PHOTO STYLE TRANSFER WITH CONSISTENCY LOSSES

Xu YAO

Gilles PUY
Parick PÉREZ

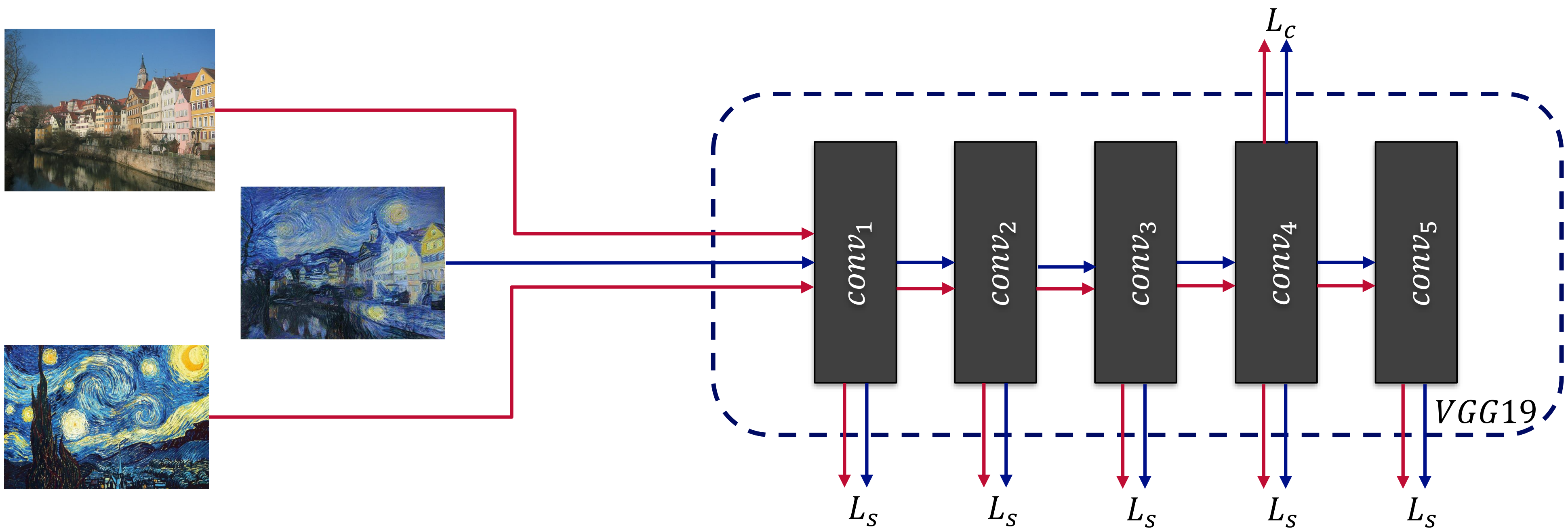
Style Transfer

L. A. Gatys et al., « Image style transfer using convolutional neural networks », CVPR 2016



Neural Style Transfer

L. A. Gatys et al., « Image style transfer using convolutional neural networks », CVPR 2016



$$\mathcal{L}_t(x) = \alpha\,\mathcal{L}_c(x, x_c) + \beta\,\mathcal{L}_s(x, x_s)$$

- Two images are similar in **content** if their high-level features as extracted by a trained classifier are close in Euclidian distance.
- Two images are similar in **style** if the difference between the features' **Gram matrices** has a small Frobenius norm.

Photo Style Transfer

L. A. Gatys et al., « Image style transfer using convolutional neural networks », CVPR 2016

Style Content

- Neural Style Transfer on photorealistic style transfer tasks.

Related Work

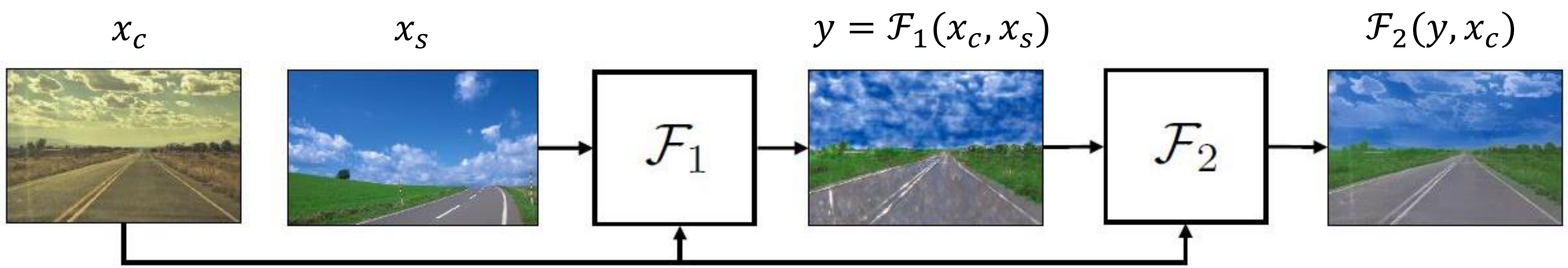
F. Luan et al., « Deep Photo Style Transfer », CVPR 2017

Style Content

- By adding a regularization term (matting Laplacian) to the objective function of neural style transfer, this method achieves photorealistic style transfer.
- Objective: $\mathcal{L}_t(x) = \alpha\,\mathcal{L}_c(x, x_c) + \beta\,\mathcal{L}_s(x, x_s) + \lambda\,\mathcal{L}_m(x), \mathcal{L}_m(x) = \mathrm{Tr}(x^T L(x_c)x)$

Related Work

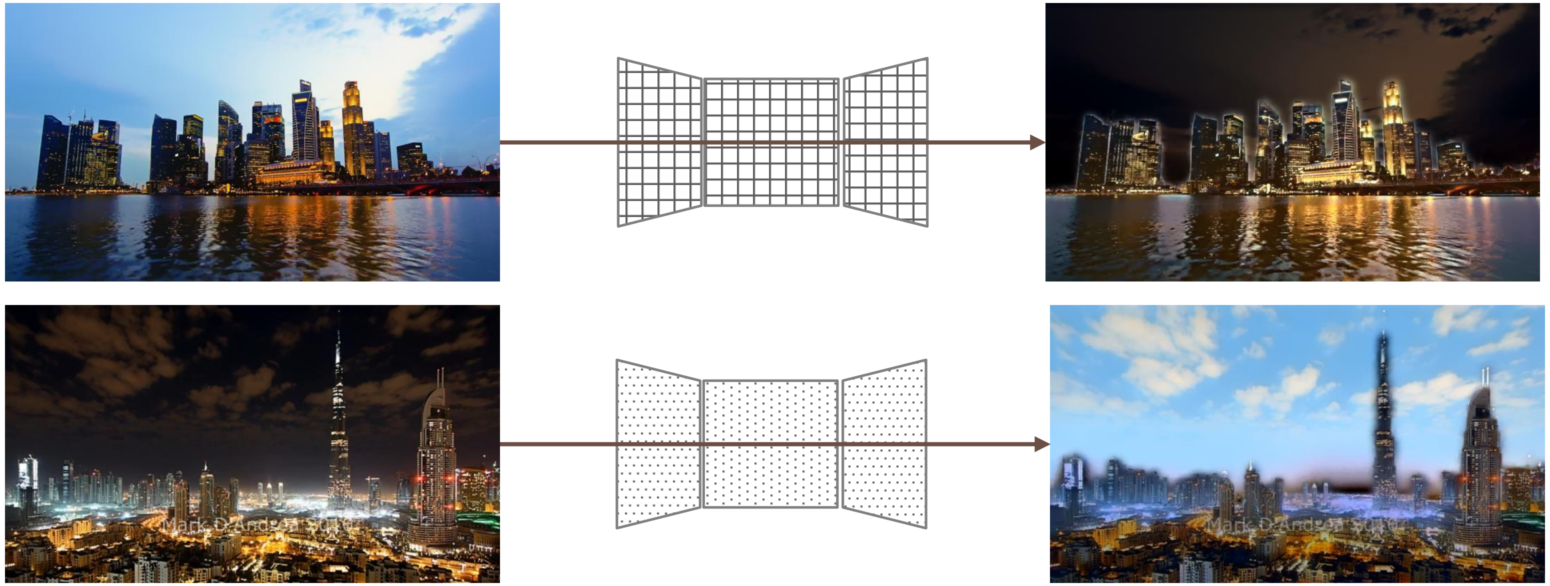
Y. Li et al., « A Closed-form Solution to Photorealistic Image Stylization », ECCV 2018



- Fast photorealistic stylization method.
- The algorithm consists of two mappings $\mathcal{F}_1$ and $\mathcal{F}_2$.
  - $\mathcal{F}_1$ maps x_c to an intermediate result in the style of x_s.
  - $\mathcal{F}_2$ removes the artifacts and produces a photorealistic stylized result.

Our Approach

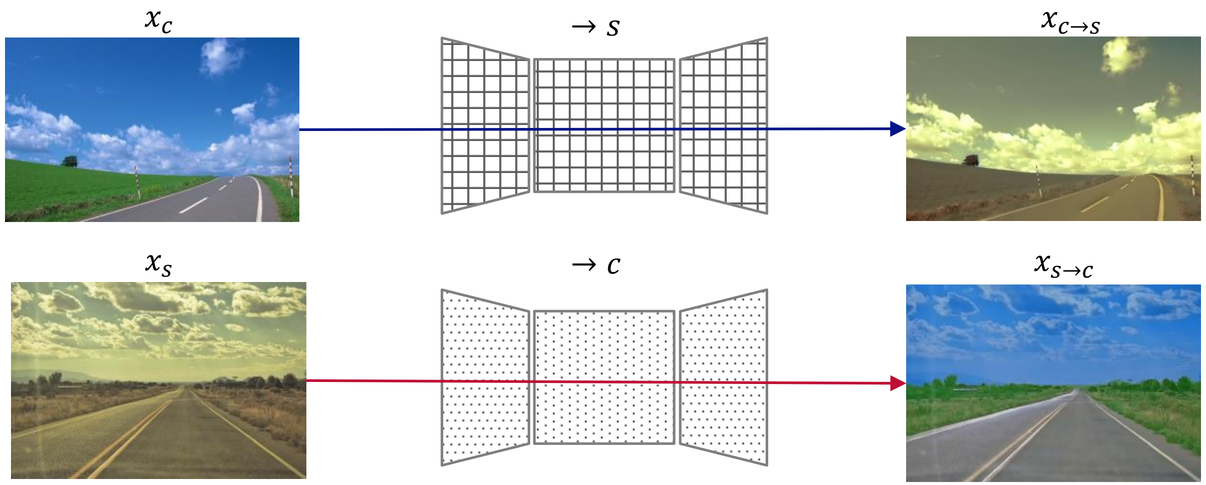
Our Approach



- With no other dataset than **two input photos**, train CNNs to transfer their styles between each other.
- The output needs to preserve the structure of the content photo, imitate the style of the reference photo, meanwhile, looks like a real photo.

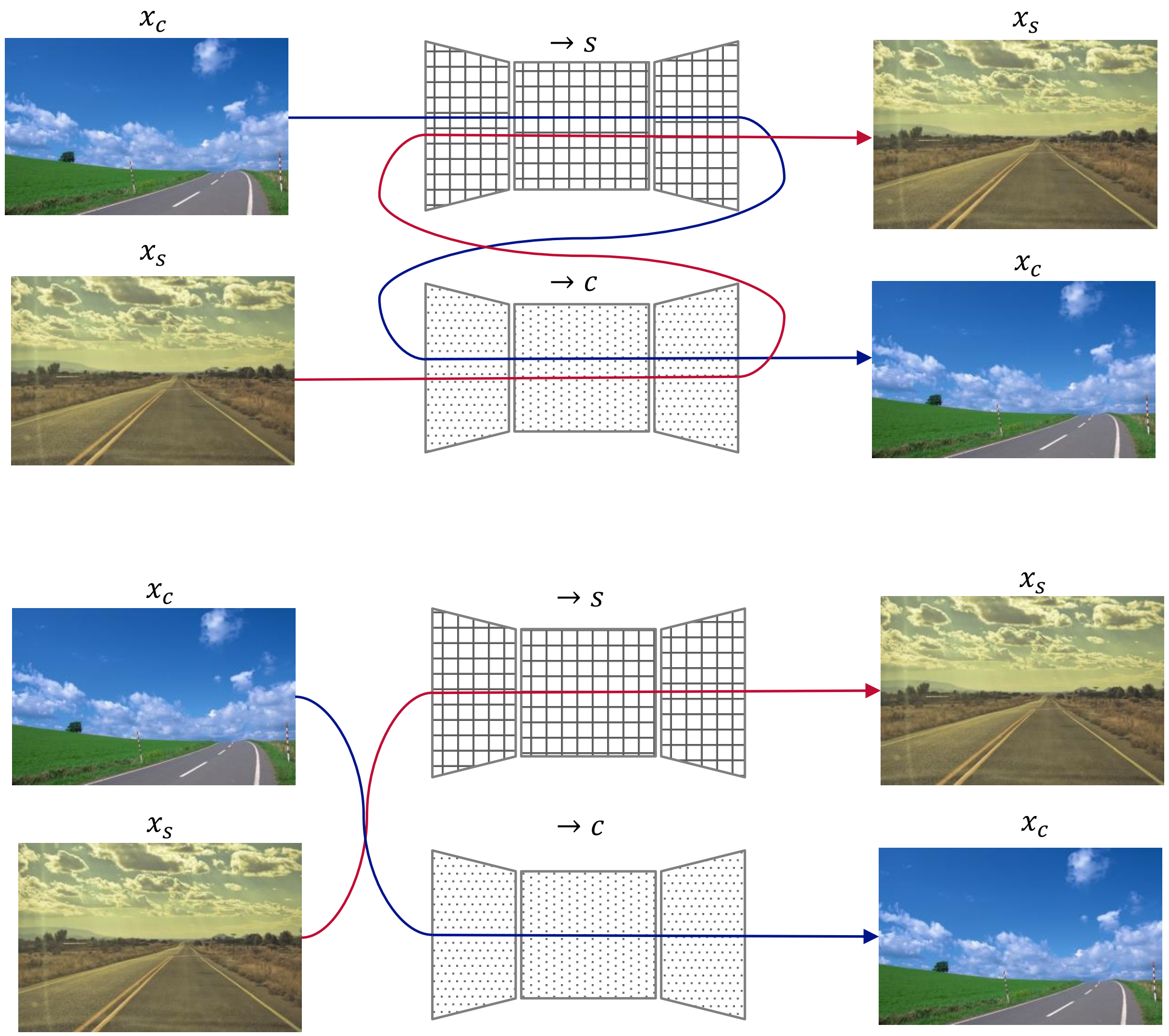
Objective

- **Style Loss**
  - The two networks ($\rightarrow$ s, $\rightarrow$ c) transfer the styles of x_s and x_c respectively.



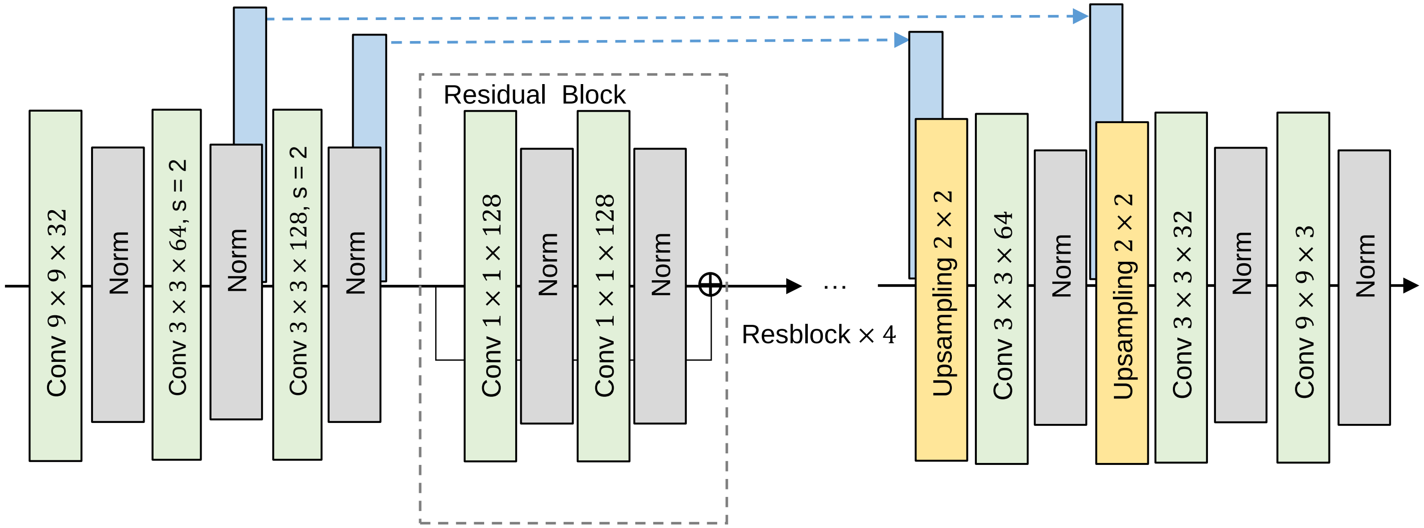
Objective

- **Content Loss**
  - **Cycle Consistency Loss**
    - If we apply the two networks ($\rightarrow s$, $\rightarrow c$) consecutively on x_c, we should get x_c itself (and vice-versa).
  - **Self Consistency Loss**
    - If we transfer the style of one photo to itself, we should get the original photo.



J. Zhu et al., « Unpaired image-to-image translation using cycle-consistent adversarial networks », ICCV 2017

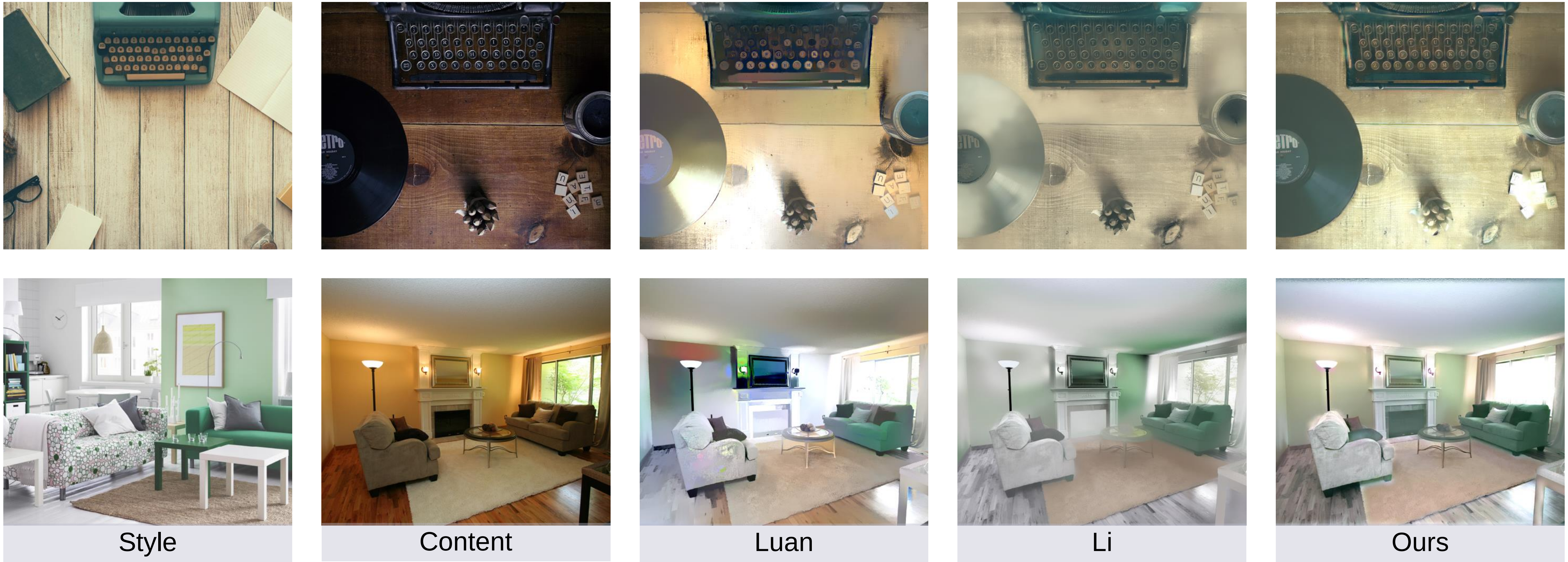
Network Architecture



- Each network transfers one style (for photos having several styles, we train one network for each style).
- During the training, the two networks share the same convolution layers but have different normalization layers which encode different styles.

V. Dumoulin et al., « A Learned Representation For Artistic Style », ICLR 2017

Results



F. Luan et al., « Deep Photo Style Transfer », CVPR 2017
Y. Li et al., « A Closed-form Solution to Photorealistic Image Stylization », ECCV 2018

Results

Style | Content | Luan | Li | Ours

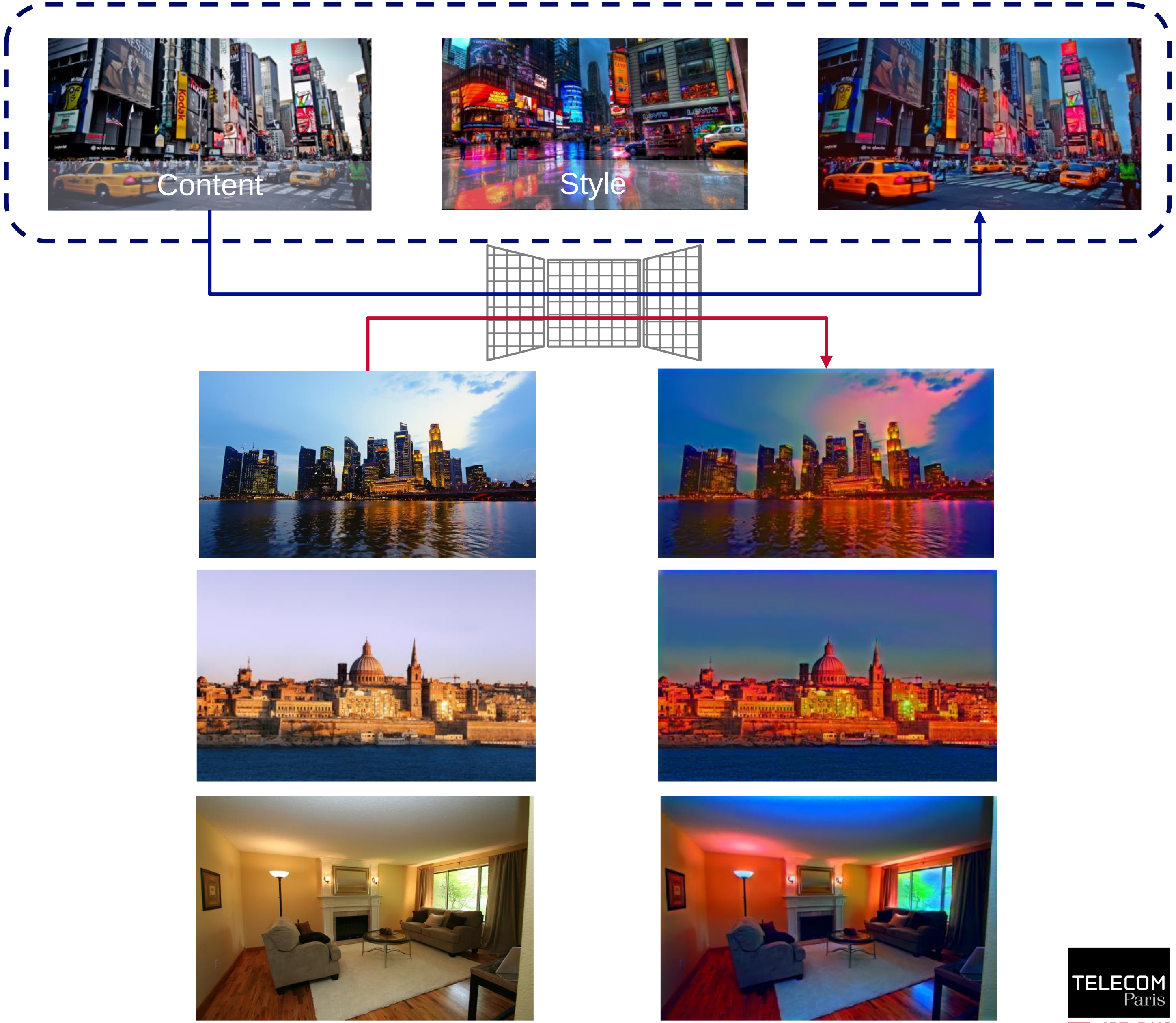
F. Luan et al., « Deep Photo Style Transfer », CVPR 2017
Y. Li et al., « A Closed-form Solution to Photorealistic Image Stylization », ECCV 2018

Two extra advantages of our method

- Generalization to unseen images
- Retraining for new styles

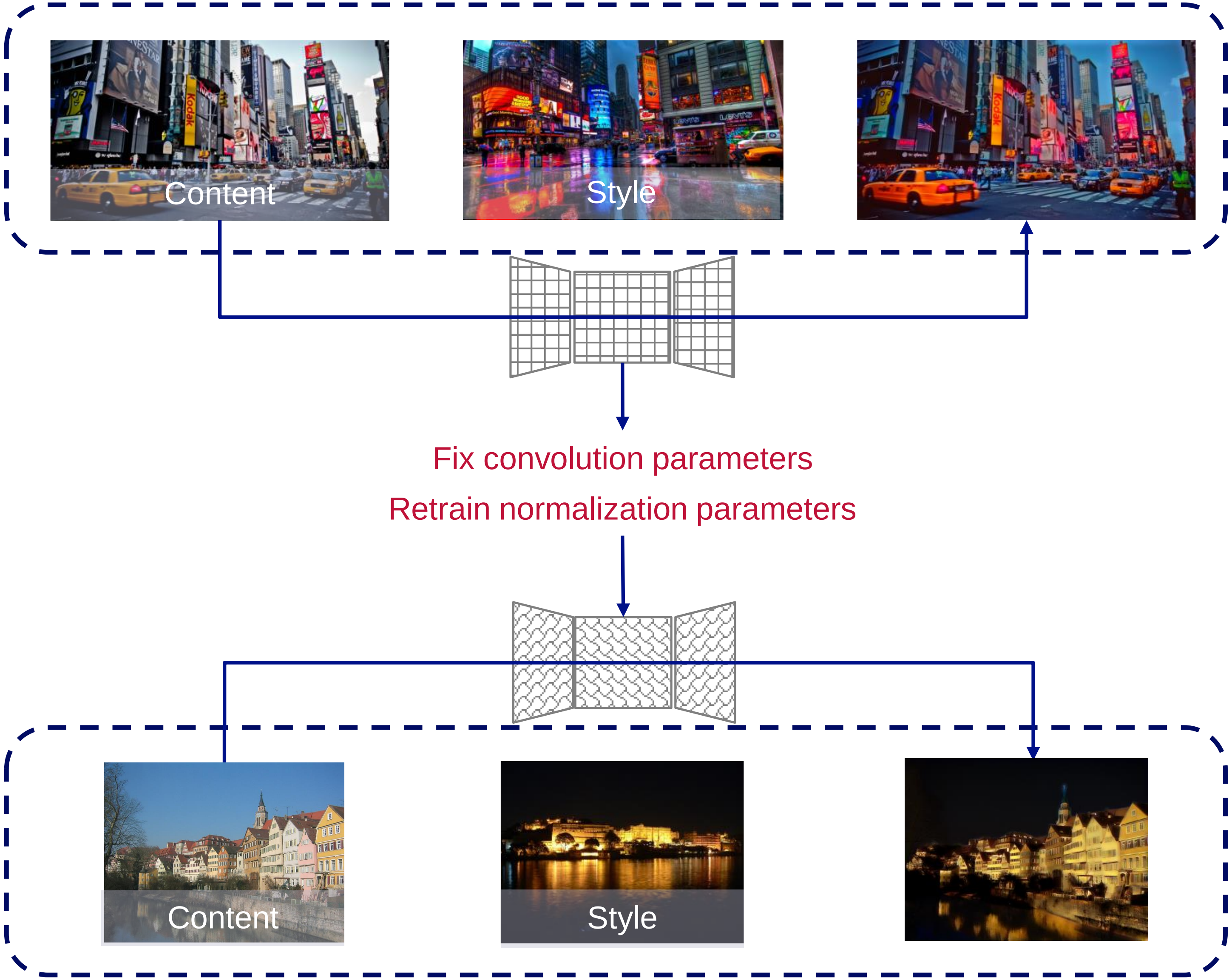
Generalization to unseen images

Even though trained on merely **two images**, our network can stylize images not viewed at training time.



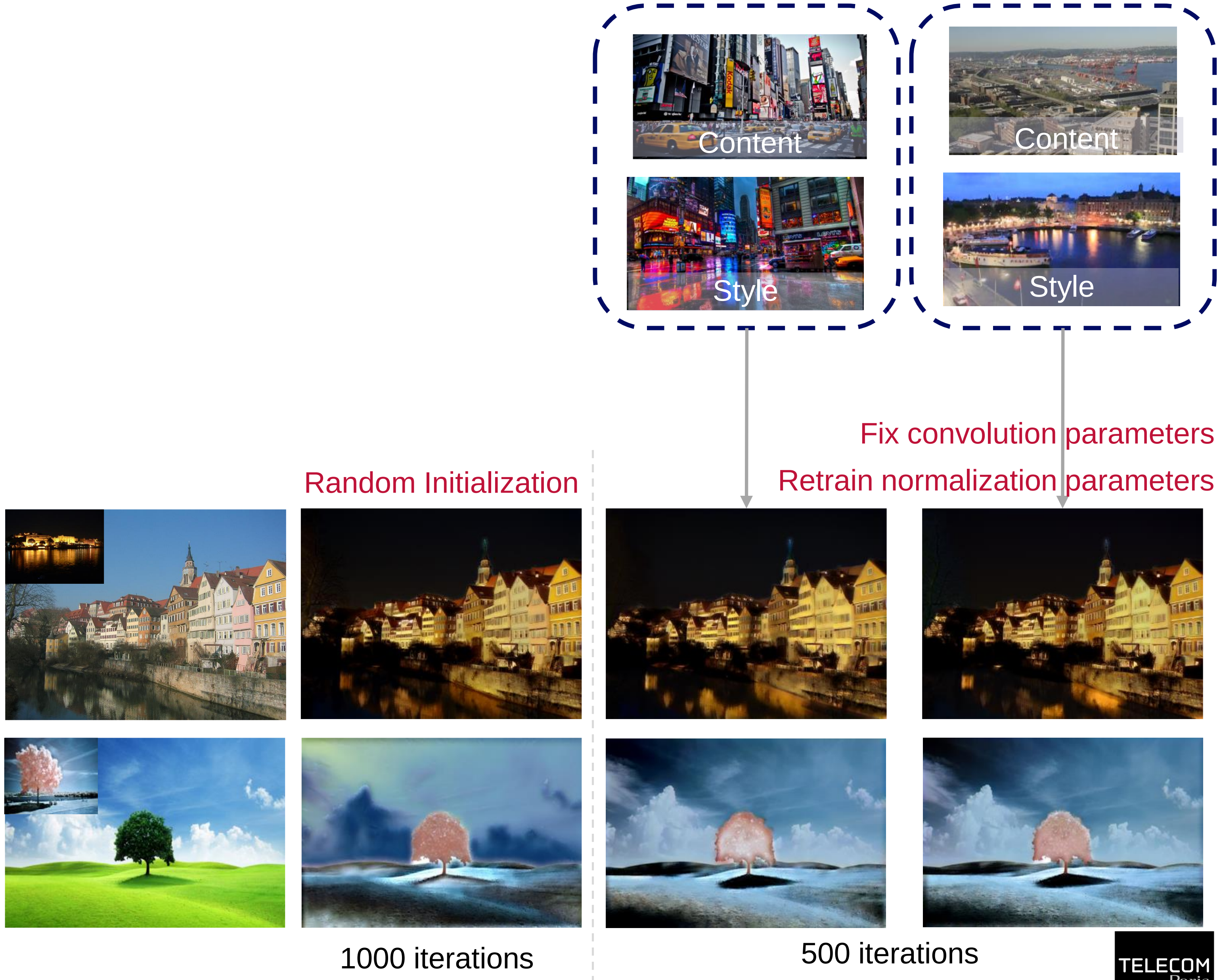
Retraining for new styles

It is enough to retrain the normalization parameters to adapt our networks to a new style, even though the convolutional layers are pre-trained using **only two images**.



Retraining for new styles

It is enough to retrain the normalization parameters to adapt our networks to a new style, even though the convolutional layers are pre-trained using **only two images**.



Failure Case

Style Content

Conclusion

- **Our approach**
  - Using only two input photos, we train a pair of deep convolution networks with consistency losses, each of which transfers the style of one photo to the other.
  - Photorealism is achieved thanks to the consistency losses.

- **Two extra advantages**
  - Even though trained on merely two images, our network can stylize images not viewed at training time.
  - To adapt the network for a new style, we only need to retrain a small subset of the parameters.

- **Future work**
  - Reduce the artefacts at the boundary of different semantic regions and reduce the overexposure that sometimes appears in small regions of the results.
  - Adapt the method for arbitrary style transfer.

Q & A


gordonlingfilm
gordonlingfilm