IEEE International Conference on Multimedia and Expo

6-10 July 2020 || London, United Kingdom || Virtual

MatchingGAN: Matching-Based Few-Shot Image Generation

Yan Hong, Li Niu, Jianfu Zhang, Liqing Zhang

MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University

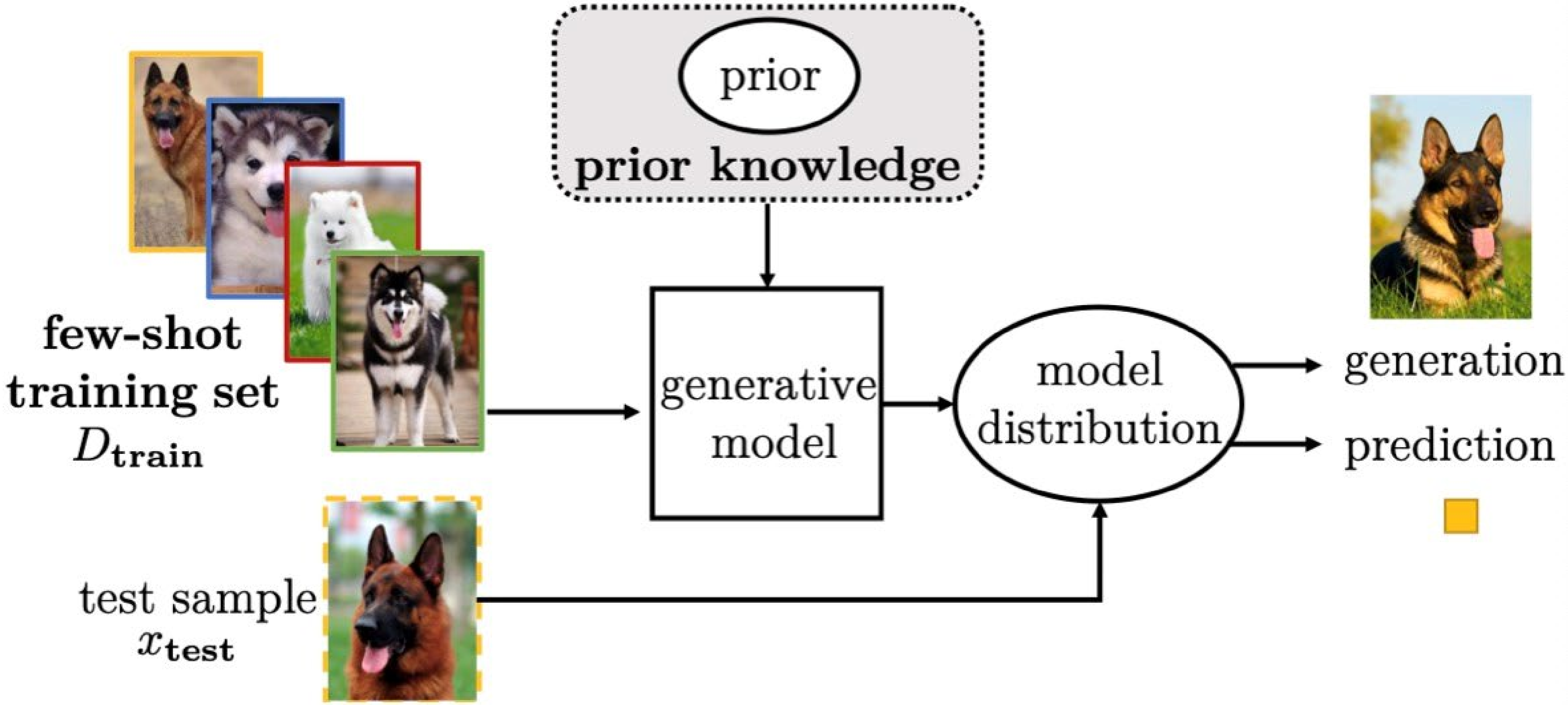
上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Catalog

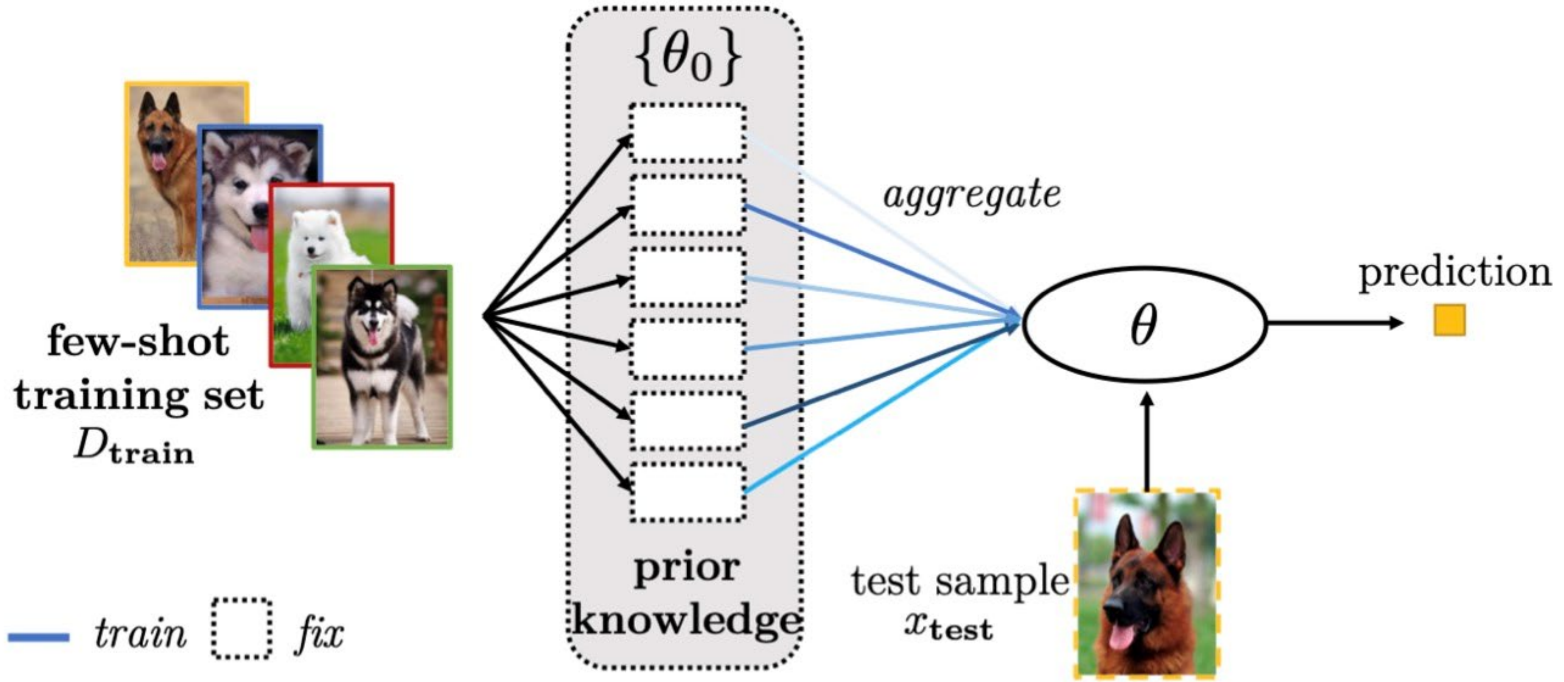
Catalog

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Few-shot Learning



Metric-based Few-shot Learning



The model can be used at testing phase without fine-tuning

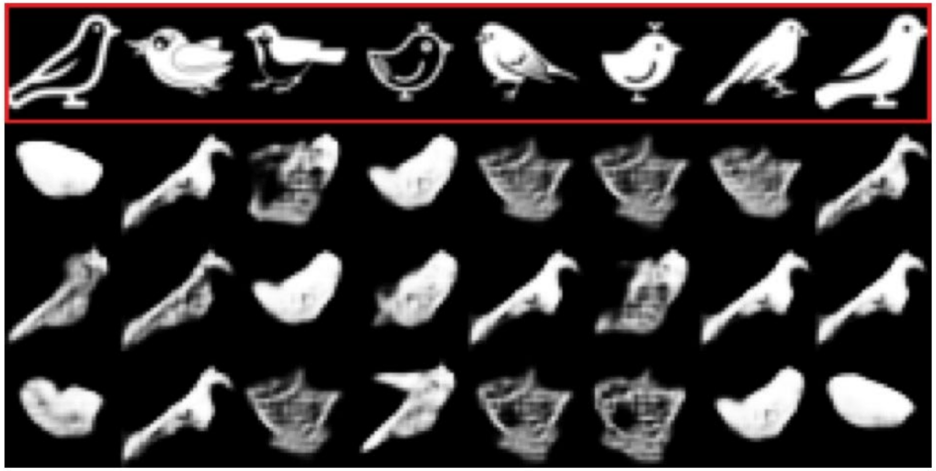
Catalog

Few-shot Image Generation

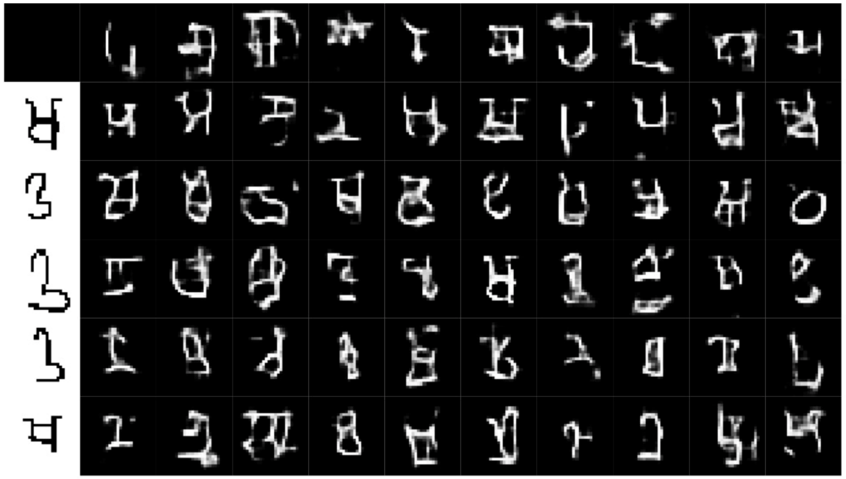
- Extensive training samples is expensive, difficulty in quick adaptation
- Few shot image generation can augment training dataset, facilitate downstream tasks like few-shot classification

The Issues in Existing Few-shot Image Generation

➤ The quality of generated images is poor



Generated images are vague



Generated images are unreasonable

The Issues in Existing Few-shot Image Generation

- The diversity of generated images is limited

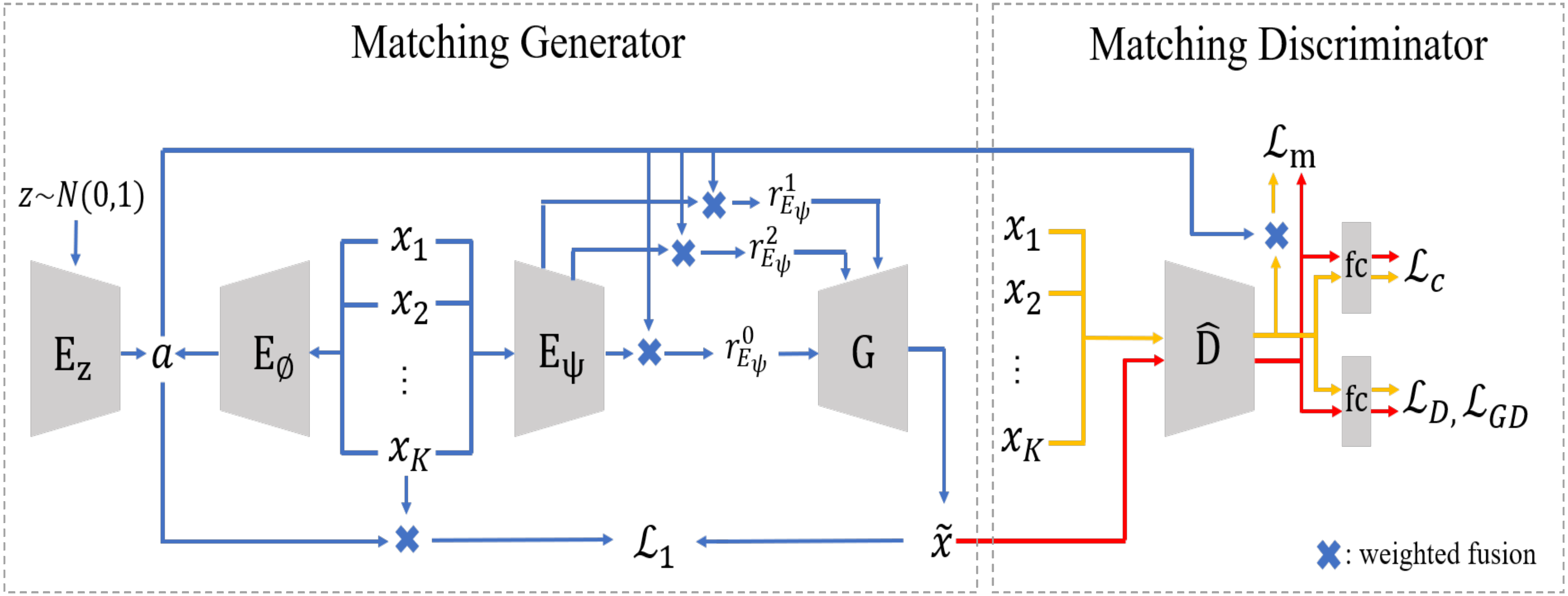
Conditional images

Generated images

Catalog

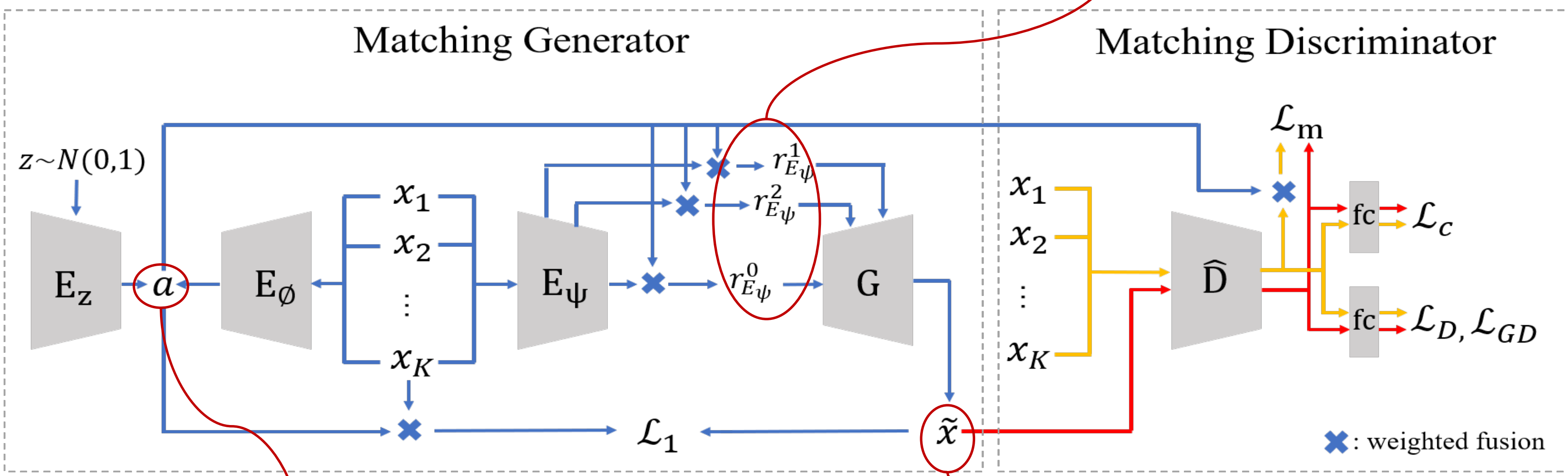
Framework

Combining matching procedure with generative adversarial network



Matching Procedure

$$\boldsymbol{r}_{E_\psi}^{j} = \sum_{i=1}^{K} \mathrm{a}(E_z(\boldsymbol{z}), E_\phi(\boldsymbol{x}_i)) E_\psi^{j}(\boldsymbol{x}_i), \ \ j = 0, \ldots, L$$

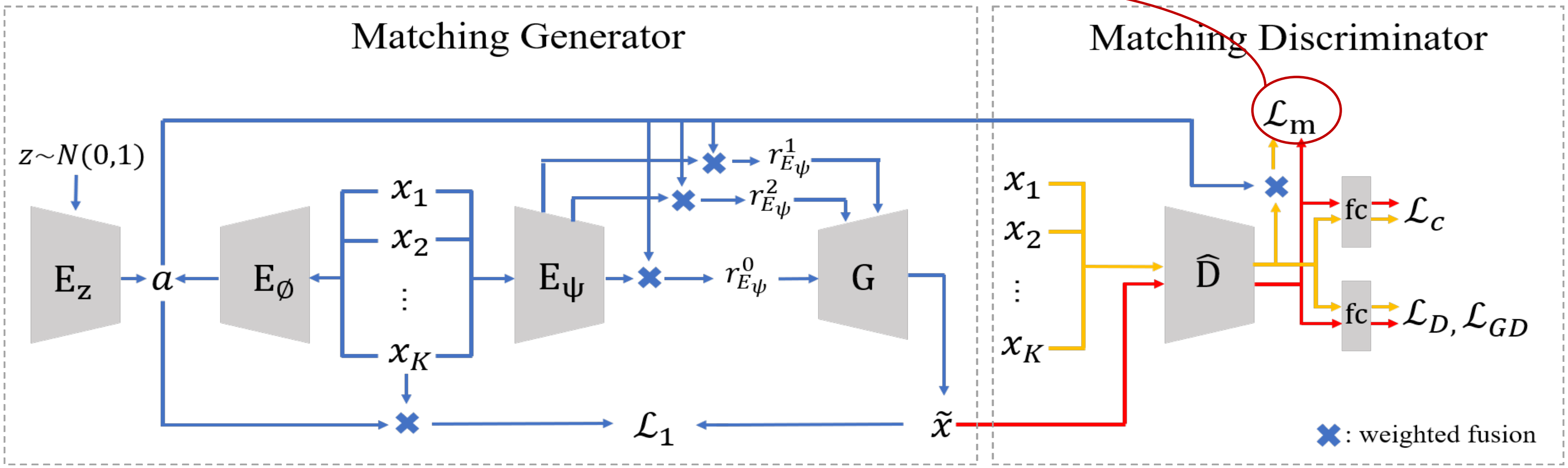


$$\mathrm{a}(E_z(\boldsymbol{z}), E_\phi(\boldsymbol{x}_i)) = \frac{e^{\cos(E_z(\boldsymbol{z}), E_\phi(\boldsymbol{x}_i))}}{\sum_{i=1}^{K} e^{\cos(E_z(\boldsymbol{z}), E_\phi(\boldsymbol{x}_i))}}$$

$$\tilde{\boldsymbol{x}} = G(\boldsymbol{r}_{E_\psi}^{0}, \boldsymbol{r}_{E_\psi}^{1}, \ldots, \boldsymbol{r}_{E_\psi}^{L})$$

Matching Procedure

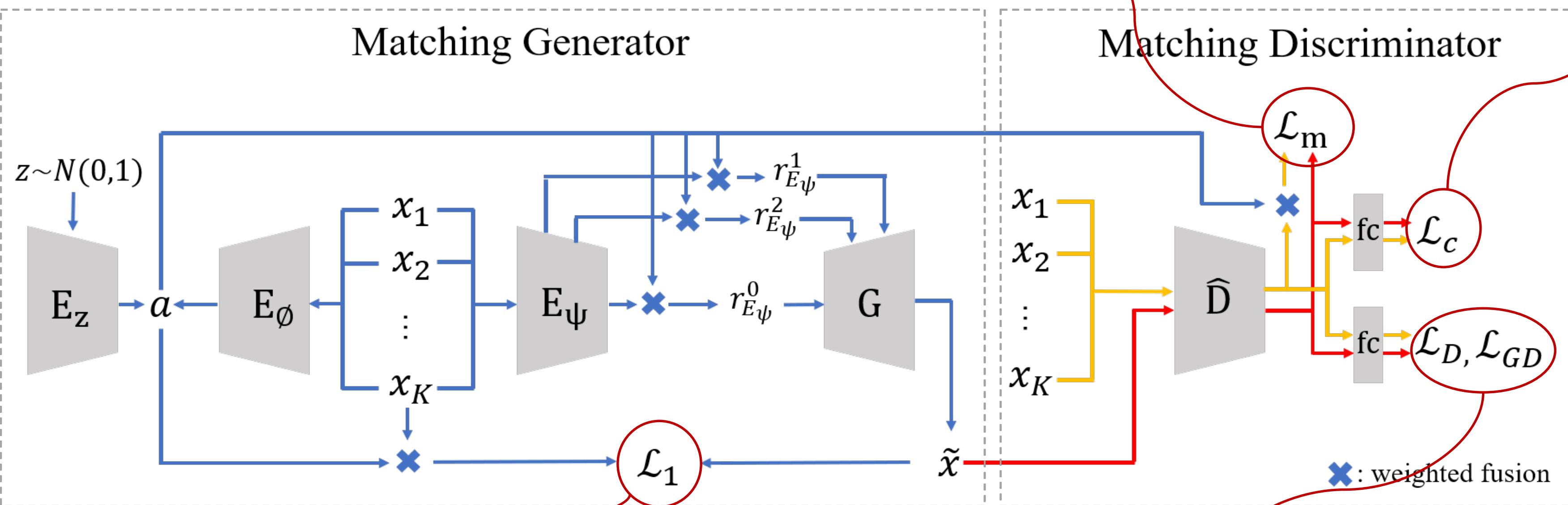
$$\mathcal{L}_m = \|\sum_{i=1}^{K} \mathrm{a}(E_z(\boldsymbol{z}), E_\phi(\boldsymbol{x}_i))\hat{D}(\boldsymbol{x}_i) - \hat{D}(\tilde{\boldsymbol{x}})\|_1$$



Optimization

$$\mathcal{L}_m = ||\sum_{i=1}^{K} \mathrm{a}(E_z(\boldsymbol{z}), E_\phi(\boldsymbol{x}_i))\hat{D}(\boldsymbol{x}_i) - \hat{D}(\tilde{\boldsymbol{x}})||_1$$

$$\mathcal{L}_c = -\log p(c(\boldsymbol{x})|\boldsymbol{x})$$



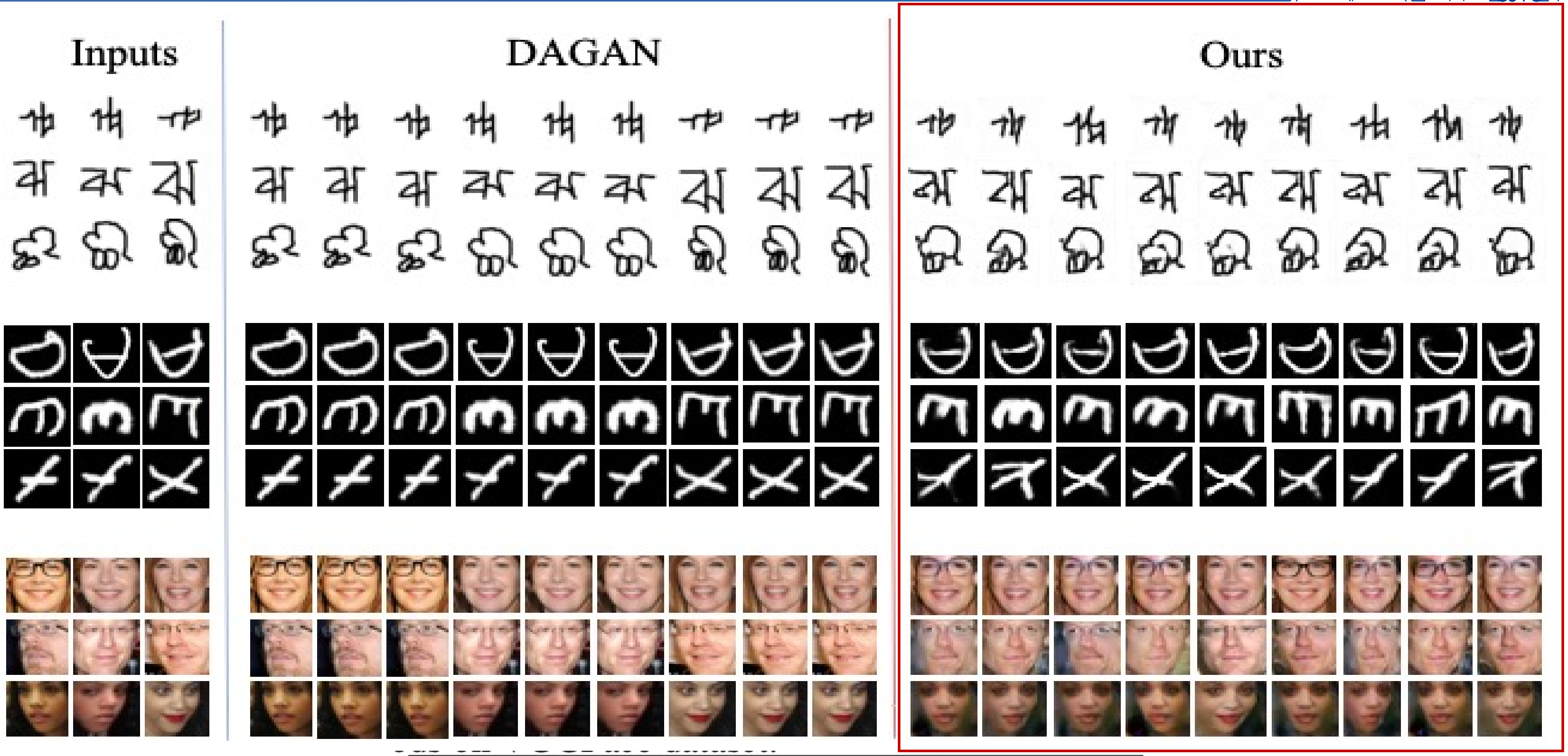
$$\mathcal{L}_1 = \sum_{i=1}^{K} \mathrm{a}(E_z(\boldsymbol{z}), E_\phi(\boldsymbol{x}_i))||\boldsymbol{x}_i - \tilde{\boldsymbol{x}}||_1$$

$$\mathcal{L}_D = \mathbb{E}_{\tilde{\boldsymbol{x}}}[\max(0, 1 + \mathrm{D}(\tilde{\boldsymbol{x}}))] + \mathbb{E}_{\boldsymbol{x}_i}[\max(0, 1 - \mathrm{D}(\boldsymbol{x}_i))]$$

$$\mathcal{L}_{GD} = -\mathbb{E}_{\tilde{\boldsymbol{x}}}[\mathrm{D}(\tilde{\boldsymbol{x}})]$$

Catalog

Visualization Results



Methods	FID (↓)	IS (↑)
FIGR [3]	154.21	5.19
GMN [4]	201.12	6.38
DAGAN [6]	121.43	4.12
Ours	108.56	8.32

Classification in Low-data Setting

Method	Dataset	Accuracy		
		5	10	15
Standard	Omniglot	66.22	81.87	83.31
FIGR [3]	Omniglot	69.23	83.12	84.89
GMN [4]	Omniglot	67.74	84.19	85.12
DAGAN [6]	Omniglot	88.81	89.32	95.38
Ours	Omniglot	89.03	90.92	96.29
Standard	EMNIST	83.64	88.64	91.14
FIGR [3]	EMNIST	85.91	90.08	92.18
GMN [4]	EMNIST	84.56	91.21	92.09
DAGAN [6]	EMNIST	87.45	94.18	95.58
Ours	EMNIST	91.75	95.91	96.29
Standard	VGGFace	8.82	20.29	39.12
FIGR [3]	VGGFace	6.12	18.84	32.13
GMN [4]	VGGFace	5.23	15.61	35.48
DAGAN [6]	VGGFace	19.23	35.12	44.36
Ours	VGGFace	21.12	40.95	50.12

Classification in Few-shot Setting

Methods	Dataset	5-way 5-shot	10-way 5-shot
MatchingNets [11]	Omniglot	98.70	98.91
MAML [20]	Omniglot	99.90	99.13
RelationNets [21]	Omniglot	99.80	99.22
MTL [22]	Omniglot	99.85	99.35
DN4 [23]	Omniglot	99.83	99.29
Ours	Omniglot	99.93	99.42
MatchingNets [11]	VGGFace	60.01	48.67
MAML [20]	VGGFace	61.09	47.89
RelationNets [21]	VGGFace	60.93	49.12
MTL [22]	VGGFace	63.67	51.94
DN4 [23]	VGGFace	62.89	51.58
Ours	VGGFace	65.12	53.21

Ablation Study

setting	accuracy	FID (↓)	IS (↑)
$\lambda_r = 0.01$	35.62	112.16	7.89
$\lambda_r = 0.1$	40.95	108.56	8.32
$\lambda_r = 1$	33.89	107.16	9.17
$\lambda_r = 10$	30.12	106.12	11.04
$\lambda_m = 1$	40.95	108.56	8.32
$\lambda_m = 0$	28.98	111.4	7.56
matching coefficient	40.95	108.56	8.32
random coefficient	38.12	110.98	7.92
1 connection	38.67	113.21	7.09
2 connection	40.95	108.56	8.32
3 connection	34.12	106.12	9.14

IEEE International Conference on Multimedia and Expo

6-10 July 2020 || London, United Kingdom || Virtual

Thanks for watching!

MatchingGAN: Matching-Based Few-Shot Image Generation

Yan Hong, Li Niu, Jianfu Zhang, Liqing Zhang

MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY