MOTIVATION

1. Modern datasets often contain 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, do not form mixtures naturally in their raw forms

POSTERIOR CONSISTENCY MODULE

1. Returns softmax outputs \( \hat{w} = (\hat{w}_1, \ldots, \hat{w}_K) \)
2. Feature encoding is shared with the discriminator
3. Makes 2 comparisons
   - \( p(k|x, \theta) & p(k|x, \theta) \)
   - \( p(k|x, \theta) & p(k|x, \theta) = \frac{N(z, \mu_k, \sigma_k)}{N(z, \mu_k, \sigma_k)} \cdot \frac{\sum_{k=1}^{K} N(z, \mu_k, \sigma_k)}{\sum_{k=1}^{K} N(z, \mu_k, \sigma_k)} \)
4. Loss functions:
   - \( L^x = E_{x \sim N(\mu, \sigma)} \) \quad \text{where} \quad p(k|x, \theta) \in [0, 1] \quad \text{and} \quad \sum_{k=1}^{K} p(k|x, \theta) = 1
   - \( L^x = E_{x \sim N(\mu, \sigma)} \) \quad \text{where} \quad p(k|x, \theta) \in [0, 1] \quad \text{and} \quad \sum_{k=1}^{K} p(k|x, \theta) = 1
   - \( I(x; y) = \sum_{x \in X} \sum_{y \in Y} P(x, y) \log \frac{P_{XY}(x, y)}{P(X)P(Y)} \)

GAN

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

\[
L_{\text{adversarial}} = E_{x \sim \text{data}} \sum_{k=1}^{K} \log(D(x_k)) + \log(1 - D(\hat{x}_k))
\]

ARCHITECTURE

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 \( z \)

RESULTS

- Performance on highly imbalanced dataset
- Linear interpolation over 3 modes
- Image quality

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