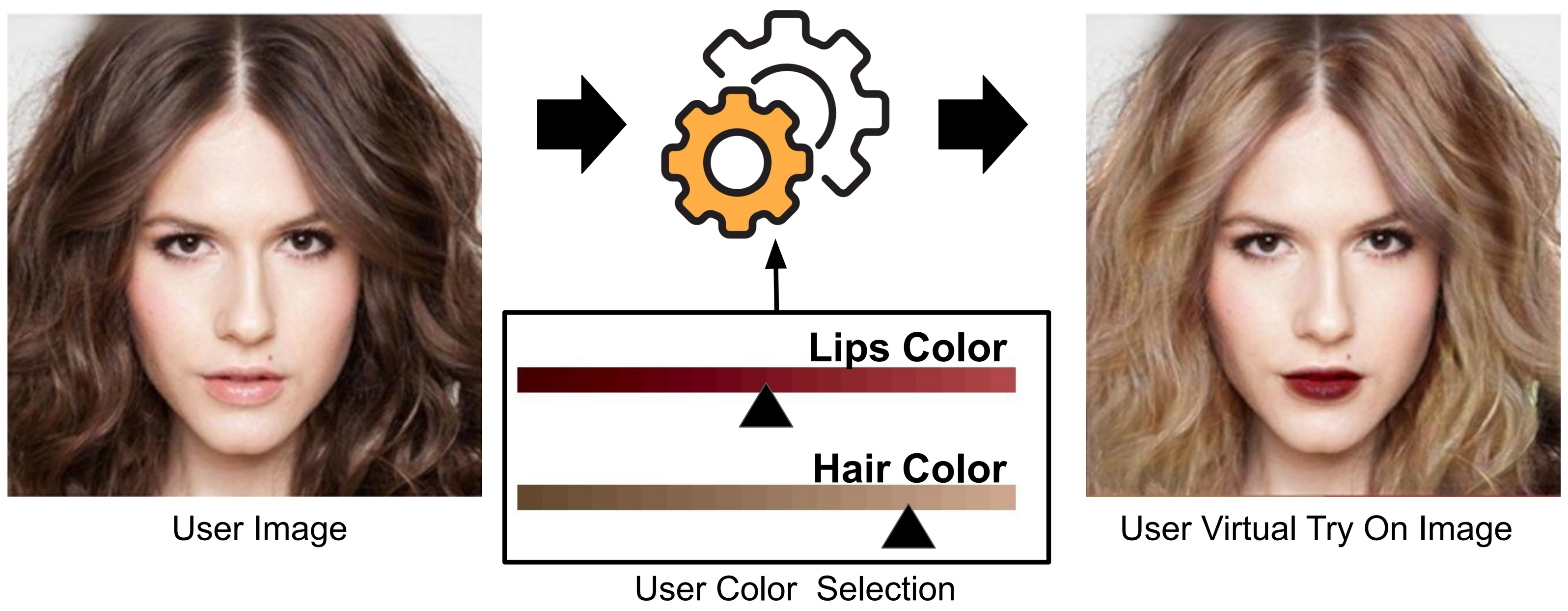
Weak Segmentation Guided GAN for Color Edition

Vincent Auriau, Emmanuel Malherbe, Matthieu Perrot

L'ORÉAL Recherche & Innovation

ICIAP 2023 UDINE

Introduction



User Image — User Color Selection — User Virtual Try On Image

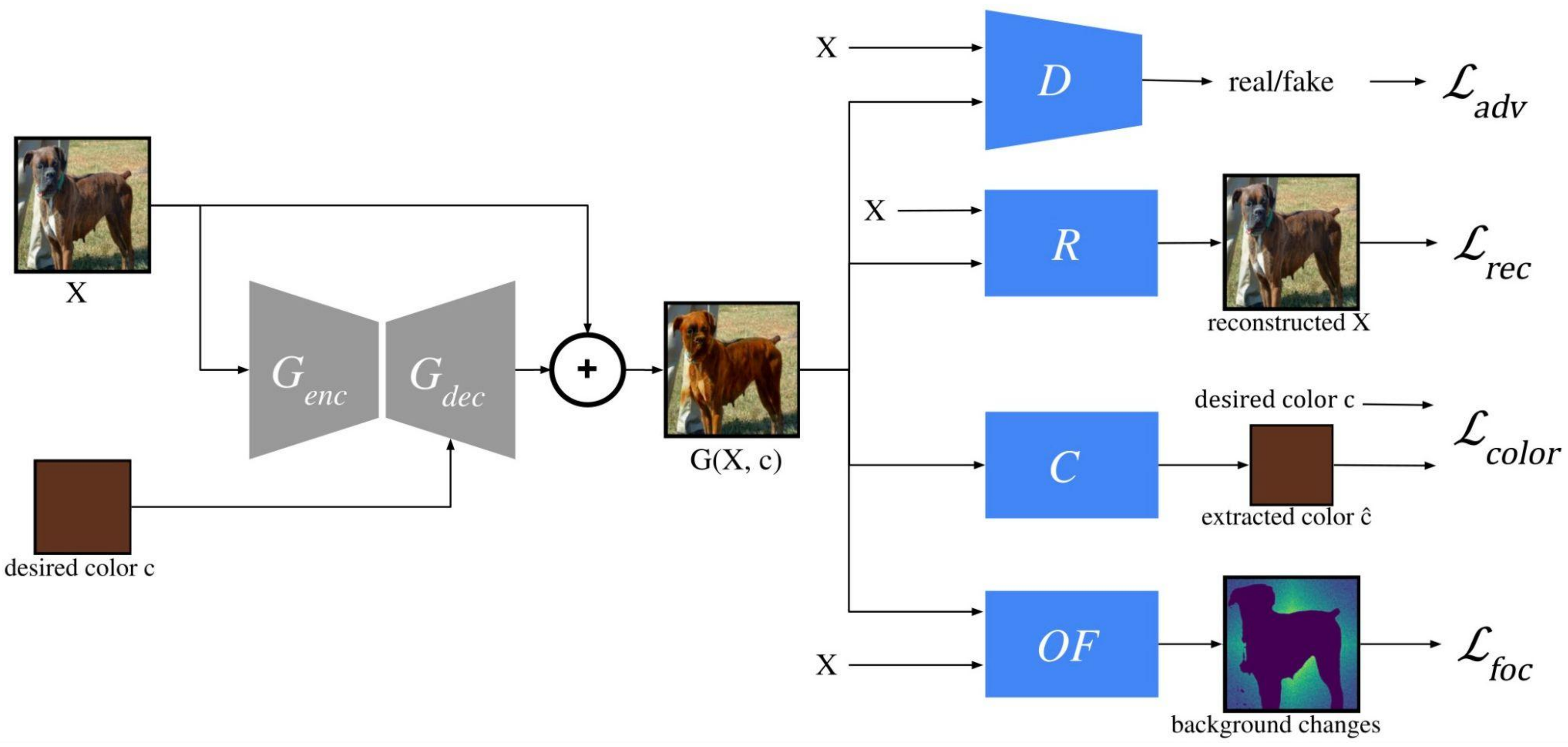
We present a model for precise color edition. From a color picked by a user, we aim at editing color accordingly on a selected object. A key application is Virtual Try On, where a customer can choose the exact hair or lips color before any product purchase.

Related Work

Required Key Property	Alpha Blending Segmentation & color shift [4]	StarGAN [1]	StyleGAN [2]	Ours
Realism: The generated image is realistic to the human eye	✗	✓	✓	✓
Stability: The geometry of the considered object and local color variations are kept unchanged	✓	✗	✓	✓
Colorization accuracy: The perceived resulting color of the object is as close as possible as the desired one	✓	✗	✗	✓
Restraint: Objects and areas that should not be modified, or "background", are left unchanged	✓	✗	✗	✓

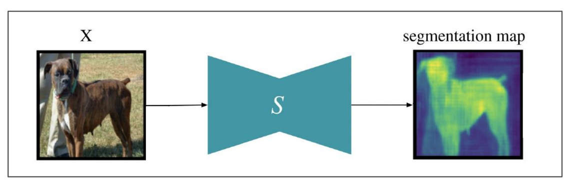
Proposed model: Weak Segmentation Guided Generative Adversarial Networks

Our framework consists of a main model, *G*, which is an encoder-decoder CNN. *G* is trained through four different modules: *D, R, C* and *OF*. These modules let *G* follow four key properties we have defined.



Our global training objective is:
$$\begin{cases} \mathcal{L}_G = \mathcal{L}_{adv} + \lambda_{color}\mathcal{L}_{color} + \lambda_{foc}\mathcal{L}_{foc} + \lambda_{rec}\mathcal{L}_{rec} \\ \mathcal{L}_D = -\mathcal{L}_{adv} \end{cases}$$

- Only requires a segmentation function ***S*** (e.g. CNN model)
- No label is needed



Realism

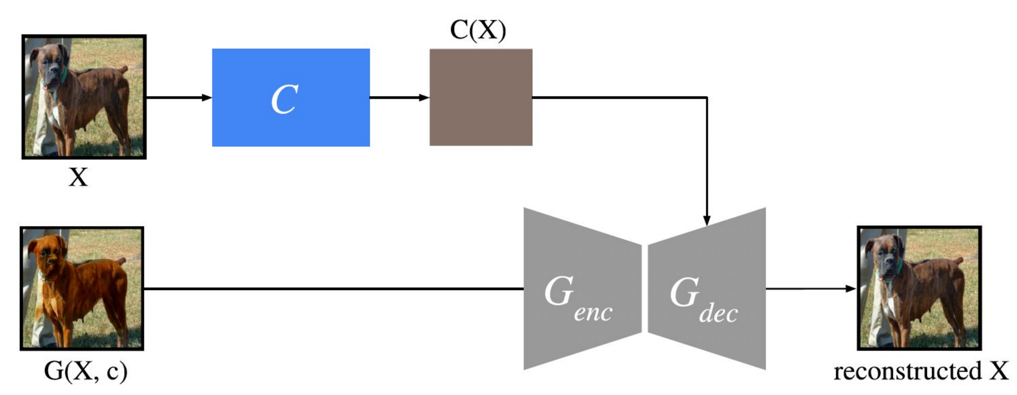
Discriminator:

CNN trained with gradient-penalty Wasserstein loss

$$\mathcal{L}_{adv} = D(\hat{X}) - D(X) + \lambda_{gp}\nabla_{\tilde{X}} D(\tilde{X})$$

Stability

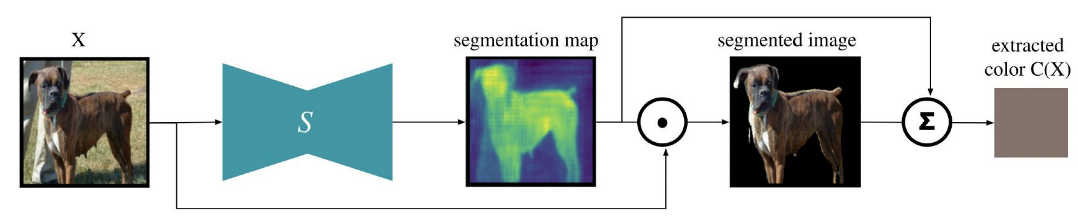
Reconstruction Module:



$$\mathcal{L}_{rec} = ||X - G(G(X,c), C(X))||^2$$

Colorization Accuracy

Color Estimator:

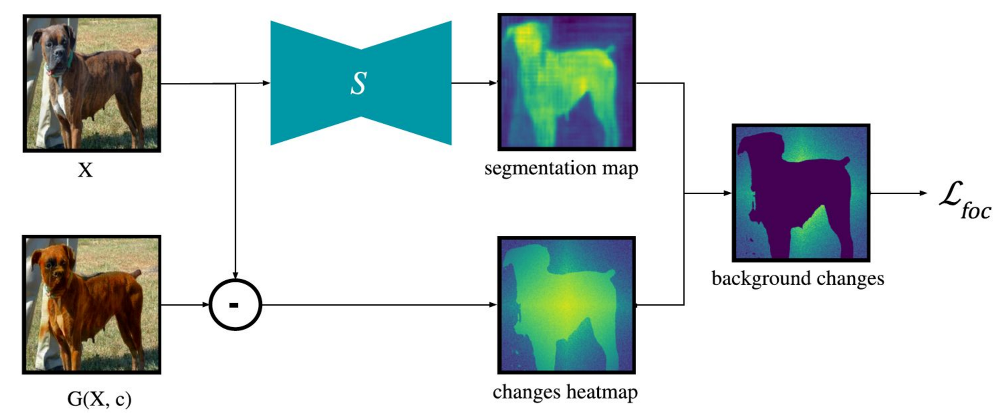


$$C(X) = \mathbb{E}_{object}[X] = \frac{\sum_{i,j} S(X)_{i,j} * x_{i,j}}{\sum_{i,j} S(X)_{i,j}}$$

$$\mathcal{L}_{color} = ||C(\hat{X}) - c||_{L^*a^*b^*_{76}}$$

Restraint

Object Focus Module:



$$\mathcal{L}_{foc} = ||(X - \hat{X})) \odot (\mathbf{1} - S(X))||^2$$

Experiments

Evaluation Protocol:

We select $\mathcal{X}_1$ and $\mathcal{X}_2$ two random subsets of our dataset consisting of 10k images kept for testing. We build the dataset $\mathcal{X}_{1\to 2}$ composed of the generations $G(X_1^i, C(X_2^i))$ for each $(X_1^i, X_2^i) \in (\mathcal{X}_1, \mathcal{X}_2)$. We compute the following metrics:

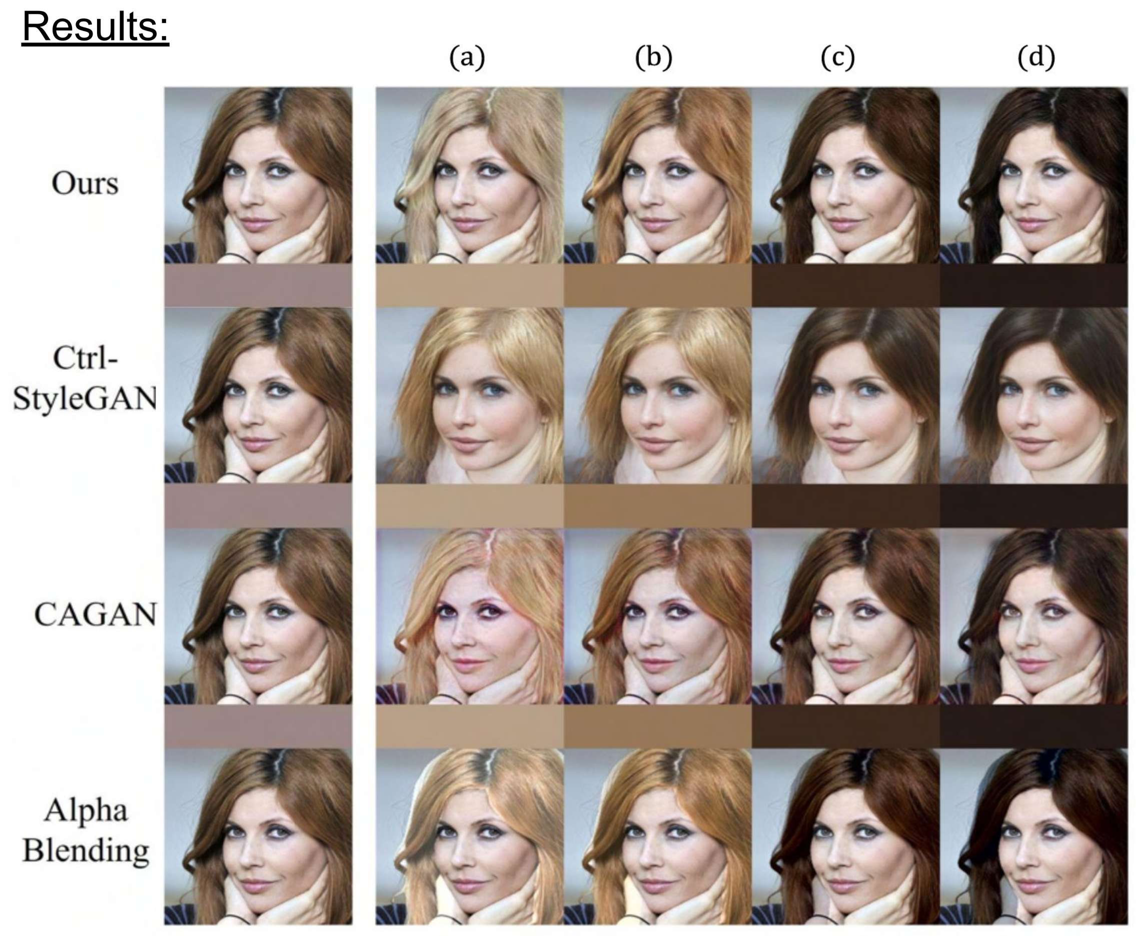
- $FID_{1\to 2} = FID(\mathcal{X}_{1\to 2}, \mathcal{X}_2)$ with *FID* the Fréchet Inception Distance to account for realism
- $L_{1\to 2}^{color} = \frac{1}{n}\sum_{i=1}^{n} ||C(X_{1\to 2}^i) - C(X_2^i)||_{L^*a^*b^*_{76}}$ to account for color precision
- $L_{bg}^{avg} = \frac{1}{3n}\sum_{i=1}^{n} \frac{||(X_{1\to 2}^i - X_1^i) \odot (\mathbf{1} - M_1^i)||_1}{||\mathbf{1} - M_1^i||_1}$ with M_1^i the ground truth mask associated to X_1^i

The results presented are computed on hair with the CelabA-HQ open dataset.

Compared approaches:

- **Alpha Blending [1]**, pairing segmentation and RGB shift.
- **CA-GAN [3]**, a straightforward adjustment of StarGAN.
- An adaptation of **StyleGAN [3]** for color edition.

Results:



Original — Color variations controlled by indicated RGB value

Analysis:

Our model shows best balance between color precision and realism

Identity is altered due to projection

Face and background colors are also edited

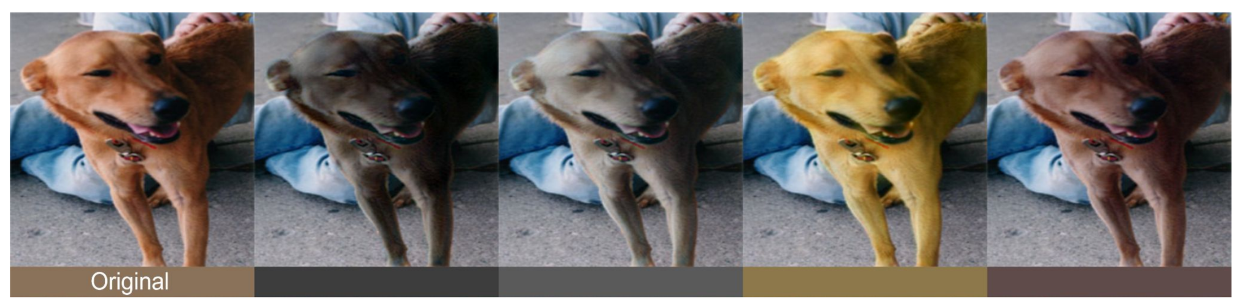
Artifacts are created on the edges of the segmentations mask

Model	$L_{bg}^{avg} \times 10^3$	$FID_{1\to 2}$	$L_{1\to 2}^{color}$	User % Most Realistic
Ours	0.92	0.050	5.03	61%
Ctrl- StyleGAN	9.51	0.670	17.2	0%
CAGAN	121	0.096	11.3	0%
Alpha Blending	0.00	1.301	0.00	34%

Extension to other usecases

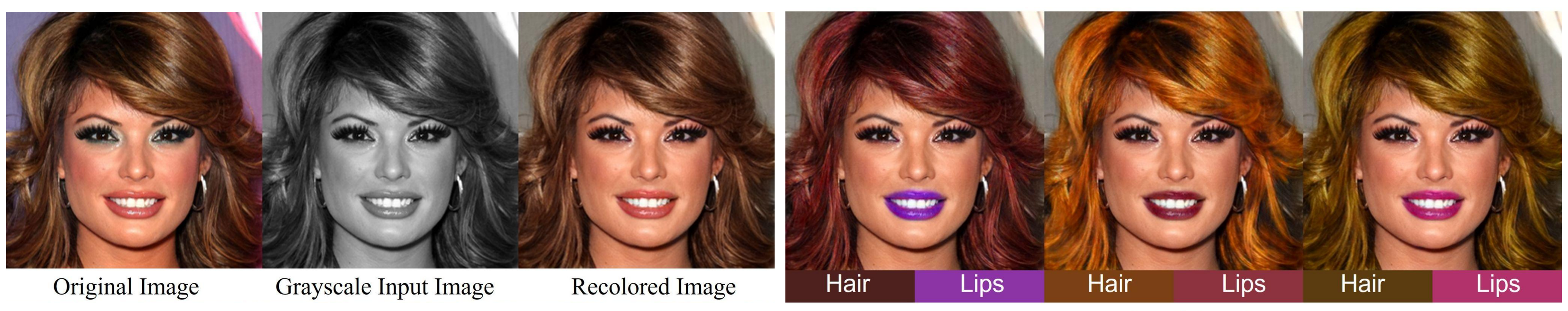
Our framework can be easily generalized to any dataset, and also handles several object colorizations at once.

Cats & Dogs Dataset



Garments Dataset

Grey scale images Colorization:



Original Image — Grayscale Input Image — Recolored Image — Other Colorations

Conclusion

Our new object-aware colorization model based on segmentation, introduces specific losses to improve the realism and the color precision of the results. It can be used on any type of image of objects and **only needs a pretrained and differentiable segmentation** model to work. We have shown **state of the art results** on several datasets, with the best balance between **realism** of generated images, **accurate color** of transformed object as well as **maintained geometry and background**.

References

[1] Choi, Y; ., M., Kim, M., Ha, J., Kim, S., Choo, J.: StarGAN Unified generative adversarial networks for multi-domain image-to-image translation

[2] Karras, T., Laine, S., Aila, T.: A style-based generator architecture for generative adversarial networks. In: IEEE CVPR 2019

[3] Kips, R., Gori, P., Perrot, .M., Bloch, I.: CA-GAN: weakly supervised color aware GAN for controllable makeup transfer

[4] Reinhard, E., Ashikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. IEEE Computer Graphics and Applications