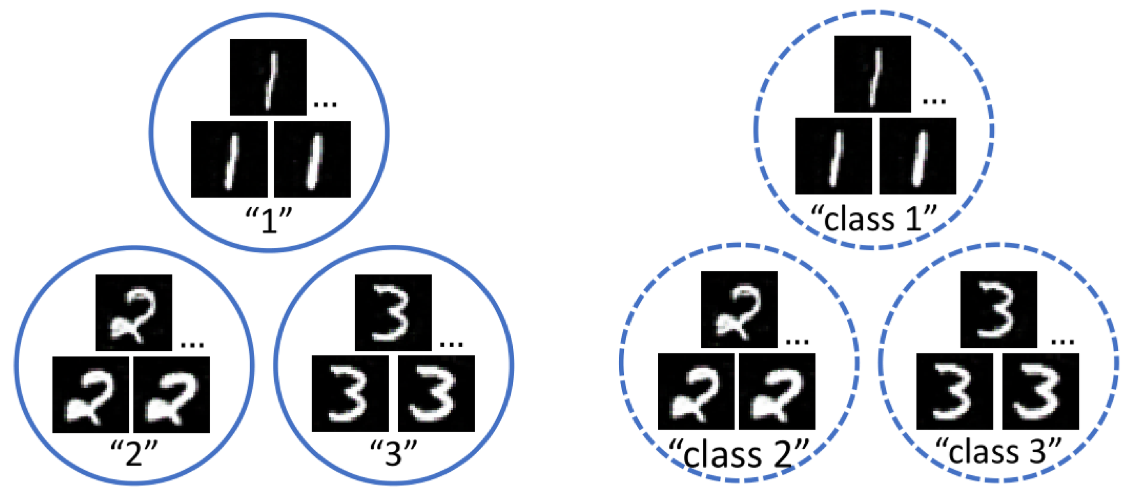
Robust Conditional GAN from Uncertainty-Aware Pairwise Comparisons

Ligong Han,[1] Ruijiang Gao,[2] Mun Kim,[1] Xin Tao,[3] Bo Liu,[4] and Dimitris Metaxas[1]
[1]Rutgers University [2]The University of Texas at Austin [3]Tencent YouTu Lab
[4]JD Finance America Corporation

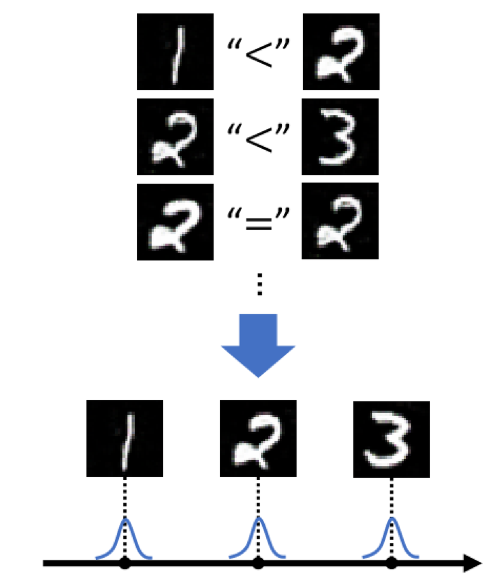
Project Website

Motivation

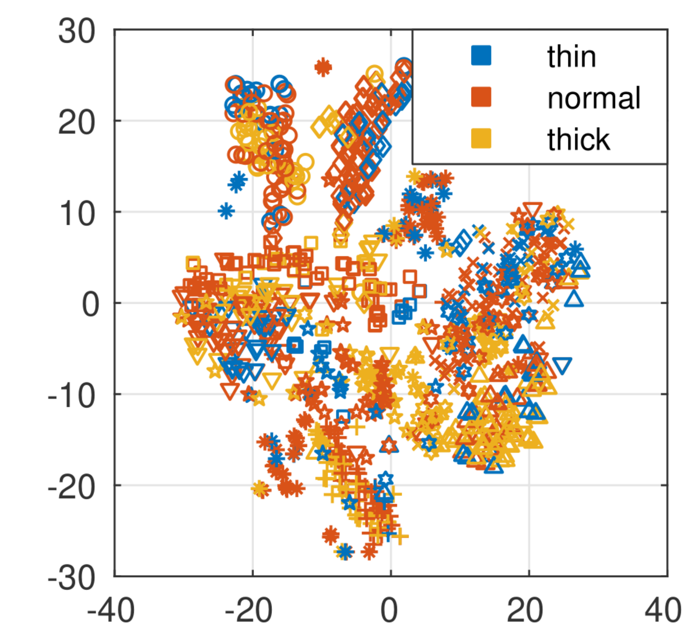
- Traditional conditional GANs require many labeled data
- Unsupervised conditional GANs leverage unsupervised or self-supervised learning methods to obtain pseudo-labels
- For continuous-valued attributes or attributes that are NOT salient, e.g. stroke thickness of MNIST digits
- We consider *weak supervisions* in the form of *pairwise comparisons*



Supervised Unsupervised



Weakly Supervised (Pairwise Comparison)

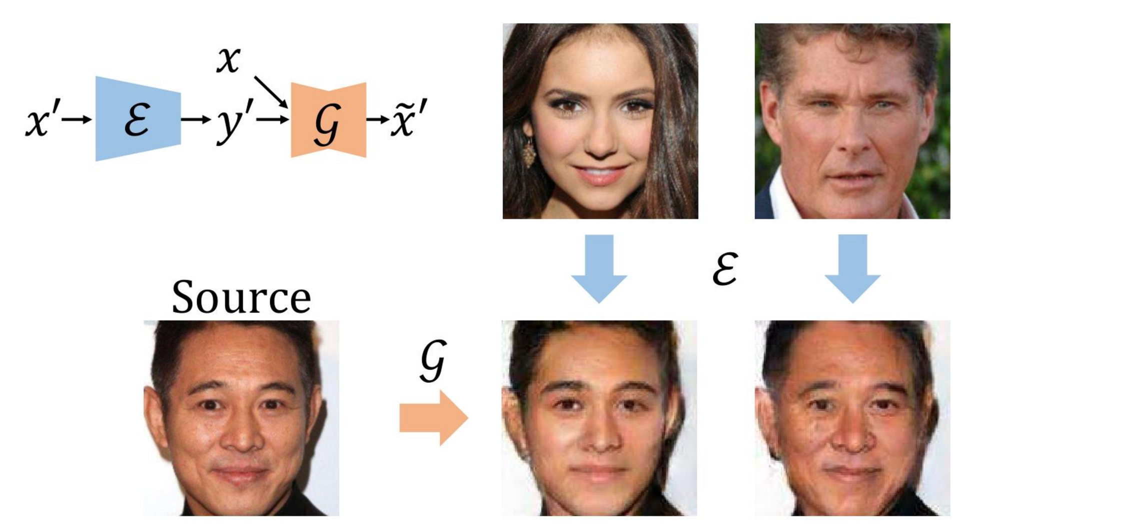


t-SNE of raw MNIST pixels, different shapes correspond to different numbers

Framework

Generative Process

- Substitute the full supervision with the attribute ratings learned from weak supervisions
- Elo rating network with uncertainty estimations: learning intrinsic scores from pairwise comparisons
- Uncertainty-aware noise-robust conditional GANs

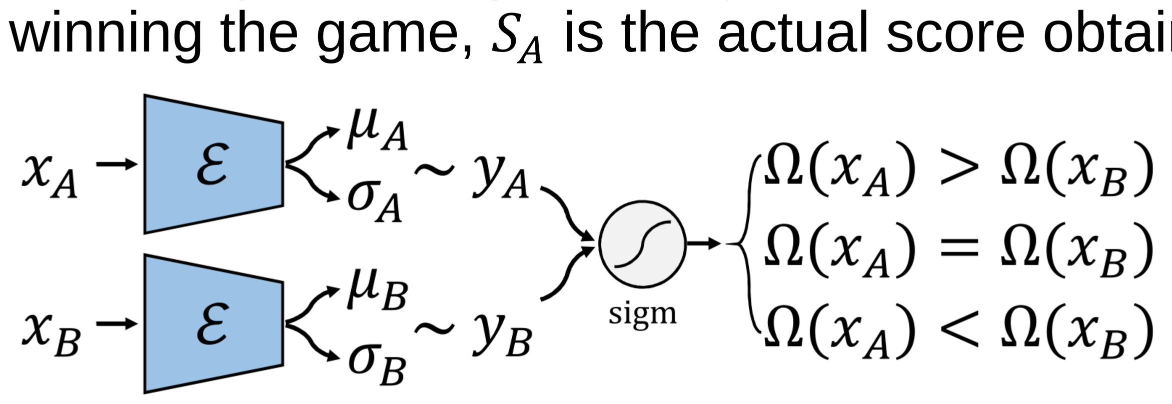


Elo Rating Network

- Elo rating system (Elo 1978)

$$P_A = \frac{1}{1 + 10^{(y_B - y_A)/400}}$$
$$y'_A = y_A + K(S_A - P_A)$$

P_A is the predicted probability of `Player A` winning the game, S_A is the actual score obtained



- Loss function

$$P_A(\Omega(x_A) > \Omega(x_B)|x_A, x_B) = \int \mathrm{sigm}(y_A - y_B) dy_A dy_B$$

$$\begin{cases} \mathcal{L}_{rank}^{MC} = -\mathbb{E}_{x_A, x_B \sim C}\left[S_A \log P_A^{MC} + S_B \log P_B^{MC}\right] \\ \mathcal{L}_{rank}^{UB} = -\mathbb{E}_{x_A, x_B \sim C}\left[\frac{1}{M}\sum_{m=1}^{M} S_A \log P_{A,y} + S_B \log P_{B,y}\right] \end{cases}$$

$$\mathcal{L}_{\mathcal{E}} = \mathcal{L}_{rank} + \underbrace{\mathrm{D}_{KL}(q_\theta(w) \| p(w|\text{data}))}_{KL}$$

- Uncertainty

$$\hat{\sigma}^2(y) \approx \frac{1}{T}\sum_{t=1}^{T}\mu_t^2 - \left(\frac{1}{T}\sum_{t=1}^{T}\mu_t\right)^2 + \frac{1}{T}\sum_{t=1}^{T}\sigma_t^2$$

Robust Conditional GAN

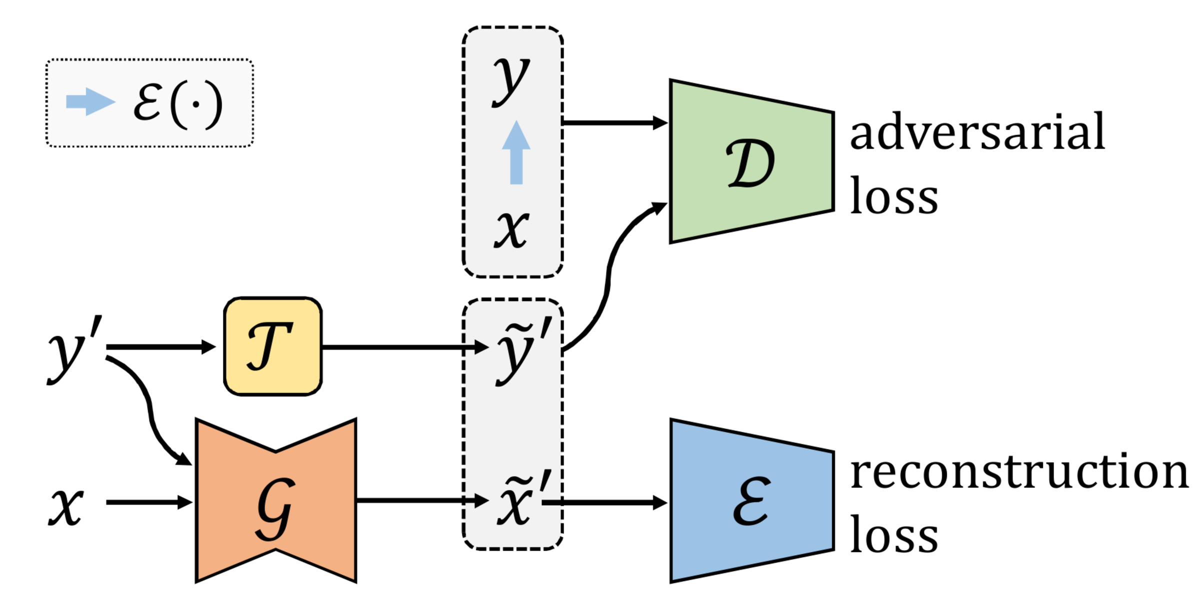
- Loss function

$$\mathcal{L}_{CGAN} = \mathbb{E}_{x,y \sim p(x,y)} \log(\mathcal{D}(x,y)) + \mathbb{E}_{x \sim p(x), y' \sim p(y'), \tilde{y}' \sim \mathcal{T}(y')} \log(1 - \mathcal{D}(\mathcal{G}(x,y'), \tilde{y}'))$$

$$\mathcal{L}_{rec}^{y} = \mathbb{E}_{x \sim p(x), y' \sim p(y')} \frac{1}{2\hat{\sigma}'^2}\|\mathcal{E}(\mathcal{G}(x,y')) - y'\|_2^2 + \frac{1}{2}\log\hat{\sigma}'^2$$

$$\mathcal{L}(\mathcal{G},\mathcal{D}) = \mathcal{L}_{CGAN} + \lambda_{rec}\mathcal{L}_{rec}^{y} + \lambda_{cyc}\mathcal{L}_{cyc}$$

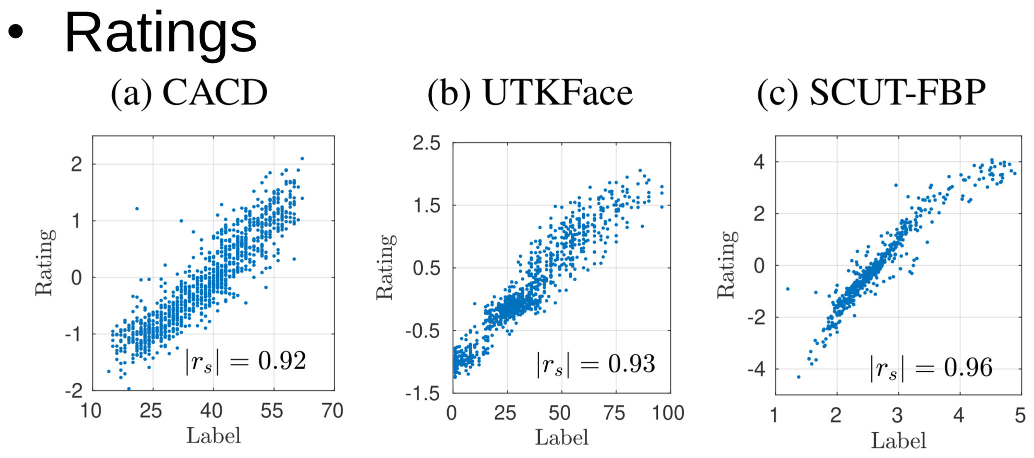
$$\mathcal{G}^* = \arg\min_{\mathcal{G}}\max_{\mathcal{D}} \mathcal{L}(\mathcal{G},\mathcal{D})$$



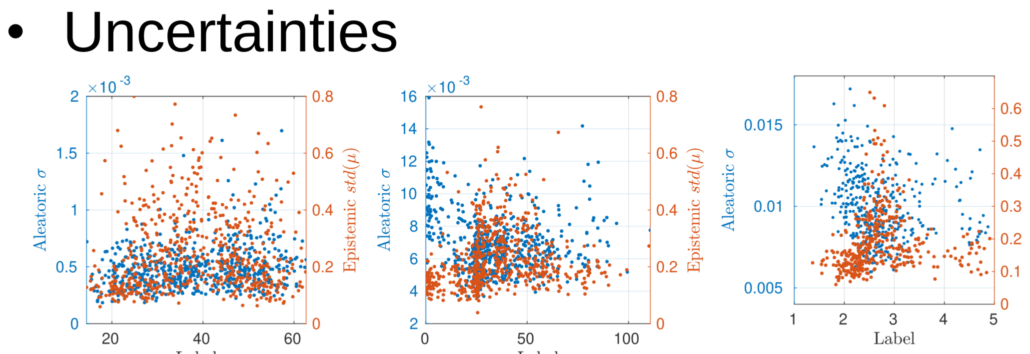
Experiment

Learning from Pairwise Comparisons

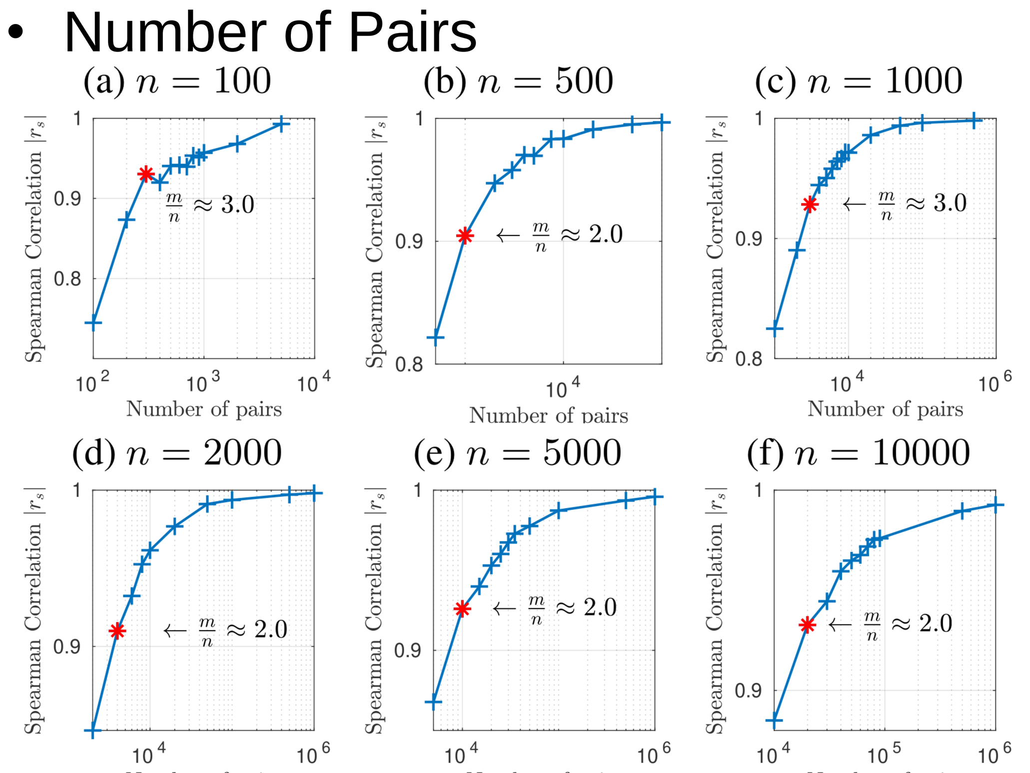
- Ratings



- Uncertainties



- Number of Pairs



Conditional Image Synthesis & Ablation Studies

- Uncertainty

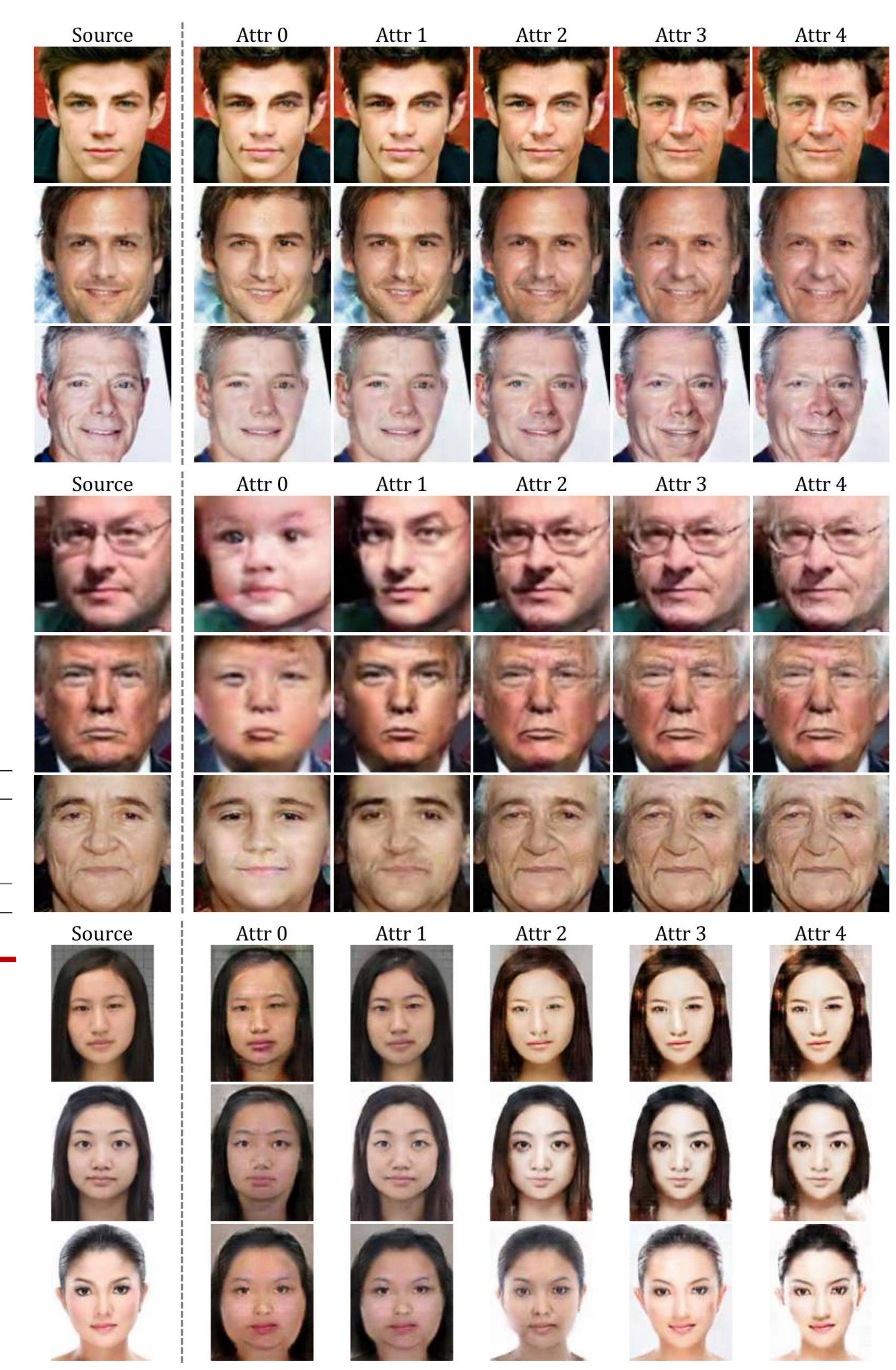
Model	CACD Acc (%)	CACD IS	CACD FID	UTKFace Acc (%)	UTKFace IS	UTKFace FID
CNN-CGAN	35.04	2.14±0.02	31.08±0.54	40.12	2.69±0.03	26.58±0.51
BNN-CGAN	37.64	2.38±0.04	27.36±0.36	38.54	2.72±0.03	26.56±0.40
BNN-RCGAN	41.02	2.45±0.03	30.22±0.51	43.64	2.84±0.04	25.25±0.39

- Loss Terms

Loss				CACD			UTKFace		
CGAN	rec	cyc	idt	Acc (%)	IS	FID	Acc (%)	IS	FID
✓	✓	✓	✓	48.08	2.87±0.04	27.90±0.44	62.74	3.50±0.04	21.63±0.52
✓	✗	✓	✓	39.50	2.93±0.04	25.68±0.46	56.90	3.38±0.05	24.98±0.88
✓	✓	✗	✓	50.86	3.10±0.04	25.93±0.55	60.56	3.39±0.05	23.70±0.65
✓	✓	✓	✗	48.60	3.05±0.03	26.81±0.59	63.92	3.60±0.05	27.65±0.75
✓	✓	✗	✗	48.98	3.01±0.03	26.90±0.67	66.34	3.65±0.04	25.39±0.86
✓	✗	✓	✗	24.28	3.06±0.04	24.01±0.66	50.42	3.02±0.04	48.80±1.70
✓	✗	✗	✓	43.86	2.94±0.05	24.27±0.58	62.42	3.54±0.04	32.87±1.47
✓	✗	✗	✗	20.08	1.59±0.02	293.03±1.40	34.88	2.16±0.04	187.98±2.17

- Supervision

		No Supervision		Full Supervision		Weak Supervision	
Dataset	Real	CycleGAN	BiGAN	Disc-CGAN	Cont-CGAN	DFI	PC-GAN
CACD	94.37(train) 49.00(val)	20.52	19.66	46.02	41.62	20.92	48.44
UTK	98.19(train) 76.80(val)	19.46	20.50	71.44	59.16	22.90	63.88
SCUT-FBP	100.00(train) 58.00(val)	19.75	20.38	29.63	46.25	22.69	40.00
Average Rank	–	5.67	5.33	2.00	2.33	4.00	**1.67**



Conclusion

- Learning weakly supervised conditional GANs from pairwise comparisons for image attribute editing
- A novel Elo rating network to learn intrinsic ratings with uncertainty estimations from pairwise comparisons
- Competitive with fully-supervised, surpassing unsupervised by a large margin