Painting as your like: Colorization and Neural Style Transfer

Changjie Lu 1129503

3700 Final Presentation

2021.12.5

Introduction

- Image colorization and Neural Style Transfer are awesome!

Examples

Colorization and Neural Style Transfer

Convolutional Neural Network

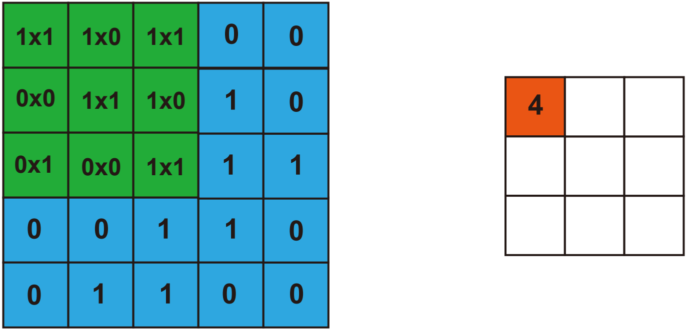

Figure: The moving window(kernal): Green, original matrix: Blue, output: Orange

Introduction

- LAB[Papadakis et al., 2000, Segnini et al., 1999, Yam and Papadakis, 2004]
  L* represents Lightness, a* represents the color from green to red, and b* represents the color from blue to yellow.
  LAB knows the **texture**.
- RGB
  Red, Green and Blue for human perception.
- LAB<->RGB[Leon et al., 2006]

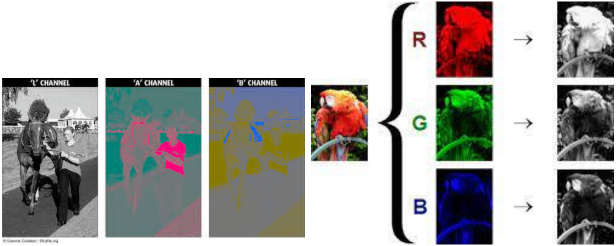

Figure: https://commons.wikimedia.org/wiki/File:RGB$_{c}hannels_{s}eparation.png$
http://shutha.org/node/851

Related Work

- Colorization
  Colorful Image Colorization[Zhang et al., 2016]
  User-guided Colorization[Zhang et al., 2017]
  ChromaGAN [Vitoria et al., 2020]
  Instance-aware[Su et al., 2020]
- Neural Style Transfer
  Domain-Aware [Hong et al., 2021] Adaptive attention mechanism[Liu et al., 2021]

Colorization Network[Su et al., 2020]

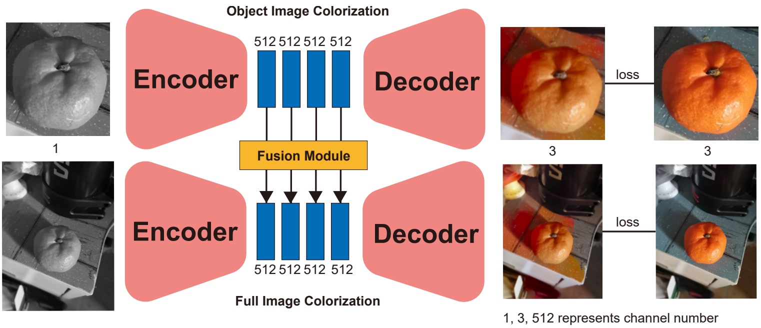

Figure: Model architecture for Instance-aware colorization network[Su et al., 2020]

Visual Comparison

(a) gray (b) cGAN (c) CIC (d) Inst (e) CIC2 (f) gt

Quantitative Comparison

	ChromaGAN		Inst		CIC2		CIC	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
NCD	19.90	0.926	22.45	0.924	21.49	0.908	19.52	0.905
COCO	25.04	0.951	23.52	0.875	23.36	0.870	21.97	0.857
Imagenet1k	25.30	0.950	22.88	0.876	23.48	0.870	21.95	0.857
VOC	25.77	0.955	24.16	0.893	24.49	0.887	22.96	0.876

Table: Quantitative Comparison on the image evaluation index of PSNR↑ and SSIM↑. We do this comparison on four datasets, containing over 10k image.[Anwar et al., 2020, Lin et al., 2014, Everingham et al., 2010, Deng et al., 2009],[Vitoria et al., 2020, Su et al., 2020, Zhang et al., 2017, Zhang et al., 2016]

PSNR and SSIM

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2 \tag{1}$$

$$\begin{aligned} PSNR &= 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \\ &= 20 \cdot \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \\ &= 20 \cdot \log_{10}(MAX_I) - 10 \cdot \log_{10}(MSE) \end{aligned} \tag{2}$$

where m,n is the weight and weight of the image.

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{3}$$

where x,y are two moving window. $c_1 = (k_1L)^2$, $c_2 = (k_2L)^2$ $c_2 = (k_2L)^2$ two variables to stabilize the division with weak denominator; L is the dynamic range of the pixel-values;$k_1 = 0.01$ and $k_2 = 0.03$ $k_2 = 0.03$ by default.

Strength and Weakness

- Strength
  Cleverly introduces the Mask-RCNN to detect objects in pictures, which solves the poor color performance of instances in the past.
  Proposes the fusion module in the fusion of object and background.
- Weakness
  Simple design of the loss function(only smooth l1 loss)
  The low efficiency of the fusion module

Neural Style Transfer Network[Liu et al., 2021]

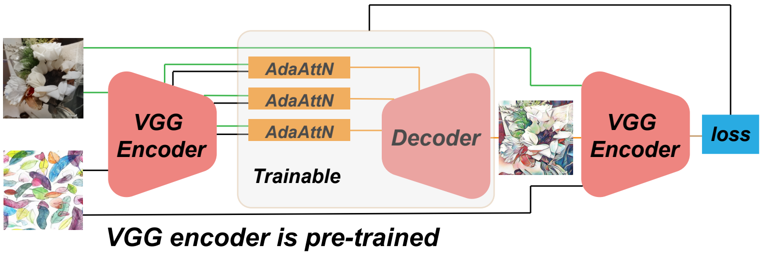

Figure: Model architecture for attention neural style transfer network[Liu et al., 2021]

Visual Comparison

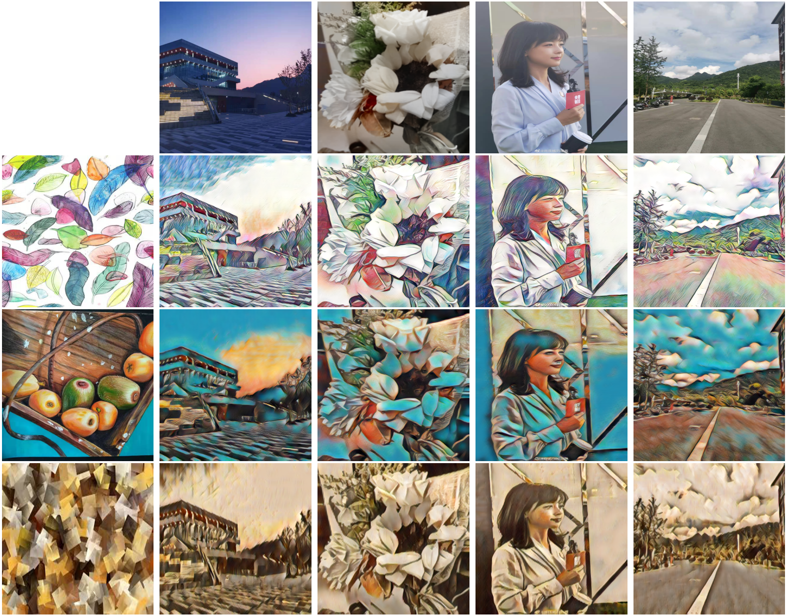

Figure: Visual results of Neural Style Transfer. The images at the left column are style including color pencil, oil painting, cubism, impressionism, sketch and expressionism. The image at the first row are content. The other image are style transfer image.

Strength and Weakness

- Strength
  This paper introduces an adaptive attention mechanism network using multi-scale feature outputs.
- Weakness
  In terms of the network not having panoramic perception, the network is deficient in learning spatial information of images, especially those with a particular color gradient in the background.

Conclusion

- Study the algorithm of image coloring and style transfer.
- Autoencoder and decoder construction
- The choice of latent space
- Adaptive attention mechanisms
- Semantic segmentation can solve the weakness of both paper.

References I

Anwar, S., Tahir, M., Li, C., Mian, A., Khan, F. S., and Muzaffar, A. W. (2020).
Image colorization: A survey and dataset.
arXiv preprint arXiv:2008.10774.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009).
Imagenet: A large-scale hierarchical image database.
In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee.

Everingham, M., Van Gool, L., Williams, C. K., Winn, J., and Zisserman, A. (2010).
The pascal visual object classes (voc) challenge.
International journal of computer vision, 88(2):303–338.

Hong, K., Jeon, S., Yang, H., Fu, J., and Byun, H. (2021).
Domain-aware universal style transfer.
In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 14609–14617.

Leon, K., Mery, D., Pedreschi, F., and Leon, J. (2006).
Color measurement in l a b units from rgb digital images.
Food research international, 39(10):1084–1091.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L. (2014).
Microsoft coco: Common objects in context.
In *European conference on computer vision*, pages 740–755. Springer.

Liu, S., Lin, T., He, D., Li, F., Wang, M., Li, X., Sun, Z., Li, Q., and Ding, E. (2021).
Adaattn: Revisit attention mechanism in arbitrary neural style transfer.
In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6649–6658.

References II

Papadakis, S. E., Abdul-Malek, S., Kamdem, R. E., and Yam, K. L. (2000).
A versatile and inexpensive technique for measuring color of foods.
Food technology (Chicago), 54(12):48–51.

Segnini, S., Dejmek, P., and Öste, R. (1999).
A low cost video technique for colour measurement of potato chips.
LWT-Food Science and Technology, 32(4):216–222.

Su, J.-W., Chu, H.-K., and Huang, J.-B. (2020).
Instance-aware image colorization.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7968–7977.

Vitoria, P., Raad, L., and Ballester, C. (2020).
Chromagan: Adversarial picture colorization with semantic class distribution.
In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2445–2454.

Yam, K. L. and Papadakis, S. E. (2004).
A simple digital imaging method for measuring and analyzing color of food surfaces.
Journal of food engineering, 61(1):137–142.

Zhang, R., Isola, P., and Efros, A. A. (2016).
Colorful image colorization.
In *European conference on computer vision*, pages 649–666. Springer.

Zhang, R., Zhu, J.-Y., Isola, P., Geng, X., Lin, A. S., Yu, T., and Efros, A. A. (2017).
Real-time user-guided image colorization with learned deep priors.
arXiv preprint arXiv:1705.02999.