RelGAN: Multi-Domain Image-to-Image Translation via Relative Attributes Supplementary Material

Po-Wei Wu[1] Yu-Jing Lin[1] Che-Han Chang[2] Edward Y. Chang[2,3] Shih-Wei Liao[1]
[1]National Taiwan University [2]HTC Research & Healthcare [3]Stanford University
maya6282@gmail.com r06922068@ntu.edu.tw chehan_chang@htc.com echang@cs.stanford.edu liao@csie.ntu.edu.tw

1. Network Architecture

Figure 1 shows the schematic diagram of RelGAN. Table 1 and 2 show the network architecture of RelGAN.

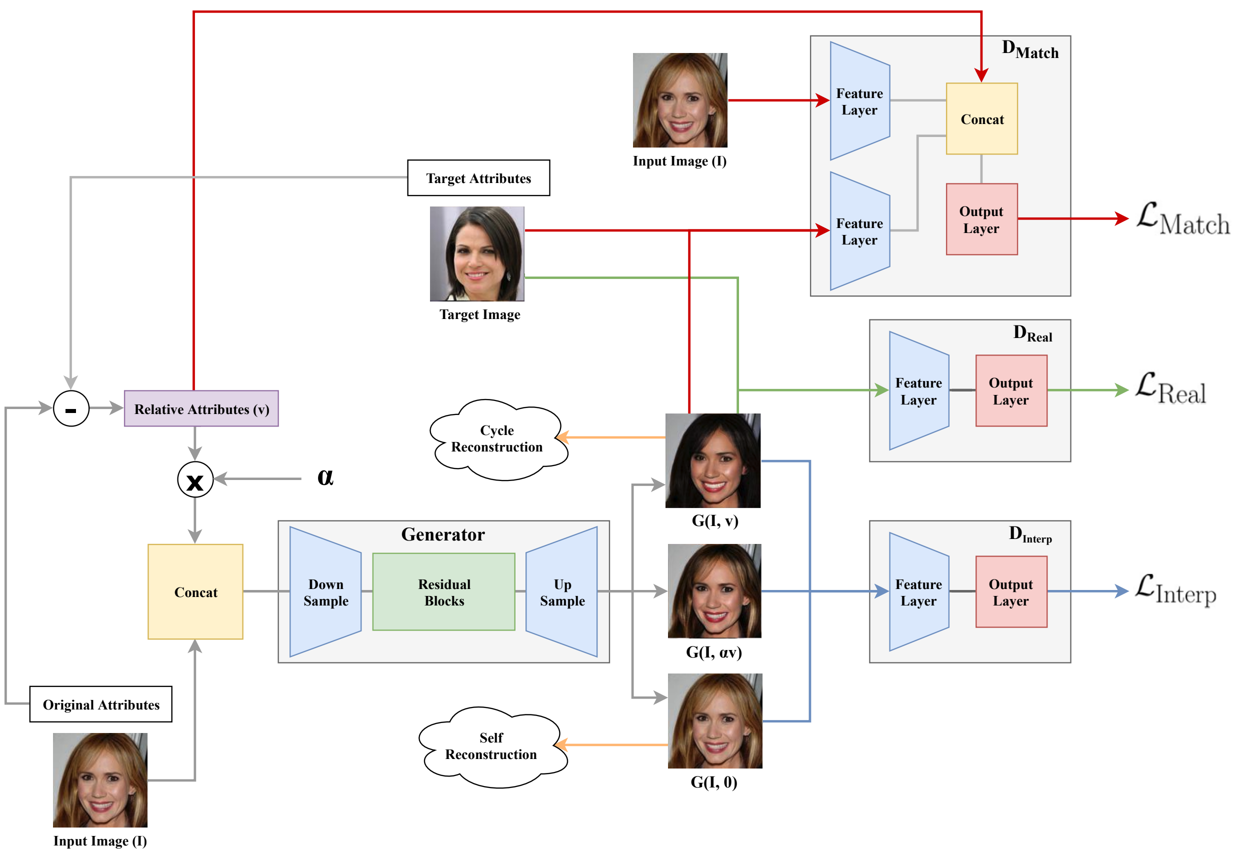


Figure 1. Detailed schematic diagram of RelGAN. D_{Real}, D_{Match}, and D_{Interp} share the weights of the feature layers.

Component	Input → Output Shape	Layer Information
Down Sample	$(h, w, 3+n) \rightarrow (h, w, 64)$	Conv-(F=64,K=7,S=1,P=3),SN,ReLU
	$(h, w, 64) \rightarrow (h, w, 128)$	Conv-(F=128,K=4,S=2,P=1),SN,ReLU
	$(h, w, 128) \rightarrow (h, w, 256)$	Conv-(F=256,K=4,S=2,P=1),SN,ReLU
Residual Blocks	$(h, w, 256) \rightarrow (h, w, 256)$	Residual Block: Conv-(F=256,K=3,S=1,P=1),SN,ReLU
	$(h, w, 256) \rightarrow (h, w, 256)$	Residual Block: Conv-(F=256,K=3,S=1,P=1),SN,ReLU
	$(h, w, 256) \rightarrow (h, w, 256)$	Residual Block: Conv-(F=256,K=3,S=1,P=1),SN,ReLU
	$(h, w, 256) \rightarrow (h, w, 256)$	Residual Block: Conv-(F=256,K=3,S=1,P=1),SN,ReLU
	$(h, w, 256) \rightarrow (h, w, 256)$	Residual Block: Conv-(F=256,K=3,S=1,P=1),SN,ReLU
	$(h, w, 256) \rightarrow (h, w, 256)$	Residual Block: Conv-(F=256,K=3,S=1,P=1),SN,ReLU
Up Sample	$(h, w, 256) \rightarrow (h, w, 128)$	Conv-(F=128,K=4,S=2,P=1),SN,ReLU
	$(h, w, 128) \rightarrow (h, w, 64)$	Conv-(F=64,K=4,S=2,P=1),SN,ReLU
	$(h, w, 64) \rightarrow (h, w, 3)$	Conv-(F=3,K=7,S=1,P=3),Tanh

Table 1. **Generator network architecture.** We use switchable normalization, denoted as SN, in all layers except the last output layer. n is the number of attributes. F is the number of filters. K is the filter size. S is the stride size. P is the padding size. The number of trainable parameters is about 8M.

Component	Input → Output Shape	Layer Information
Feature Layers	$(h, w, 3) \rightarrow (h/2, w/2, 64)$	Conv-(F=64,K=4,S=2,P=1),Leaky ReLU
	$(h/2, w/2, 64) \rightarrow (h/4, w/4, 128)$	Conv-(F=128,K=4,S=2,P=1),Leaky ReLU
	$(h/4, w/4, 128) \rightarrow (h/8, w/8, 256)$	Conv-(F=256,K=4,S=2,P=1),Leaky ReLU
	$(h/8, w/8, 256) \rightarrow (h/16, w/16, 512)$	Conv-(F=512,K=4,S=2,P=1),Leaky ReLU
	$(h/16, w/16, 512) \rightarrow (h/16, w/16, 1024)$	Conv-(F=1024,K=4,S=2,P=1),Leaky ReLU
	$(h/32, w/32, 1024) \rightarrow (h/64, w/64, 2048)$	Conv-(F=2048,K=4,S=2,P=1),Leaky ReLU
D_{Real}: Output Layer	$(h/64, w/64, 2048) \rightarrow (h/64, w/64, 1)$	Conv-(F=1,K=1)
D_{Interp}: Output Layer	$(h/64, w/64, 2048) \rightarrow (h/64, w/64, 64)$	Conv-(F=64,K=1)
	$(h/64, w/64, 64) \rightarrow (h/64, w/64, 1)$	Mean(axis=3)
D_{Match}: Output Layer	$(h/64, w/64, 4096+n) \rightarrow (h/64, w/64, 2048)$	Conv-(F=1,K=1),Leaky ReLU
	$(h/64, w/64, 2048) \rightarrow (h/64, w/64, 1)$	Conv-(F=1,K=1)

Table 2. **Discriminator network architecture.** We use Leaky ReLU with a negative slope of 0.01. n is the number of attributes. F is the number of filters. K is the filter size. S is the stride size. P is the padding size. The number of trainable parameters is about 53M.

2. Additional Results

Figure 2 and 3 show comparison results between StarGAN, AttGAN, and RelGAN on the hair color tasks. The residual heat maps show that RelGAN preserves the smile attribute while the other methods strengthen the smile attribute. Figure 4 and 5 show more comparison results. Figure 6, 7, 8, 9, 10, 11, 12, and 13 show additional results on facial attribute transfer. Figure 14, 15, 16, and 17 show additional results on facial attribute interpolation. All the input images are from the CelebA-HQ dataset.

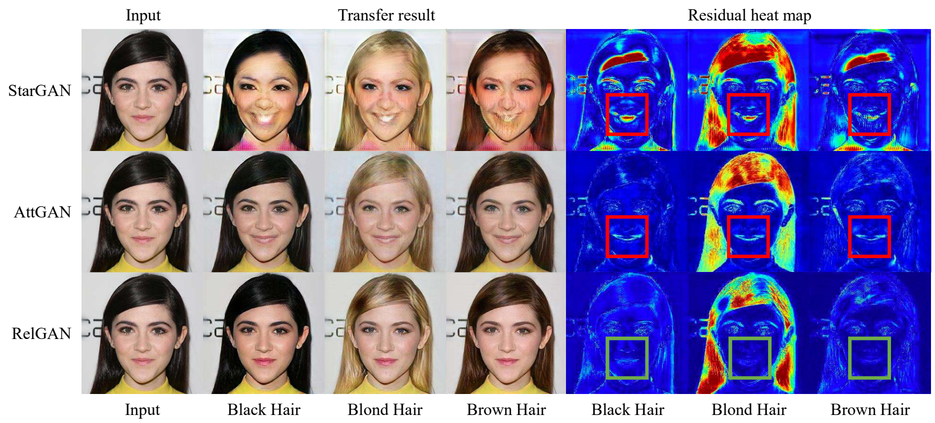


Figure 2. Hair color transfer results of StarGAN, AttGAN, and RelGAN. The residual heat maps visualize the differences between the input and the output images.

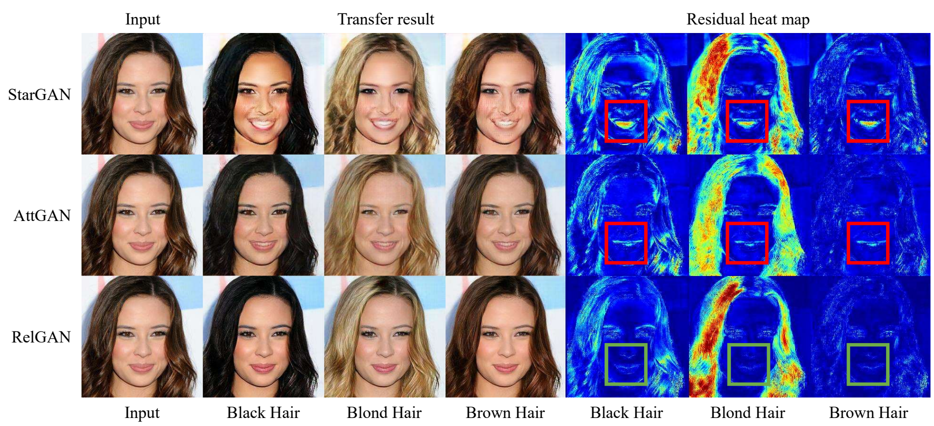


Figure 3. Hair color transfer results of StarGAN, AttGAN, and RelGAN. The residual heat maps visualize the differences between the input and the output images.

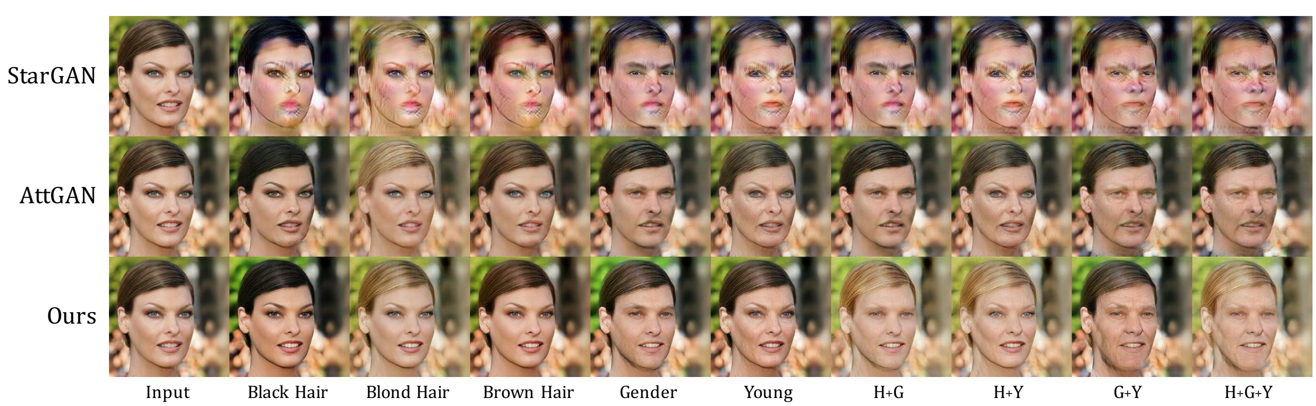


Figure 4. Facial attribute transfer results of StarGAN, AttGAN, and RelGAN. Please zoom in for more details. In the case of changing hair color, RelGAN preserves the smile attribute, while both StarGAN and AttGAN make the woman look unhappy due to their target-attribute-based formulation.

Figure 5. Facial attribute transfer results of StarGAN, AttGAN, and RelGAN. Please zoom in for more details.

Figure 6. Single attribute transfer results. From left to right: input, black hair, blond hair, brown hair, gender, and mustache.

Figure 7. Single attribute transfer results. From left to right: input, black hair, blond hair, brown hair, gender, and mustache.

Figure 8. Single attribute transfer results. From left to right: input, pale skin, smiling, bangs, glasses, and age.

Figure 9. Single attribute transfer results. From left to right: input, pale skin, smiling, bangs, glasses, and age.

Figure 10. Single attribute transfer results. Form left to right: input, gender, and mustache.

Figure 11. Single attribute transfer results. Form left to right: input, gender, and mustache.

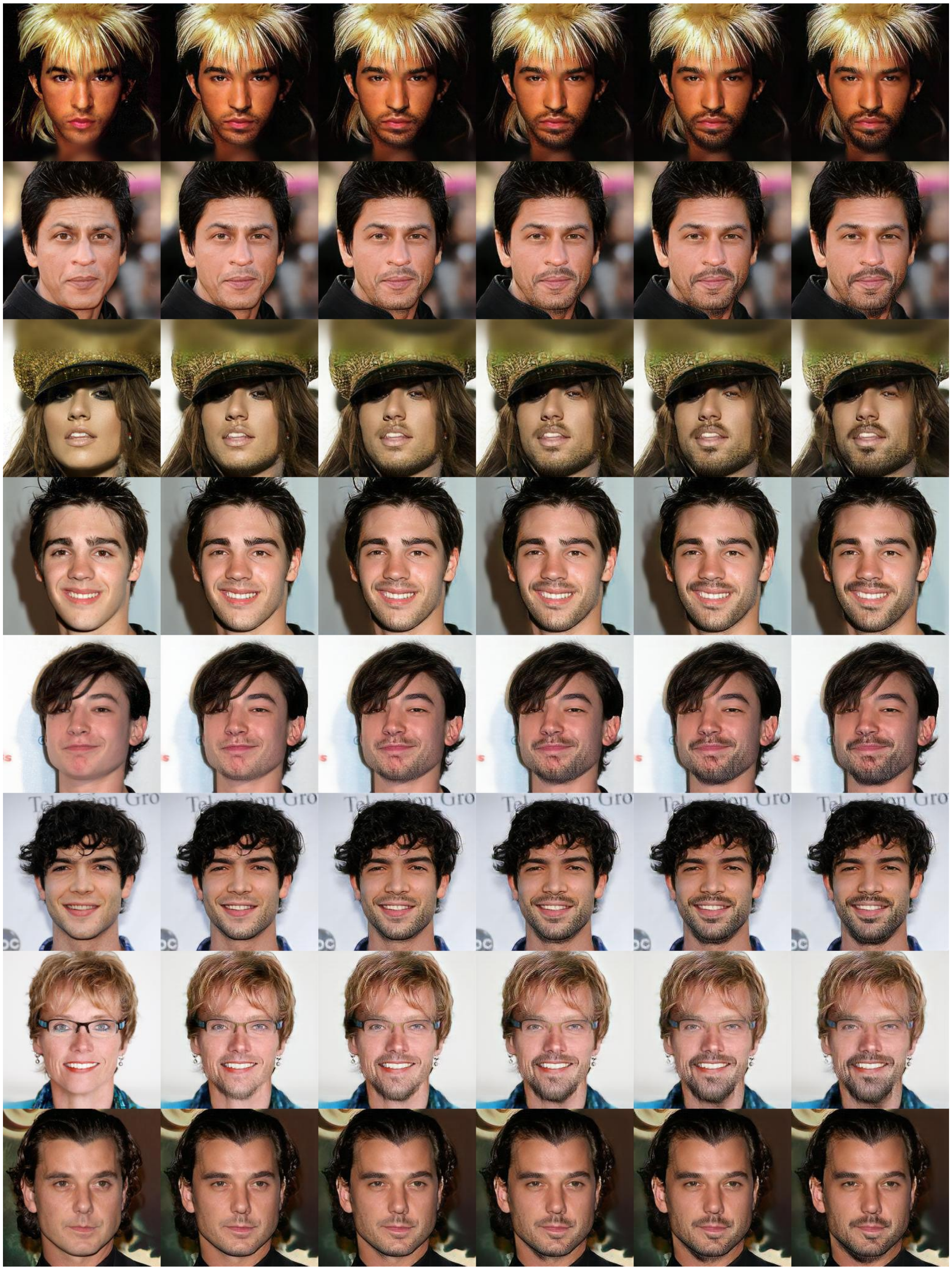
Figure 12. Single attribute transfer results. From left to right: input, glasses, and age.

Figure 13. Single attribute transfer results. From left to right: input, glasses, and age.

Figure 14. Single attribute interpolation results (age). RelGAN generates different levels of attribute transfer by varying α.

$\alpha = 0$ $\alpha = 0.4$ $\alpha = 0.55$ $\alpha = 0.7$ $\alpha = 0.88$ $\alpha = 1.0$

Figure 15. Single attribute interpolation results (gender). RelGAN generates different levels of attribute transfer by varying α.



$\alpha = 0$ $\alpha = 0.4$ $\alpha = 0.55$ $\alpha = 0.7$ $\alpha = 0.88$ $\alpha = 1.0$

Figure 16. Single attribute interpolation results (mustache). RelGAN generates different levels of attribute transfer by varying α.

$\alpha = 0$ $\alpha = 0.6$ $\alpha = 0.7$ $\alpha = 0.8$ $\alpha = 0.9$ $\alpha = 1.0$

Figure 17. Single attribute interpolation results (smile). RelGAN generates different levels of attribute transfer by varying α.