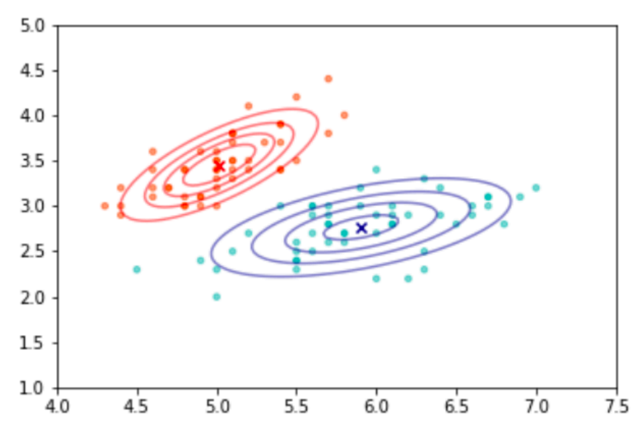
GAN-based Gaussian Mixture Model Responsibility Learning

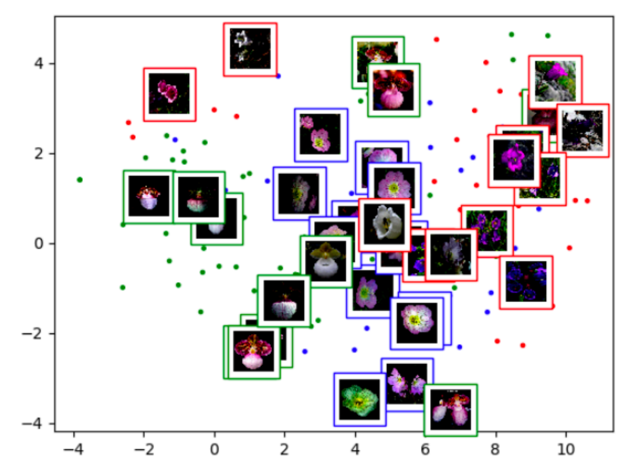
Wanming Huang, UTS; Richard Yi Da Xu, UTS; Shuai Jiang, UYS; Xuan Liang, UTS; Ian Oppermann, NSW Data Analytics Center

MOTIVATION

1. Modern datasets often contains multiple unlabelled modes
2. Gaussian Mixture Model modes such datasets
3. Important statistics can be retrieved, e.g., soft clustering membership, weights of each component
4. However, complex and high dimensional data, such as images, does not form mixtures naturally in their raw forms



Low-dimensional data clustering



High-dimensional data clustering

AIM

1. Transform the data x into its latent representation z deterministically
2. Model z with Gaussian Mixture

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

- ➢ Generative adversarial networks
- ➢ Posterior consistency module (PCM) that maps x to z

POSTERIOR CONSISTENCY MODULE

1. Returns softmax outputs $\widehat{w} = (\widehat{w}_1, \cdots, \widehat{w}_K)$
2. Feature encoding is shared with the discriminator
3. Makes 2 comparisons
   - $p(k|\hat{x},\theta)$ & $p(k|x,\theta)$
   - $p(k|\hat{x},\theta)$ & $p(k|z,\theta) =$

$$\left(\frac{N(z|\mu_1,\sigma_1)}{\sum_{k=1}^{K} N(z|\mu_k,\sigma_k)}, \cdots, \frac{N(z|\mu_K,\sigma_K)}{\sum_{k=1}^{K} N(z|\mu_k,\sigma_k)}\right)$$

4. Loss functions:

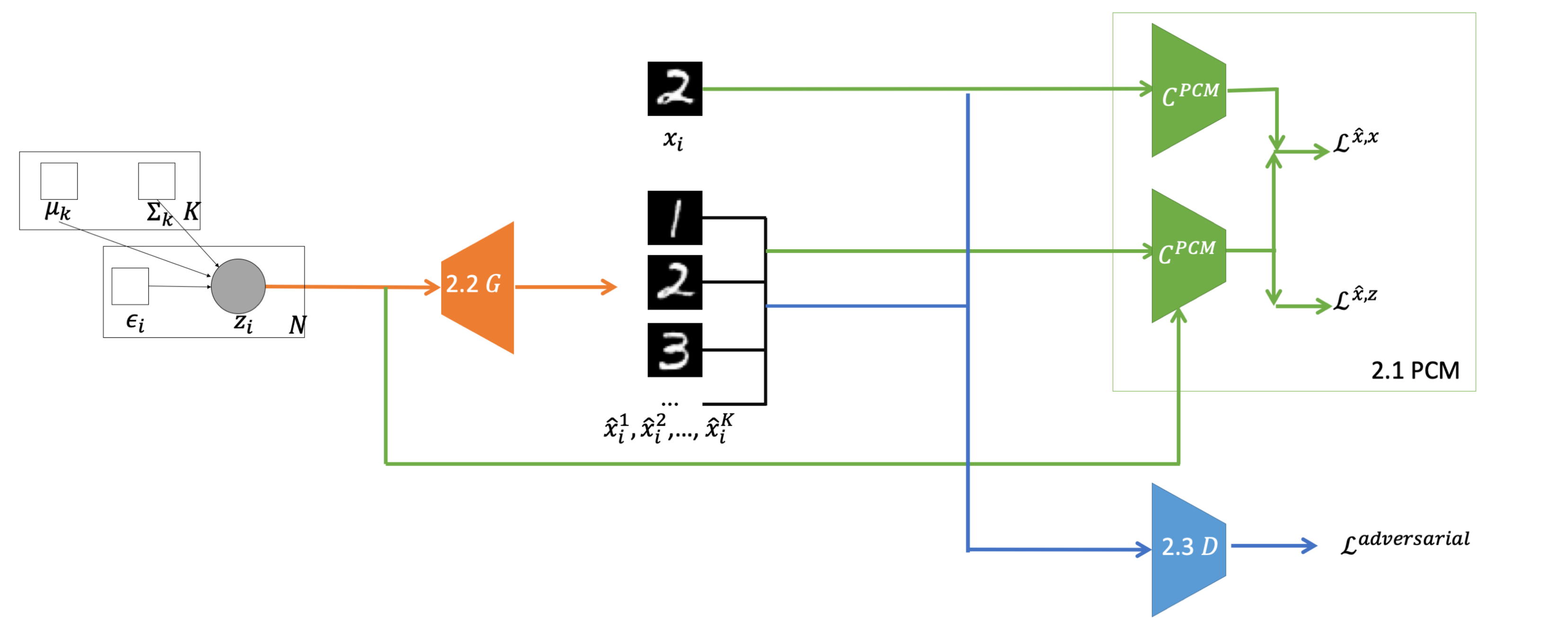
- $\mathcal{L}^{\hat{x},z} = \mathbb{E}_{\{z_1\sim\mathcal{N}(\mu_1,\Sigma_1),\cdots,z_K\sim\mathcal{N}(\mu_K,\Sigma_K)\}}[\frac{1}{K}\sum_{k=1}^{K} I(p(k|z_k,\theta), C_{PCM}^{\theta}(\hat{x}^k))]$
- $\mathcal{L}^{\hat{x},x} = \mathbb{E}_{x_i\sim p_{data}}[\sum_{k=1}^{K} I(C_{PCM}^{\theta}(x_i), C_{PCM}^{\theta}(\hat{x}))]$
- $I(X;Y) = \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}} p_{(X,Y)}(x,y)\log(\frac{p_{(X,Y)}(x,y)}{p_X(x)p_Y(y)})$

GAN

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

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

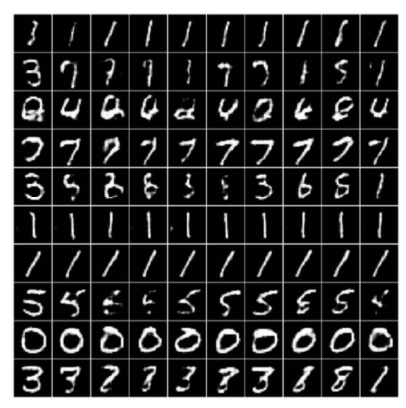
ARCHITECTURE



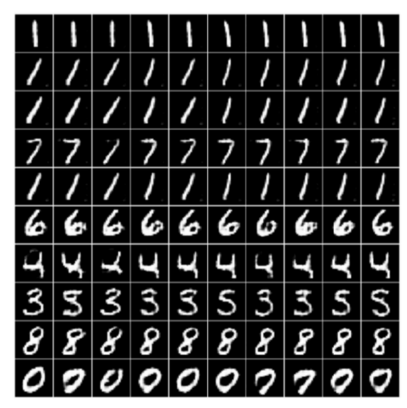
RESULTS

- Performance on highly imbalanced dataset

Vanilla GAN

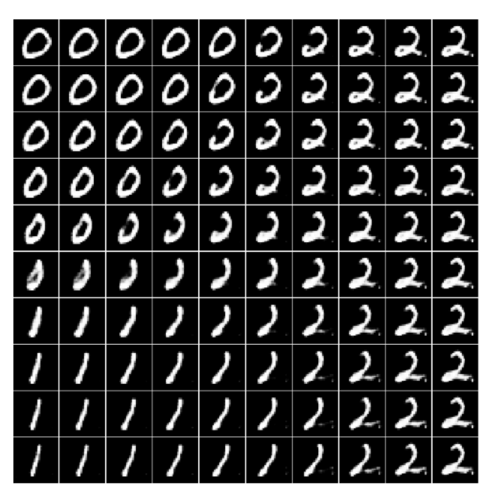


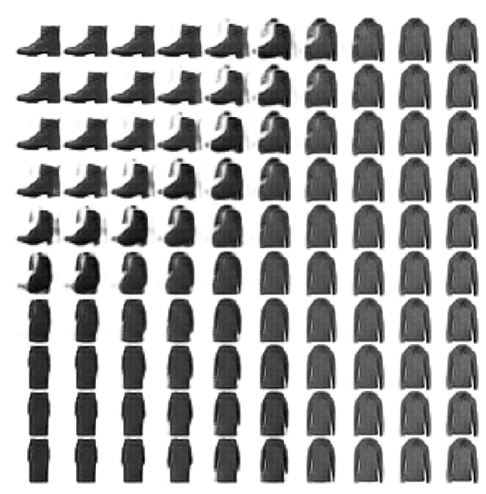
Gaussian mixture GAN w/o PCM



Proposed

- Linear interpolation over 3 modes





- Image 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 ± 0.1596**	**205.9776 ± 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

CONCLUSIONS

1. The latent space of GAN is modelled by Gaussian mixture
2. a posterior consistency module was innovated to help the model to better approximate GMM's responsibility distribution

References

- M. Ben-Yosef and D. Weinshall, "Gaussian mixture generative adversarial networks for diverse datasets, and the unsupervised clustering of images," *CoRR*, vol. abs/1808.10356, 2018. [Online]. Available: http://arxiv.org/abs/1808.10356
- N. Dilokthanakul, P. A. M. Mediano, M. Garnelo, M. C. H. Lee, H. Salimbeni, K. Arulkumaran, and M. Shanahan, "Deep unsupervised clustering with gaussian mixture variational autoencoders," *CoRR*, vol. abs/1611.02648, 2016. [Online]. Available: http://arxiv.org/abs/1611.02648