CRIPAC
智能感知与计算研究中心
Center for Research on Intelligent Perception and Computing

ANHUI UNIVERSITY 1928

Attentional Wavelet Network for Traditional Chinese Paintings Transfer

Rui Wang
Institute of Automation, Chinese Academy of Sciences
Anhui University

December 8th, 2020

Background

Motivation

Our work

Conclusion

Background

Style Transfer Task:

- **Domain transfer**
  - Pix2pix [1]
  - CycleGAN [2]
  - ......
- **Arbitrary style transfer**
  - Optimization-based [3]
  - AdaIN [4]
  - WCT [5]
  - ......

1. Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros, "Image-to-Image Translation with Conditional Adversarial Nets", CVPR2017.
2. Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", ICCV2017.
3. Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, "Image Style Transfer Using Convolutional Neural Networks", CVPR2016.
4. Xun Huang, Serge Belongie, "Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization", ICCV2017.
5. Yijun Li, Chen Fang, Jimei Yang, Zhaowen Wang, Xin Lu, Ming-Hsuan Yang, "Universal Style Transfer via Feature Transforms", NIPS2017.

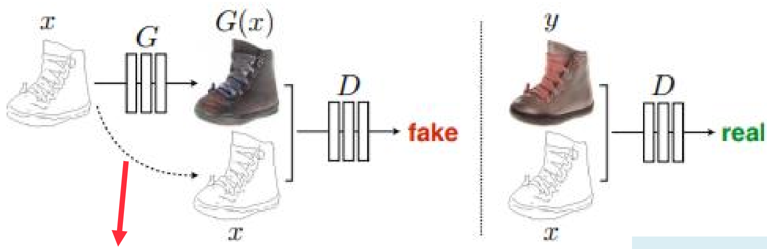
Style Transfer Task:

- **Domain transfer**
  - Pix2pix
  - CycleGAN
  - ......
- **Arbitrary style transfer**
  - Optimization-based
  - AdaIN
  - WCT
  - ......

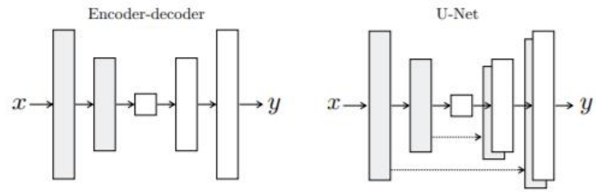
Background

Domain transfer

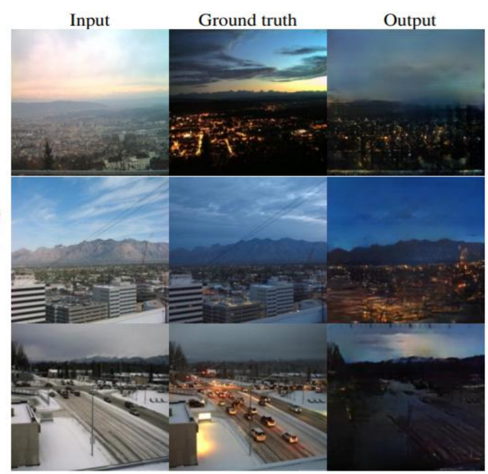
Pix2pix



L1



structure



paired

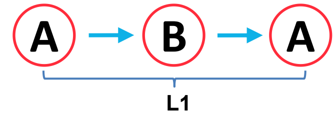
results

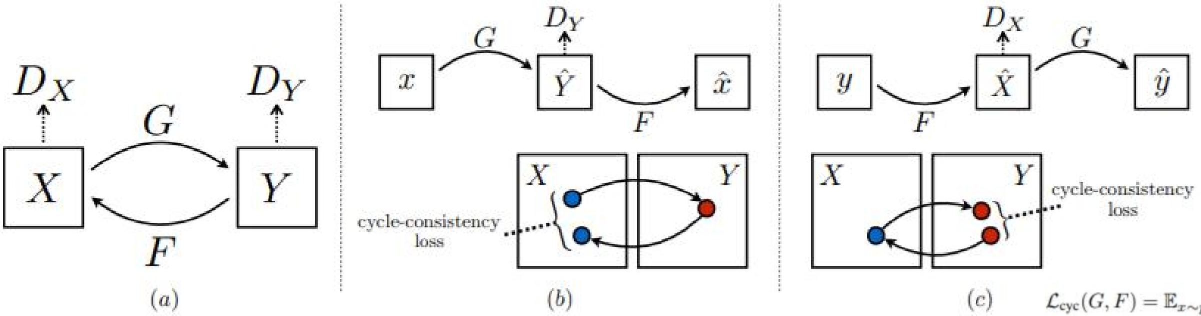
paired

Background

Domain transfer

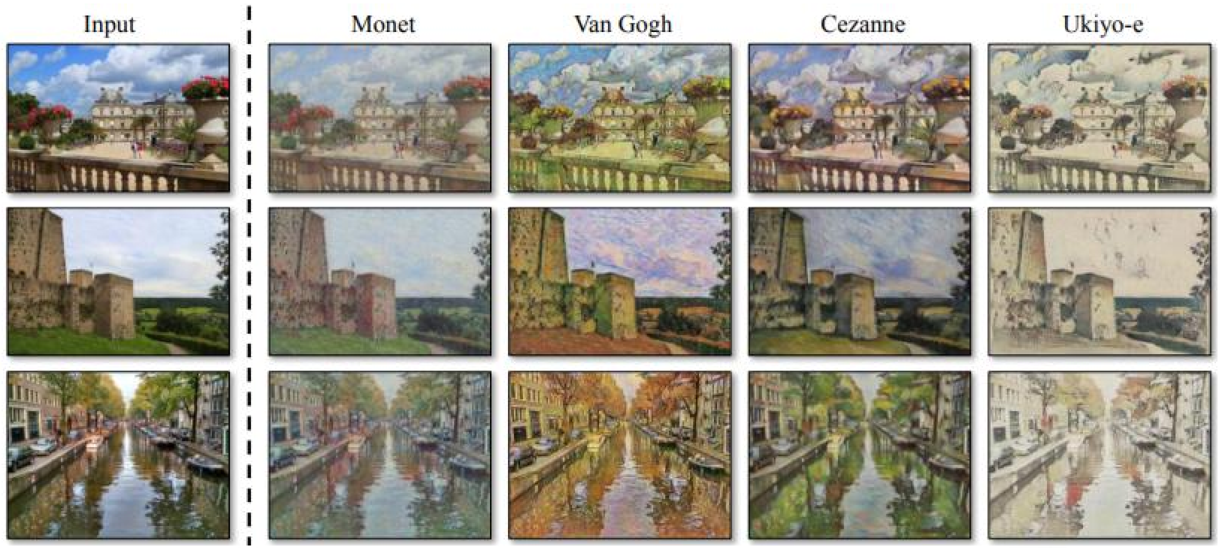
CycleGAN





$$\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{\text{data}}(y)}[\|G(F(y)) - y\|_1].$$

structure



results

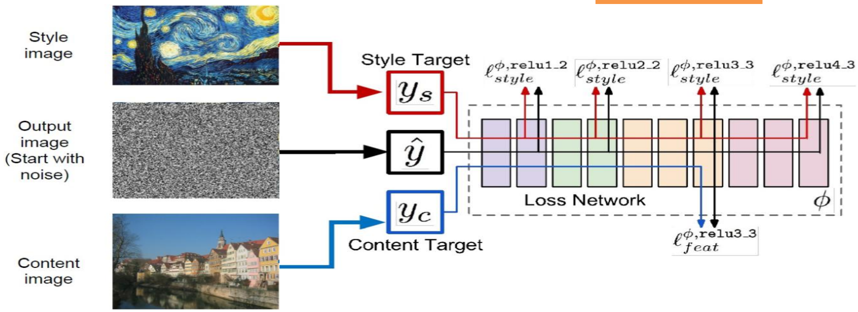
Style Transfer Task:

- Domain transfer
  - Pix2pix
  - CycleGAN
  - ......
- **Arbitrary style transfer**
  - Optimization-based
  - AdaIN
  - WCT
  - ......

Arbitrary style transfer

fixed

Gatys



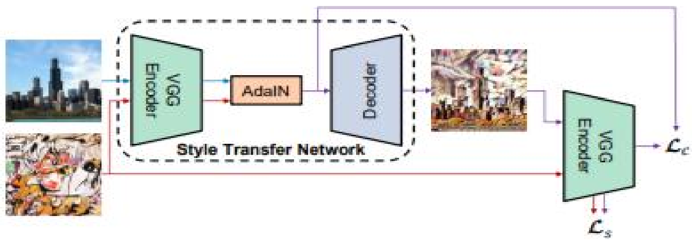
$$\mathcal{L}_{\text{content}}(\vec{p},\vec{x},l) = \frac{1}{2}\sum_{i,j}\left(F^l_{ij}-P^l_{ij}\right)^2$$

$$\mathcal{L}_{\text{style}}(\vec{a},\vec{x}) = \sum_{l=0}^{L} w_l E_l$$

$$G^l_{ij} = \sum_k F^l_{ik}F^l_{jk}$$

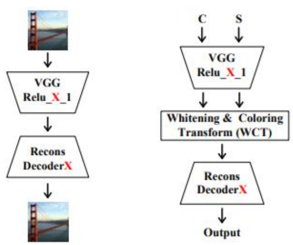
style loss

AdaIN

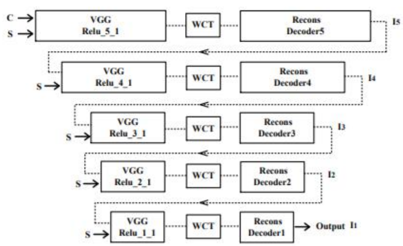


$$\text{AdaIN}(x,y) = \sigma(y)\left(\frac{x-\mu(x)}{\sigma(x)}\right)+\mu(y)$$

WCT



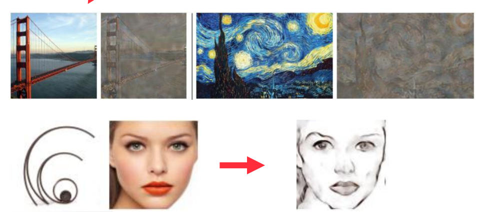
(a) Reconstruction (b) Single-level stylization



(c) Multi-level stylization

$$\hat{f}_c = E_c\, D_c^{-\frac{1}{2}}\, E_c^\top\, f_c$$

$$\hat{f}_{cs} = E_s\, D_s^{\frac{1}{2}}\, E_s^\top\, \hat{f}_c$$



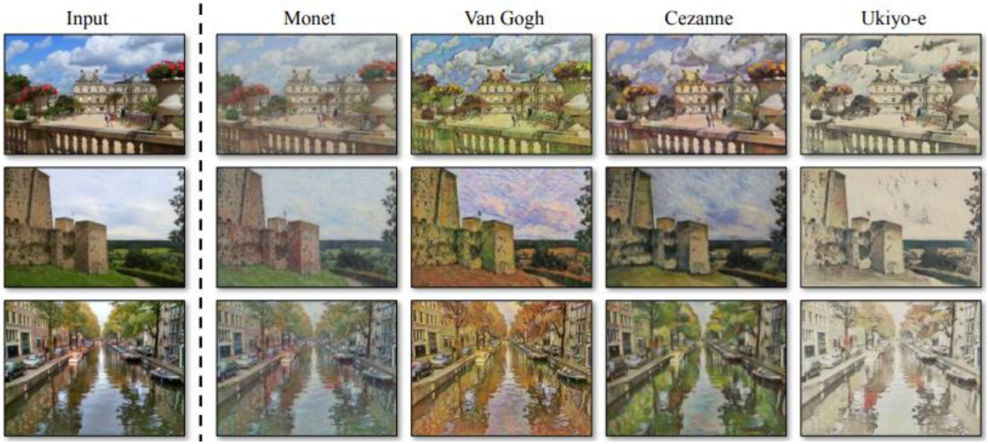
CRIPAC
智能感知与计算研究中心
Center for Research on Intelligent Perception and Computing

Motivation

- Few reasearches and dataset focu on Chinese paintings transfer task.
- Sub-style task.
- CycleGAN has "**over-grayscale**" problems.

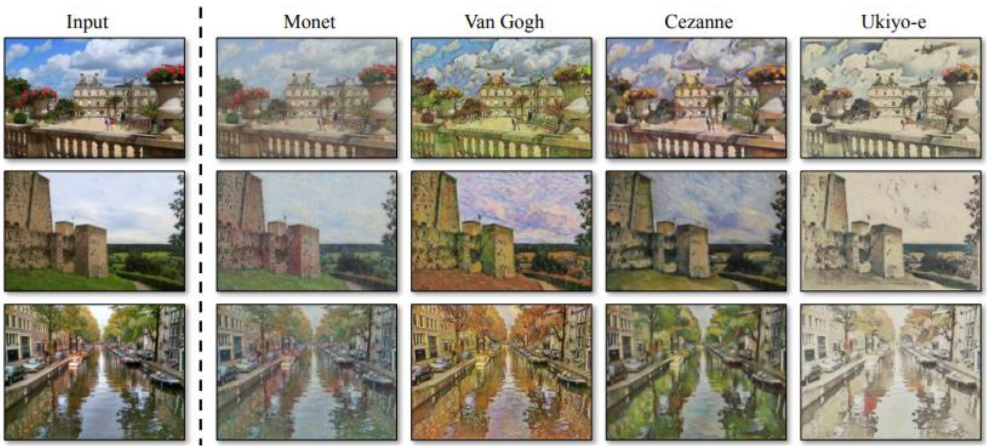
Motivation

- Few reasearches and dataset focu on Chinese paintings transfer task.
- Sub-style task.
- CycleGAN has "**over-grayscale**" problems.



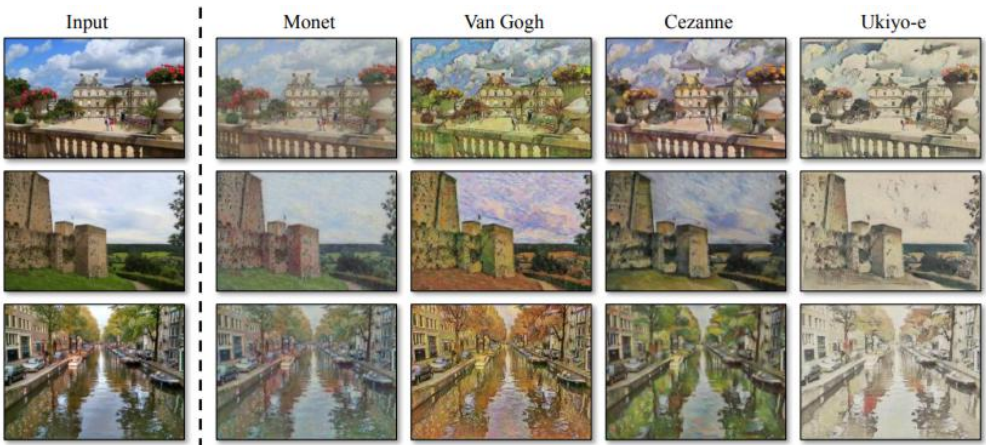
Motivation

- Few reasearches and dataset focu on Chinese paintings transfer task.
- Sub-style task.
- CycleGAN has "**over-grayscale**" problems.



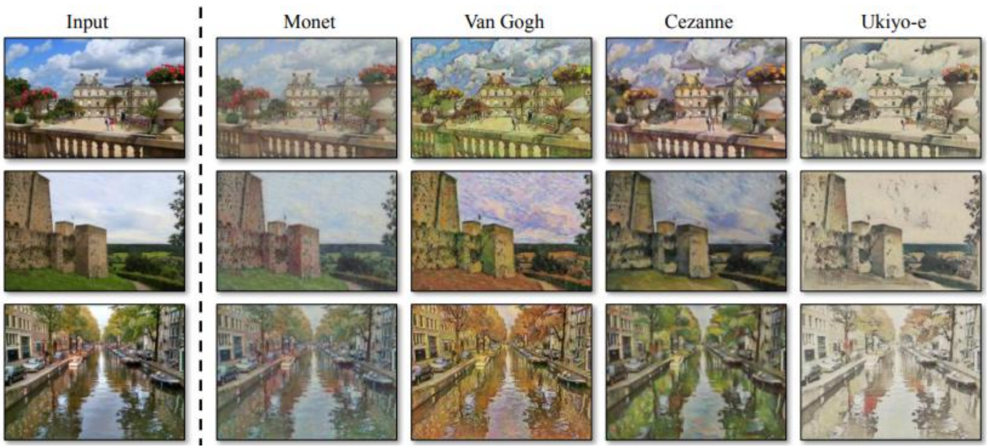
Motivation

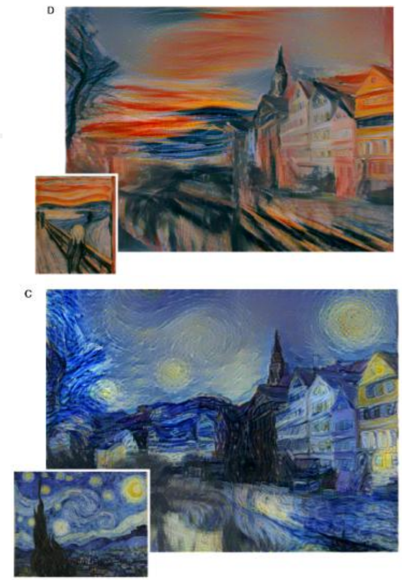
- Few reasearches and dataset focu on Chinese paintings transfer task.
- Sub-style task.
- CycleGAN has "**over-grayscale**" problems.



Motivation

- Few reasearches and dataset focu on Chinese paintings transfer task.
- Sub-style task.
- CycleGAN has "**over-grayscale**" problems.

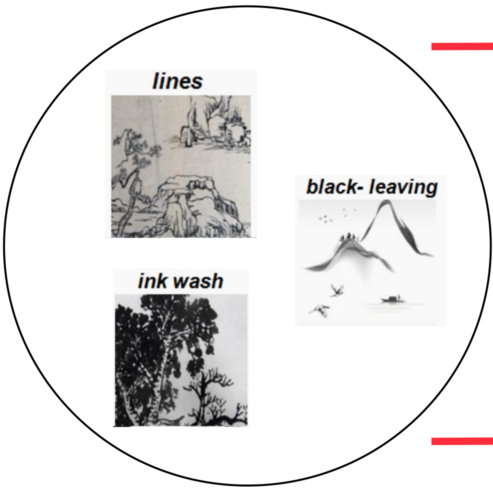




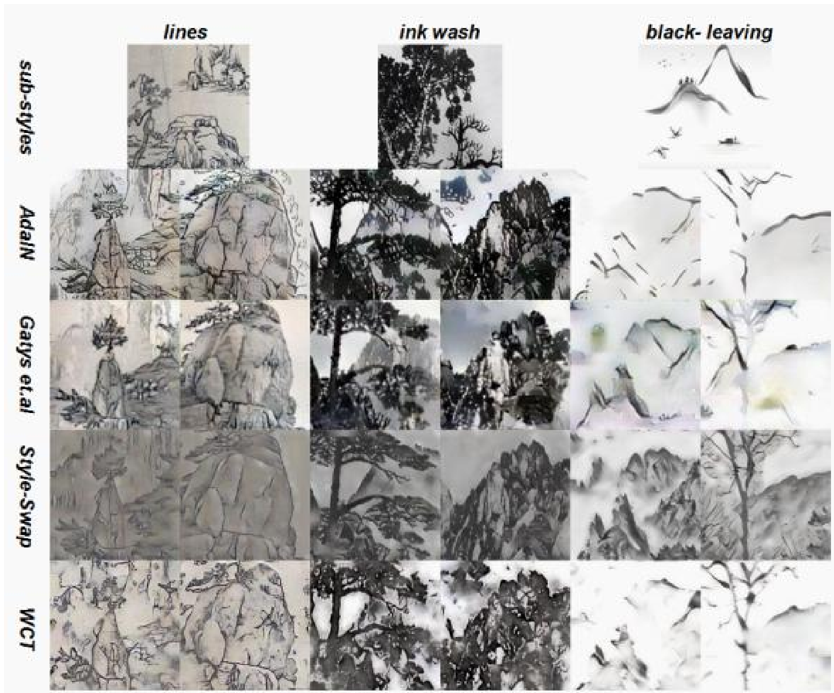
lack of Chinese paintings

Motivation

- Few reasearches and dataset focu on Chinese paintings transfer task.
- Sub-style task.
- CycleGAN has "**over-grayscale**" problems.

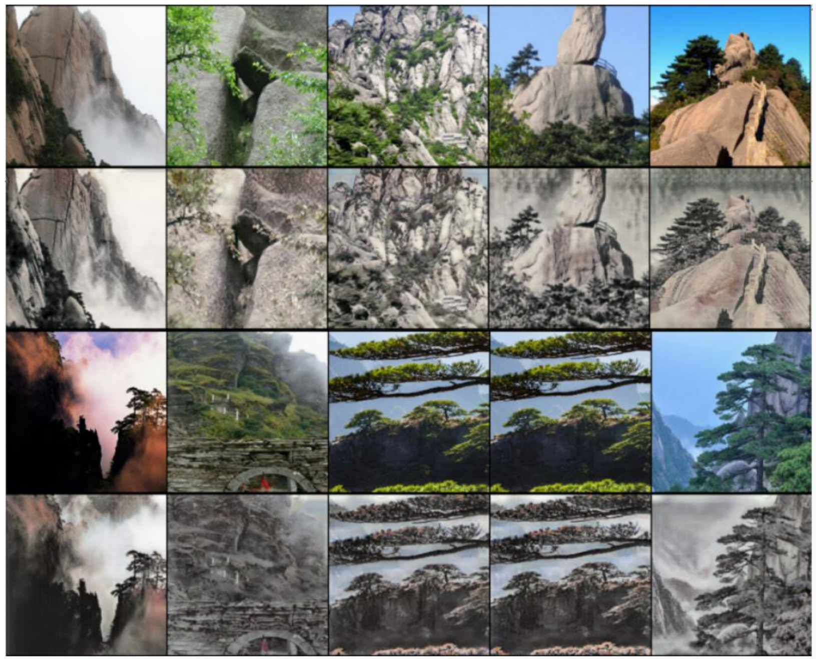


Chinese paintings



Motivation

- Few reasearches and dataset focu on Chinese paintings transfer task.
- Sub-style task.
- CycleGAN has "**over-grayscale**" problems.



Background

Motivation

Our work

Conclusion

Our work

- **Haar wavelet**
- **AWNet**
- **Results**

Our work

- **Haar wavelet**
- AWNet
- Results

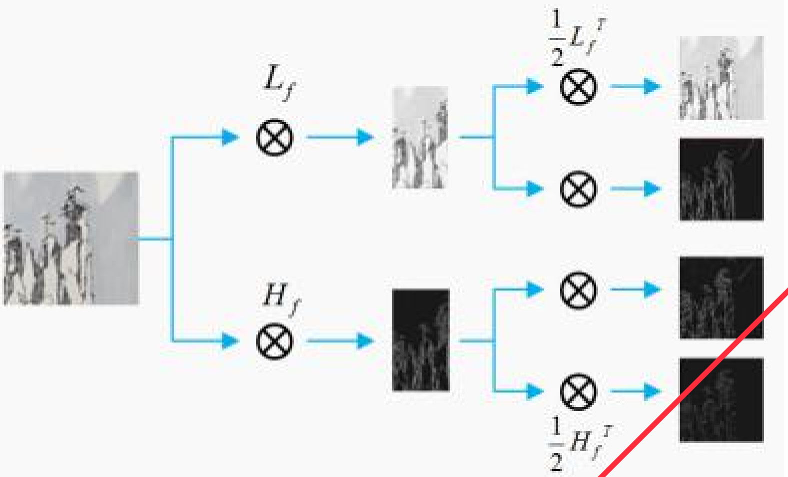
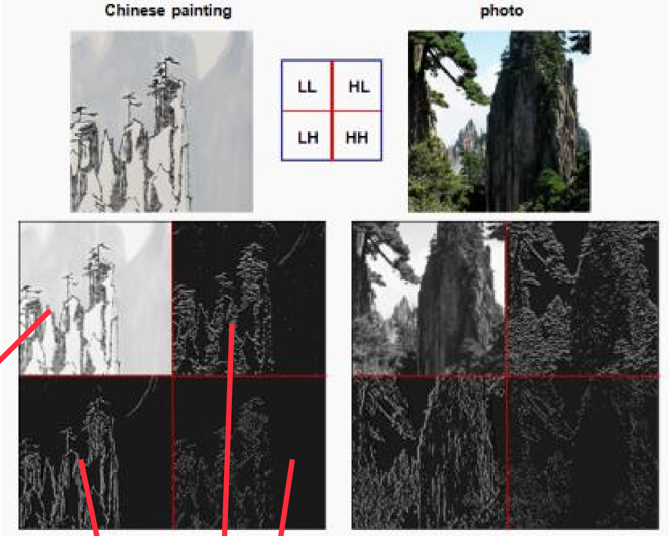


Fig. 4: Haar transform via convolution operation.



(b) Results of haar wavelet transform

Fig. 2: (a) The 'over grayscale' results of CycleGAN, i.e lack of 'Xieyi' information, (b) The results of haar wavelet transform, where the image enhancement is used for better display.

style information

details

- Haar wavelet
- **AWNet**
- Results

Total loss: $L = \alpha L_{GAN1} + \gamma L_{GAN2} + \delta(L_{cyc_1} + L_{cyc_2})$

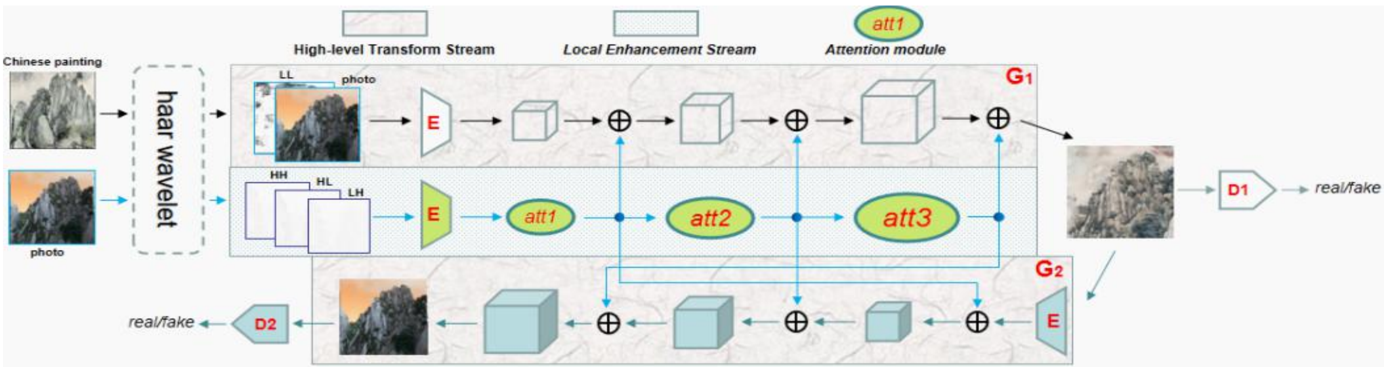


Fig. 3: Overview of our AWNet structure. Our model consists of two generators and discriminators. Each generator has 9 residual blocks while the discriminator compose of full convolutional layers. We feed photos and Chinese paintings to generators and discriminators to determine whether the output of generator is true or not. In order to capture the local details, we introduce a local enhancement stream and multi-scale self-attention modules to fuse them.

High-level Transform stream

+

Local Enhancement stream

+

Attention modules

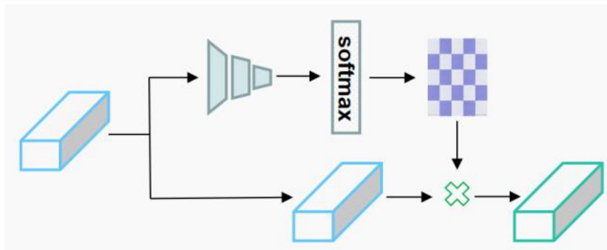


Fig. 5: Our attention module. The input of attention module is multi-channel feature, we can obtain a 1D vector after passing by several convolution and avgpooling layers and obtain the final feature via multiplying by input feature.

Our work

- Haar wavelet
- AWNet
- **Results**

Fig. 9: Our results compared with classic style transfer algorithms. AdaIN losses some content information, Gatys et.al generates blurry results, Style-Swap looks too dark and WCT's results in fragments. Moreover, none of them can reflect 'Xieyi' prospect. Our method achieves more attractive results compared with others.

Methods	Evaluation on P2ADataset	
	SSIM ↑	PSNR ↑
AdaIN	0.27	10.07
Gatys et.al	0.34	9.14
Style-Swap	0.36	**13.08**
WCT	0.18	8.67
Ours	**0.42**	11.04

TABLE I: Quantitative comparisons between ours and and the prevalent style transferring methods.

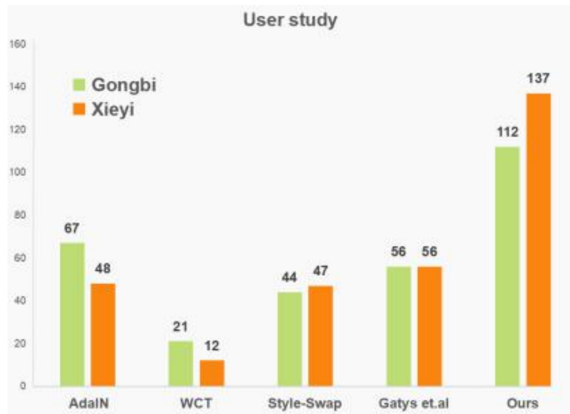


Fig. 10: Our user study. Horizontal axis shows different methods while vertical axis represents the votes on 'Gongbi' and 'Xieyi'.

Qualitative Evaluation

Quantitative Evaluation

Background

Motivation

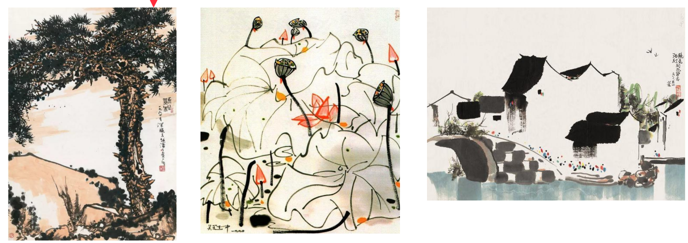
Our work

Conclusion

Conclusion

contribures:

- We propose a novel **AWNet** for photos-Chinese paintings transferring task, which can capture high-level information and local prospects simultaneously.
- To better portray the local prospects, we introduce a multi-scale **self-attention** mechanism to select details scattered in features of each layer.
- We propose a new large dataset, named **P2ADataset** contains unpaired photos and traditional Chinese paintings for photo-Chinese painting transferring task.



Future plan:

- Some failure cases
- High resolution

distorted & "ghosting"

graps

THANK YOU

CRIPAC

智能感知与计算研究中心

Center for Research on Intelligent Perception and Computing

Suggestions Questions