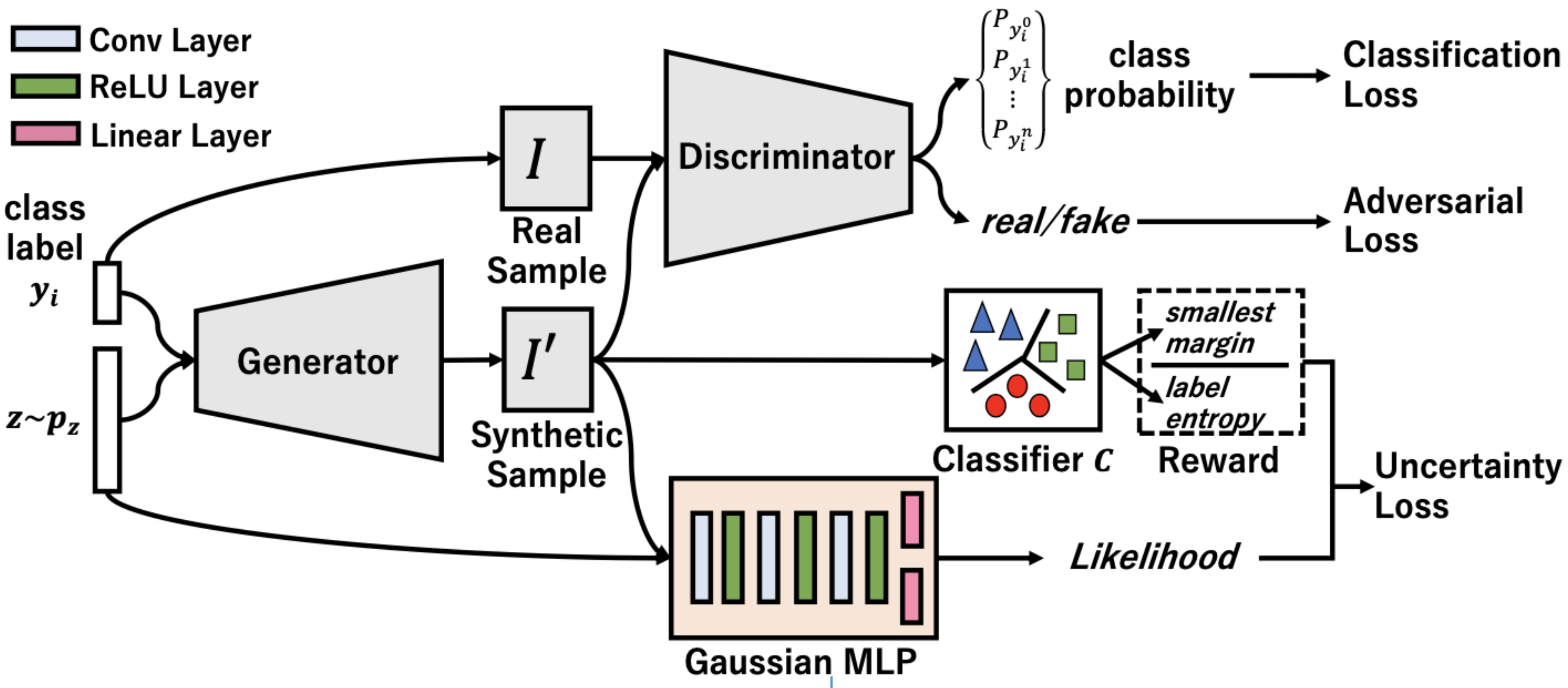
$$L^G_{\text{AC-GAN}} = \mathbb{E}[\log P(\text{real}|\widehat{\mathbf{x}}_i)] + \mathbb{E}[\log P(y_i|\widehat{\mathbf{x}}_i)].$$

→ generating labeled samples

However, for a class of samples, most generated samples may fall less informative.

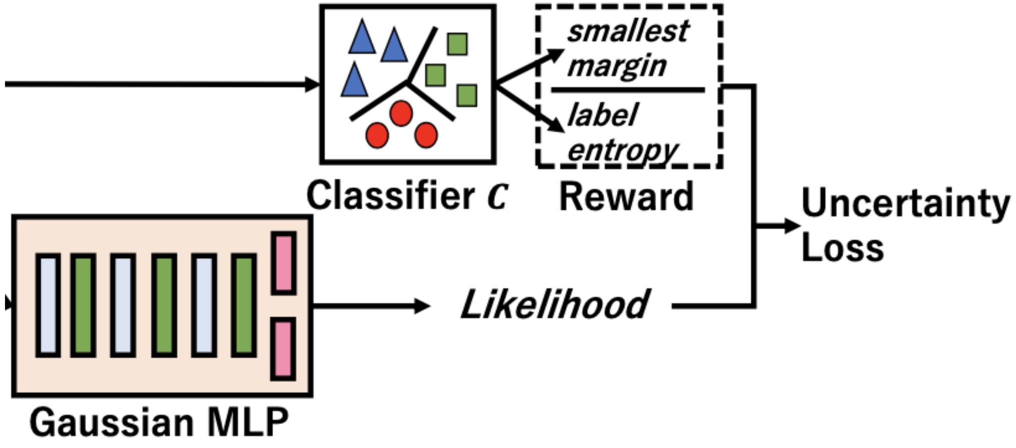
In the idea of active learning, a classifier trained with existing labeled data provides a signal to determine if a sample is informative.

ActiveGAN



calculates a likelihood of a sample to be generated from current distribution of synthetic samples.

ActiveGAN



smallest margin: $u_m = P(y'_1|\hat{x}_i) - P(y'_2|\hat{x}_i)$
label entropy: $u_{le} = -\sum_{y'} P(y'|\hat{x}_i) \log P(y'|\hat{x}_i)$

$$\mathbf{r}_m(\widehat{\mathbf{x}}_i) = \begin{cases} e^{-u_m}, & \text{if } u_m \leq \epsilon \\ C, & \text{otherwise} \end{cases}$$

$$\mathbf{r}_{le}(\widehat{\mathbf{x}}_i) = \begin{cases} e^{u_{le}}, & \text{if } u_{le} \geq \epsilon \\ C, & \text{otherwise} \end{cases}$$

$$\mathbf{r}(\widehat{\mathbf{x}}_i) = \alpha \cdot \mathbf{r}_m(\widehat{\mathbf{x}}_i) + (1-\alpha) \cdot \mathbf{r}_{le}(\widehat{\mathbf{x}}_i),$$

$$L_{\text{uncertainty}} = \sum_{\widehat{\mathbf{x}}_i} \mathbf{r}(\widehat{\mathbf{x}}_i) P(\widehat{\mathbf{x}}_i|\theta),$$

Loss function of Active GAN：

$$\begin{aligned} L^G_{\text{ActiveGAN}} &= L^G_{\text{AC-GAN}} + \lambda L_{\text{uncertainty}} \\ &= \mathbb{E}[\log P(\text{real}|\widehat{\mathbf{x}}_i)] + \mathbb{E}[\log P(y|\widehat{\mathbf{x}}_i)] \\ &\quad + \lambda \mathbb{E}[P(\widehat{\mathbf{x}}_i|\theta)\mathbf{r}(\widehat{\mathbf{x}}_i)], \end{aligned}$$

$$\begin{aligned} L^D_{\text{ActiveGAN}} &= L^D_{\text{AC-GAN}} \\ &= \mathbb{E}[\log P(\text{real}|\mathbf{x}_i)] + \mathbb{E}[\log P(\text{fake}|\widehat{\mathbf{x}}_i)] + \\ &\quad \mathbb{E}[\log P(y_i|\mathbf{x}_i)] + \mathbb{E}[\log P(y_i|\widehat{\mathbf{x}}_i)] \end{aligned}$$

ActiveGAN

Algorithm 1 ActiveGAN

Input training data $\mathbf{x}_i$ and its label y_i where $i \in [1, \ldots, N]$.

Output Ψ_d (parameters of D), Ψ_g (parameters of G) and θ (parameters of MLP)

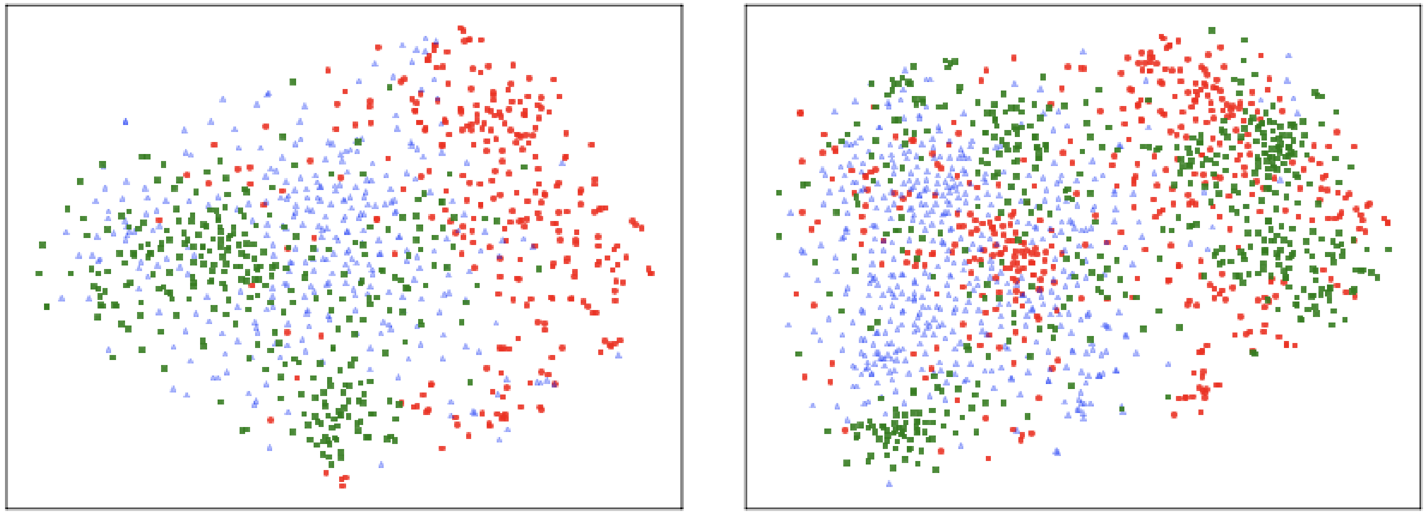
Initialize α, λ, θ, Ψ_d and Ψ_g.
Set the buffer size to be M
Train SVM with grid-search for best parameters
Train the generator G and the discriminator D with first m iterations
Save generated samples in m iterations into the buffer
repeat
    Generate a sample $\widehat{\mathbf{x}}_i \leftarrow G(z, y_i)$
    Use Equation 9 to calculate the reward $\mathbf{r}(\widehat{\mathbf{x}}_i)$ for $\widehat{\mathbf{x}}_i$.
    Use generated samples to calculate the likelihood $P(\widehat{\mathbf{x}}_i|\theta)$ for $\widehat{\mathbf{x}}_i$
    Use Equation 10 to calculate the loss L_U related to the degree of *uncertainty* for $\widehat{\mathbf{x}}_i$
    Update parameters for the generator G and MLP: $\Psi_g,\theta \leftarrow (\Psi_g,\theta) + \bigtriangledown_{\Psi_g,\theta} L^G_{\text{ActiveGAN}}(\Psi_g, \theta)$
    Update parameters for the discriminator D: $\Psi_d \leftarrow \Psi_d + \bigtriangledown_{\Psi_d} L^D_{\text{AC-GAN}}$
    Update the buffer by adding the sample $\widehat{\mathbf{x}}_i$
until

Experiment

Table 1: F-score of models on CIFAR-10, MNIST, Fashion-MNIST (F-MNIST) and Tiny-ImageNet. n represents the number of labeled images used for training.

	CIFAR-10		MNIST		F-MNIST		Tiny-ImageNet	
Method	n=5k	n=10k	n=500	n=1k	n=5k	n=10k	n=10k	n=20k
Baseline(SVM)	83.4	85.3	94.6	96.2	87.1	88.1	56.1	58.3
AC-GAN	81.4	82.7	94.1	95.8	85.4	86.4	52.2	56.1
AC-GAN+F	82.5	83.2	94.5	95.9	86.2	87.3	53.2	56.9
ActiveGAN	**84.3**	**86.3**	**95.1**	**96.5**	**87.6**	**89.0**	**57.5**	**59.4**

Experiment



(a) AC-GAN (b) ActiveGAN

Figure 4: This is t-SNE visualization of the generated samples from AC-GAN (a) and ActivGAN (b) on CIFAR-10,

green points: generated samples
blue points: training samples randomly sampled from each class
red points: real hard samples selected from test pool by using smallest margin

Experiment

Figure 5: Samples of generated images from ActiveGAN in the dataset CIFAR-10. Each column shares the same label and each row shares the same latent variables.

Thanks