GAN-based Gaussian Mixture Model Responsibility Learning

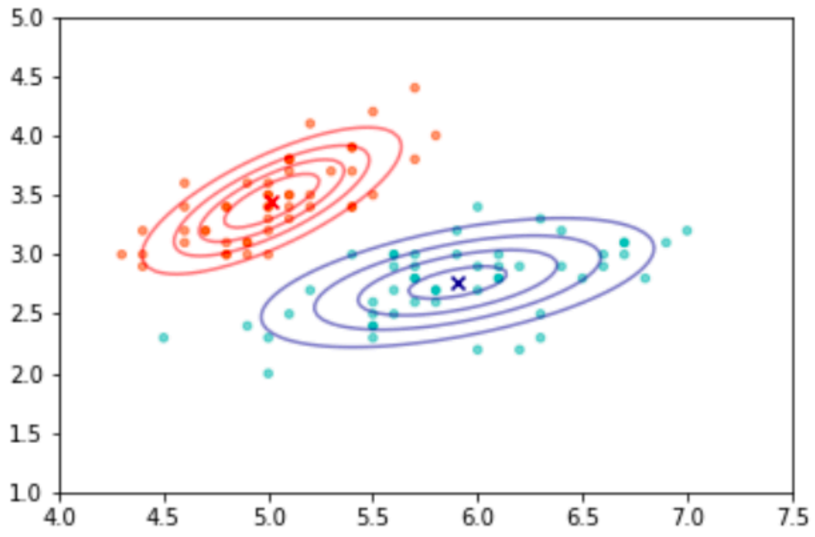
Wanming Huang, UTS
Richard Yi Da Xu, UTS
Shuai Jiang, UTS
Xuan Liang, UTS
Ian Oppermann, NSW Data Analytics Center

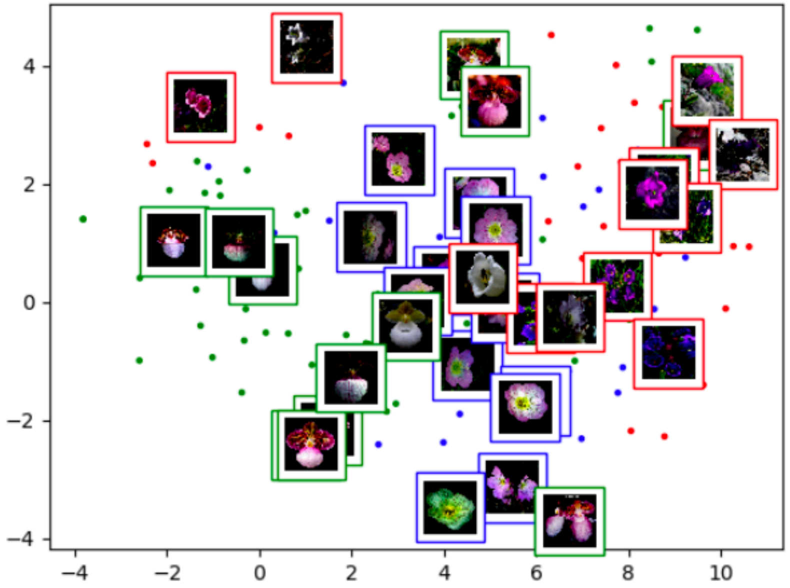
Motivation

- Modern dataset often contains multiple unlabeled modes
- Gaussian Mixture Model models such datasets
- Important statistics can be retrieved, e.g, soft clustering membership, weights of each component

Challenge

- Complex and high dimensional data, such as images, does not form mixture naturally





Solution

- Transform the data x into its latent representation z deterministically
- Model z with Gaussian Mixture

- A natural choice is variational auto-encoder, however, VAEs often lead to blur images
- Generative adversarial nets

Challenge

- The generation process of GANs is one-directional

➢ Posterior consistency module (PCM) maps x to z

Posterior Consistency Module (PCM)

- Returns softmax outputs $\widehat{w} = (\widehat{w}_1, \cdots, \widehat{w}_K)$
- Makes 2 comparisons:
- $p(k|\hat{x}, \theta)$ & $p(k|x, \theta)$
- $p(k|\hat{x}, \theta)$ & $p(k|z, \theta)$
- $p(k|z, \theta) = w \equiv (w_1, \cdots, w_K) = (\frac{N(Z|\mu_1, \sigma_1)}{\sum_{k=1}^{K} N(Z|\mu_k, \sigma_k)}, \cdots, \frac{N(Z|\mu_K, \sigma_K)}{\sum_{k=1}^{K} N(Z|\mu_k, \sigma_k)})$

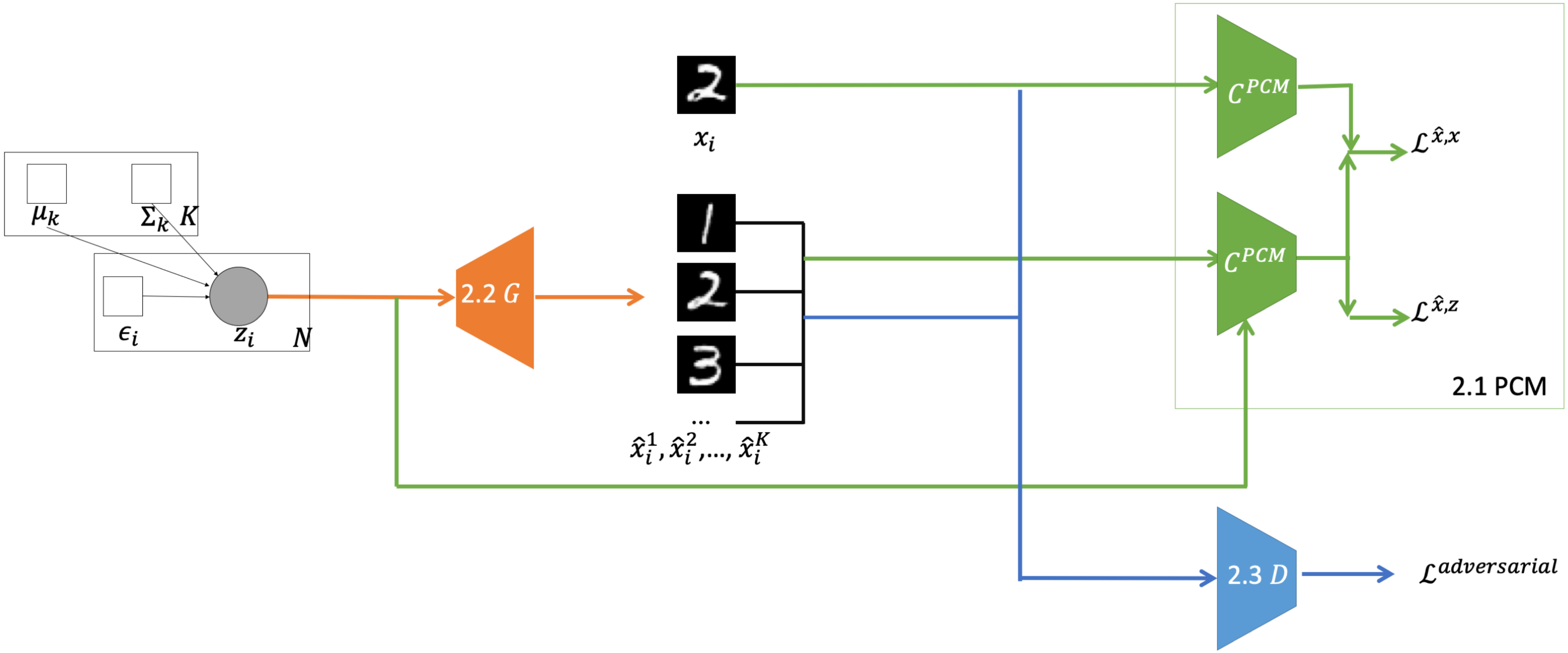
GAN

Ideally, samples generated from the same Gaussian should be similar, however, such correspondence cannot be controlled in an unsupervised training

Therefore, during training, K samples are generated from K modes, weights of each are measured by PCM

$$\mathcal{L}^{adversarial} = \mathbb{E}_{x_i \sim p_{data}}\left(\frac{1}{K}\sum_{k=1}^{K} p(k|x_i, \theta) \times (\log D(x_i) + \log(1 - D(\hat{x}^k)))\right)$$

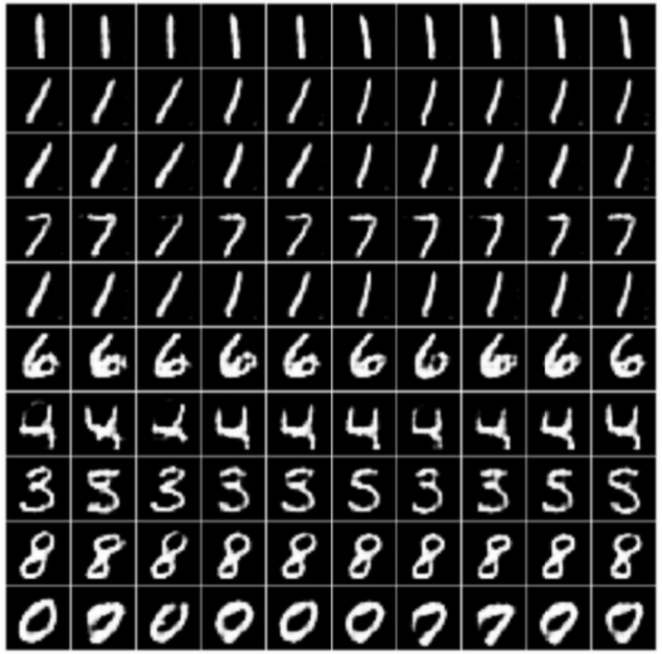
Architecture



Performance on highly imbalanced dataset

Vanilla GAN

Gaussian Mixture GAN w/o PCM



Proposed method

Generation quality

	number of parameters	Inception Score ↑	FID score ↓
Proposed (encoding not shared)	$13,005,411$	2.9664 ± 0.2188	231.0577 ± 7.5371
Proposed (encoding shared)	$8,794,835$	**3.1368** $\pm$ **0.1596**	**205.9776** $\pm$ **7.8587**
GM-GAN	$8,467,145$	2.6770 ± 0.1079	239.3936 ± 6.7672
Vanilla GAN	$8,366,145$	2.4882 ± 0.1065	247.0610 ± 7.2361

Linear interpolation over 3 modes