Neural Characteristic Function Learning for Conditional Image Generation

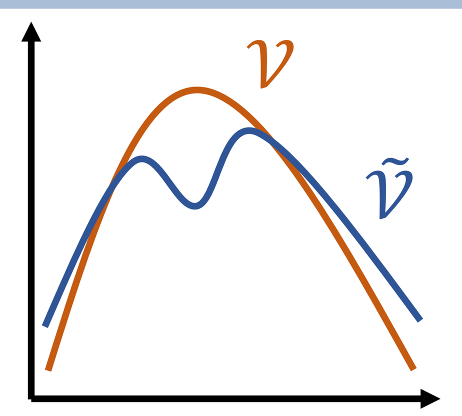
Shengxi Li[1], Jialu Zhang[1], Yifei Li[1], Mai Xu*[1], Xin Deng[2], Li Li[1]

[1]School of Electronic and Information Engineering, Beihang University, Beijing, China [2]School of Cyber Science and Technology, Beihang University, Beijing, China

{LiShengxi, JialuZhang, leafy, MaiXu, cindydeng, lili2005}@buaa.edu.cn

ICCV23 PARIS

Background



Ill-posed discrepancy measure between distributions leads to **mode collapse** and **instability** problems in training cGANs.

Scan Me.

Contribution

I. *New Metric*: We develop neural networks to upper-bound the CF discrepancy, called the neural CF (NCF) metric.

II. *New Architecture*: We establish the CCF-GAN by explicitly modelling conditional distribution.

III. *New Loss*: We propose a conditional CF loss function based on the NCF metric for joint distributions.

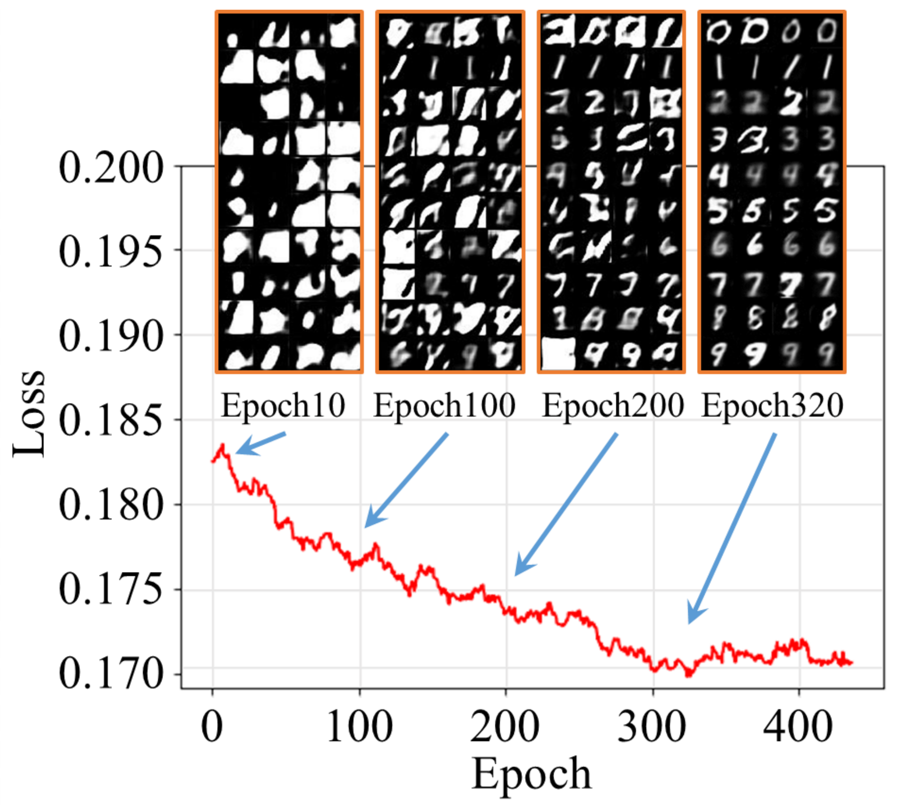
Proposed Method

A new metric

Lemma 1. *For any two random variables $\mathcal{V}, \tilde{\mathcal{V}} \in \mathbb{R}^d$, $\mathcal{L}(\mathcal{V}||\tilde{\mathcal{V}}) \geq \mathcal{D}_{\mathcal{T}}(\mathcal{V}||\tilde{\mathcal{V}})$ for any $\mathcal{T}$.*

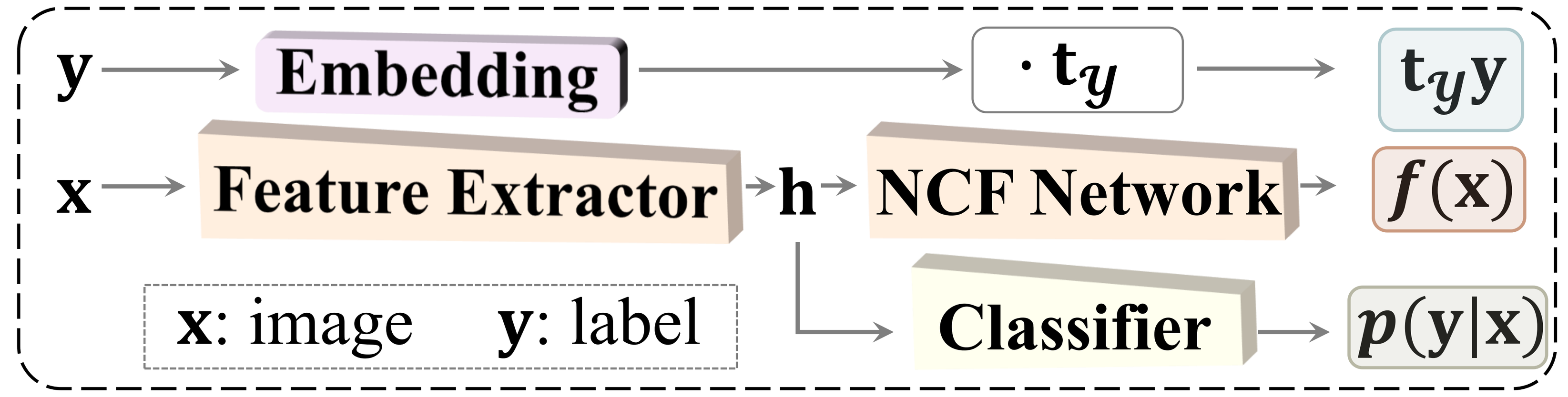
Lemma 2. *If $\mathcal{V}, \tilde{\mathcal{V}} \in \mathbb{R}^d$ are two random variables, $\mathcal{L}(\mathcal{V}||\tilde{\mathcal{V}})$ in (1) is a valid distance metric.*

$$\mathcal{L}(\mathcal{V}||\tilde{\mathcal{V}}) = \max_f \mathcal{D}_{\mathcal{F}}(\mathcal{V}||\tilde{\mathcal{V}}) = \max_f \left(\frac{1}{k}\sum_{i=1}^{k}\left|\Phi_{\mathcal{V}}^{f_i} - \Phi_{\tilde{\mathcal{V}}}^{f_i}\right|^2\right)^{\frac{1}{2}} \quad (1)$$



Optimising generator solely by $\mathcal{D}_{\mathcal{T}}(\mathcal{V}||\tilde{\mathcal{V}})$ for Gaussian $\mathcal{T}$.

A new architecture



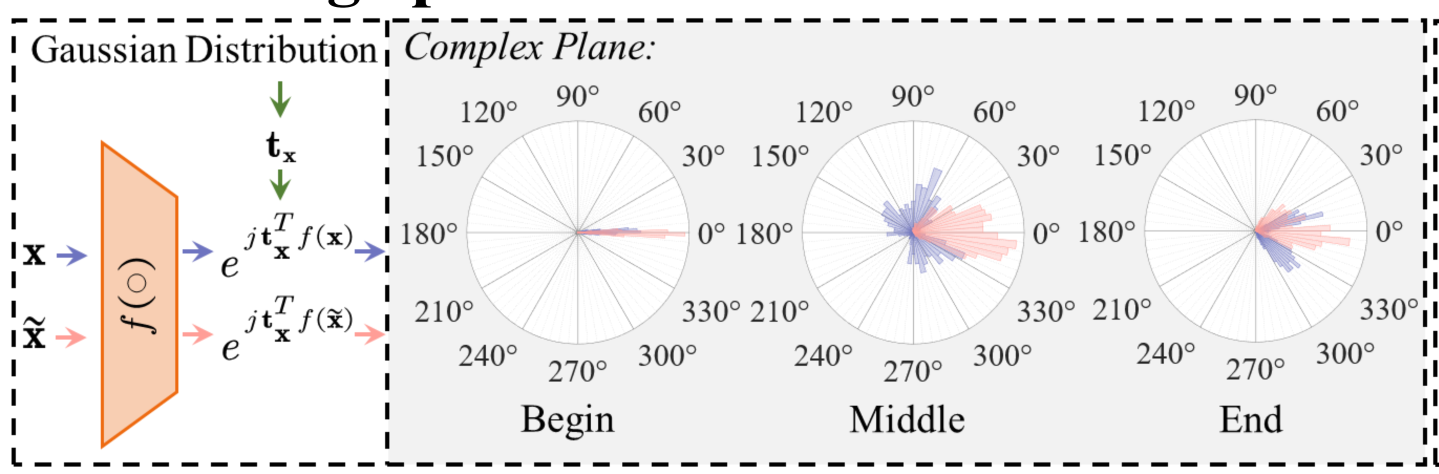
$$\Phi_{\mathcal{V}}^{f}(\mathbf{t}_y) = \Phi_{\mathcal{X},\mathcal{Y}}^{f}(\mathbf{t}_y) = \frac{1}{n}\sum_{i_x=1}^{n}\sum_{i_y=1}^{c} e^{j\mathbf{t}_y^T\mathbf{y}_{i_y}} p\left(\mathbf{y}_{i_y}\middle|\mathbf{x}_{i_x}\right) e^{jf(\mathbf{x}_{i_x})} \quad (2)$$

A new loss

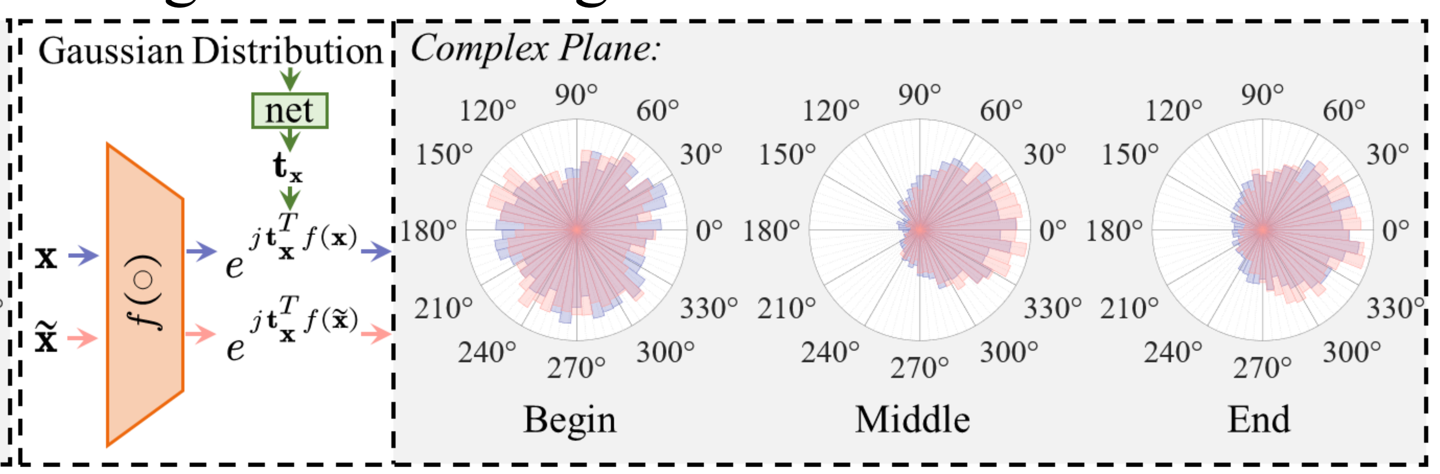
$$\min_g \mathcal{L}(\mathcal{X},\mathcal{Y}||\tilde{\mathcal{X}},\tilde{\mathcal{Y}}) = \min_g \max_f \mathcal{D}_{\mathcal{F}}(\mathcal{X},\mathcal{Y}||\tilde{\mathcal{X}},\tilde{\mathcal{Y}}) = \left(\frac{1}{k}\sum_{i=1}^{k}\left|\Phi_{\mathcal{X},\mathcal{Y}}^{f_i}(\mathbf{t}_y^i) - \Phi_{\tilde{\mathcal{X}},\tilde{\mathcal{Y}}}^{f_i}(\mathbf{t}_y^i)\right|^2\right)^{\frac{1}{2}} \quad (3)$$

Experimental Results

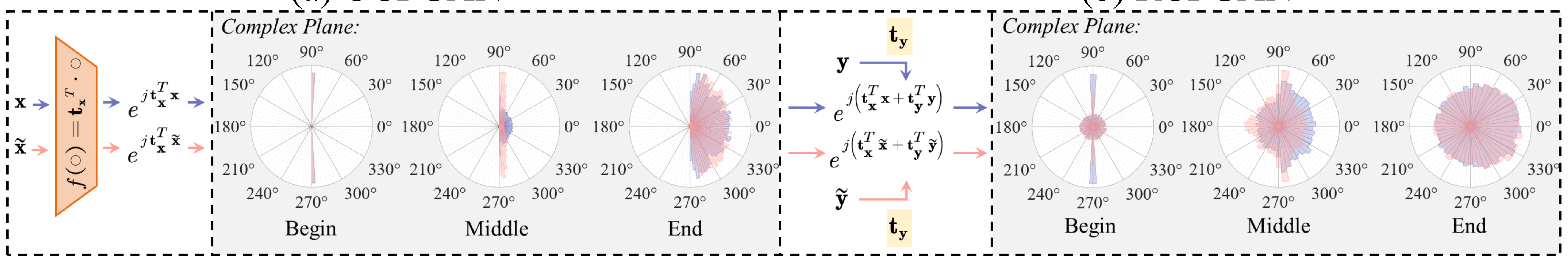
Converge process: Evolution of adversarial learning when training CF-related GANs



(a) OCFGAN

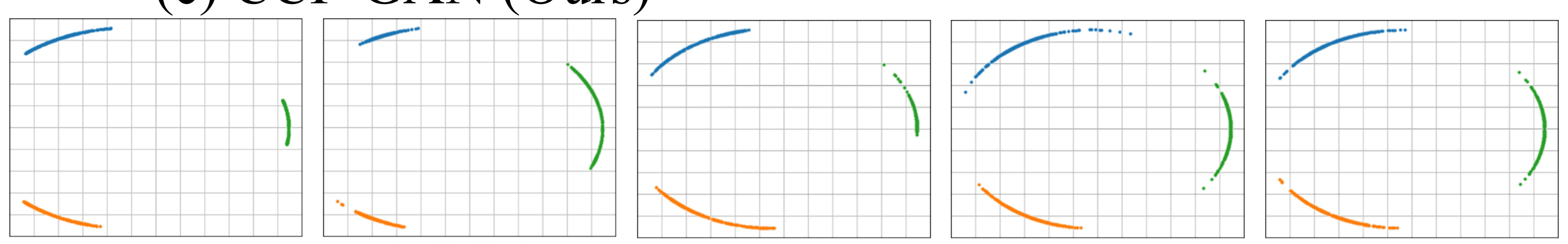


(b) RCFGAN



(c) CCF-GAN (Ours)

Learning in low-dimension support: Distribution fitting results, from the mixture of vMF distributions.

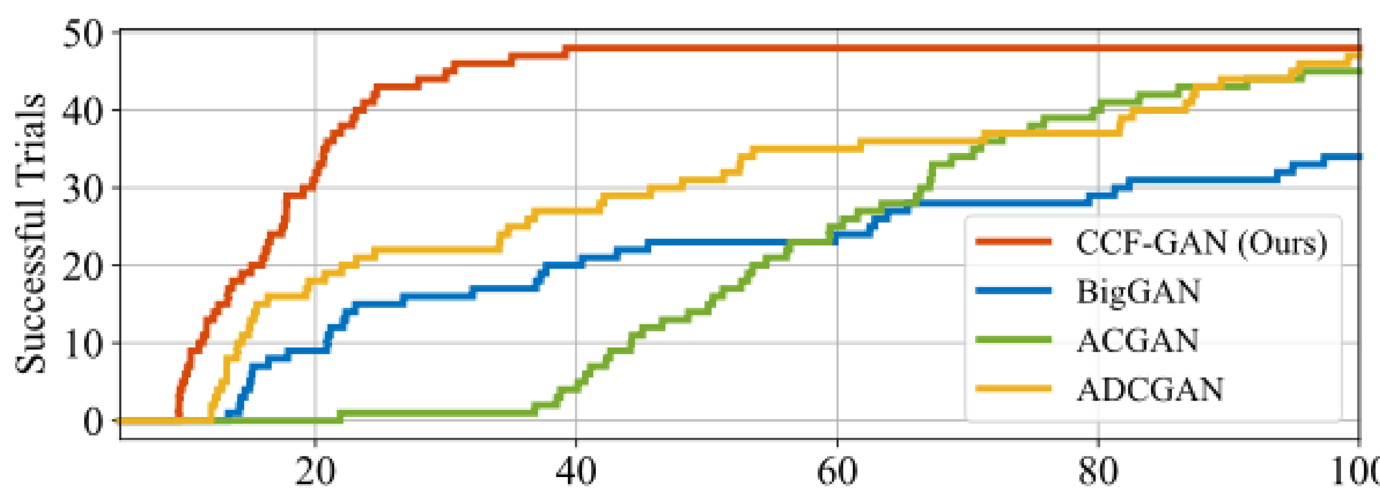


ACGAN TACGAN ADCGAN CCF-GAN **GT**

Quantitative results:

Method	CIFAR10	V200	V500	V1000	ImageNet	
	FID↓				FID↓	IS↑
BigGAN	14.73	66.23	43.10	24.07	22.77	38.05
ACGAN	8.01	95.70	31.90	—	184.41	7.26
FisherGAN	11.46	13.28	9.02	7.30	—	—
TACGAN	8.42	29.12	12.42	13.60	23.75	28.86
cRCFGAN	6.90	27.03	18.03	20.72	—	—
ContraGAN	10.60	—	—	—	19.69	31.10
ReACGAN	6.22	13.48	7.19	6.47	13.98	68.27
ADCGAN	7.17	18.64	11.34	7.94	—	—
CCF-GAN (Ours)	**6.08**	**11.61**	**6.81**	**5.70**	**11.34**	**180.84**

Stability analysis:



Qualitative results: