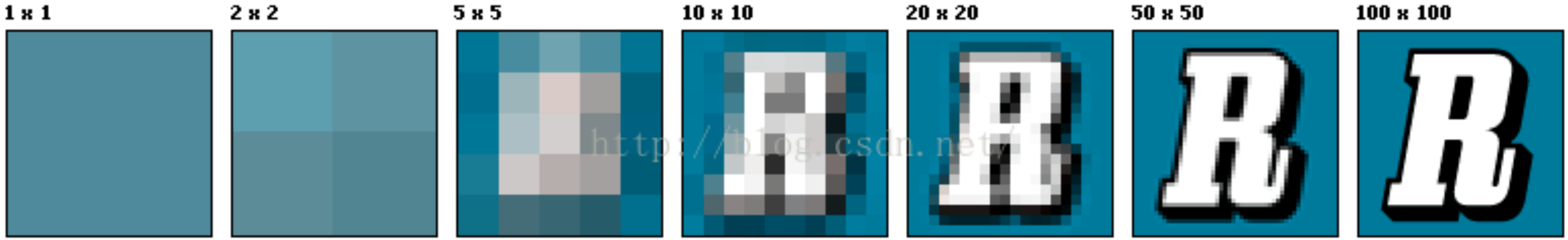
04

Super-Resolution Image Reconstruction

Super-Resolution Image Reconstruction

- **Image resolution**



https://blog.csdn.net/jiangjieqazwsx/article/details/50547485

Super-Resolution Image Reconstruction

- **Low quality and low resolution** ⋯⋯→

High quality and high resolution

Yaniv Romano and John Isidoro and Peyman Milanfar.
RAISR: Rapid and Accurate Image Super Resolution. arXiv 2016

Super-Resolution Image Reconstruction

- **Time resolution** → **Spatial resolution**

  - Image super-resolution method based on reconstruction

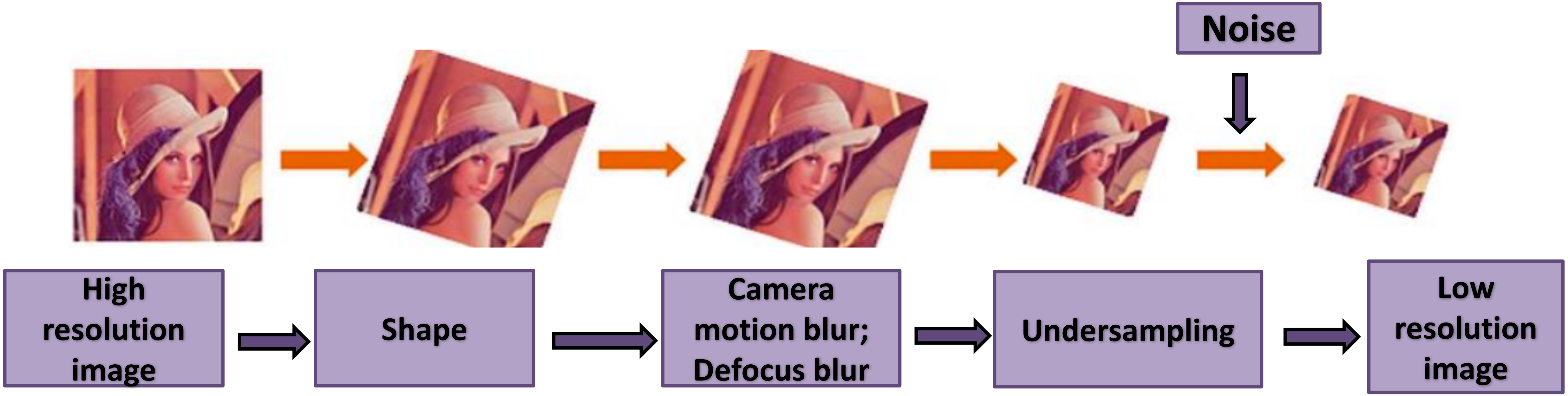
  - Image super-resolution method based on learning

Super-Resolution Image Reconstruction

- **Principle**

  - Degradation model of image degradation

$$g_k = DBM_K z + n_k$$



Single Image Super-Resolution

Learning a deep convolutional network for image super-resolution
Dong C, Chen C L, He K, et al

Photo-realistic single image super-resolution using a generative adversarial network
Ledig C, Theis L, Huszár F, et al

Recovering realistic texture in image super-resolution by deep spatial feature transform
Wang X, Yu K, Dong C, et al

Learning a deep convolutional network for image super-resolution

- **Sparse Representations**
- **Sparse Coding**

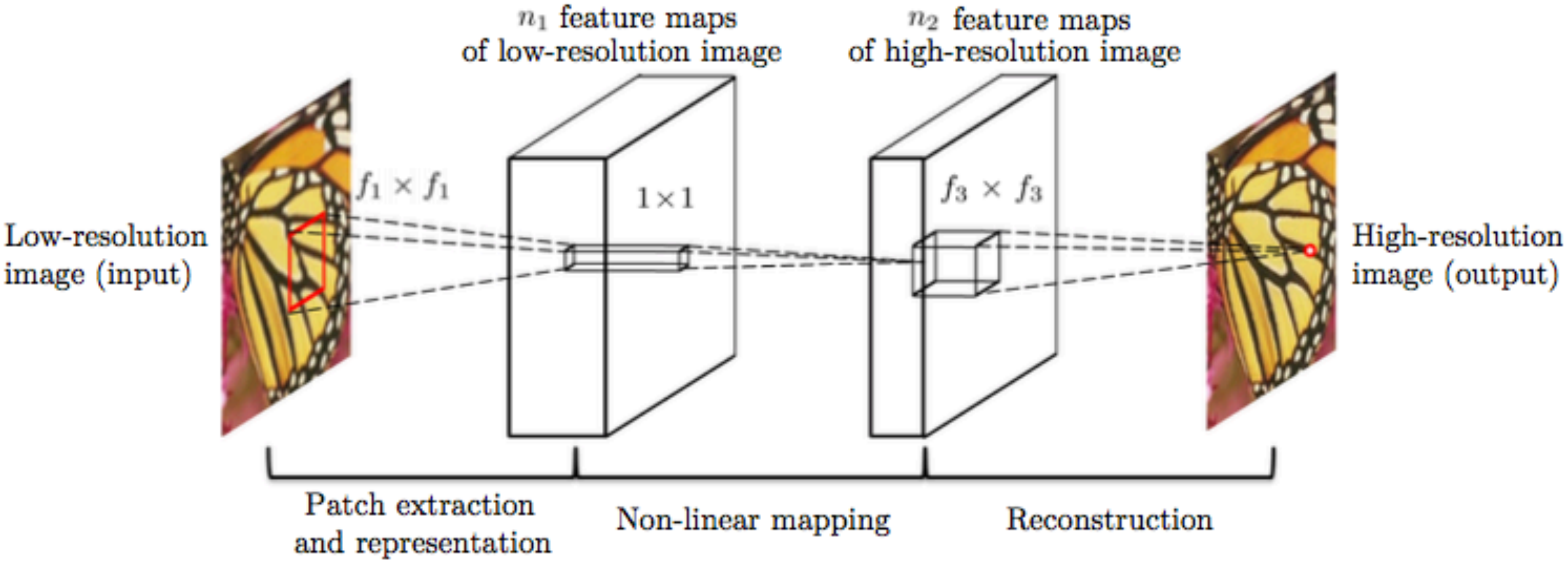
$$f = a_1 * x_1 + a_2 * x_2 + \cdots + a_n * x_n$$

Learning a deep convolutional network for image super-resolution

- **Sparse Coding**
  - Overlapping patched are densely extracted from the image and pre-processed.
  - These patches are then encoded by a low-resolution dictionary.
  - The sparse coefficients are passed into a high-resolution dictionary for reconstructing high-resolution patches.
  - The overlapping reconstructed patches are aggregated (or averaged) to produce the output.

Learning a deep convolutional network for image super-resolution

- **Patch extraction and representation**
- **Non-linear mapping**
- **Reconstruction**



Learning a deep convolutional network for image super-resolution

- **Loss function**

Mean Squared Error (MSE)

$$L(\Theta) = \frac{1}{n}\sum_{i=1}^{n} \|F_i(Y;\Theta) - X_{\mathrm{i}}\|^2$$

Learning a deep convolutional network for image super-resolution

- **Result**

Photo-realistic single image super-resolution using a generative adversarial network

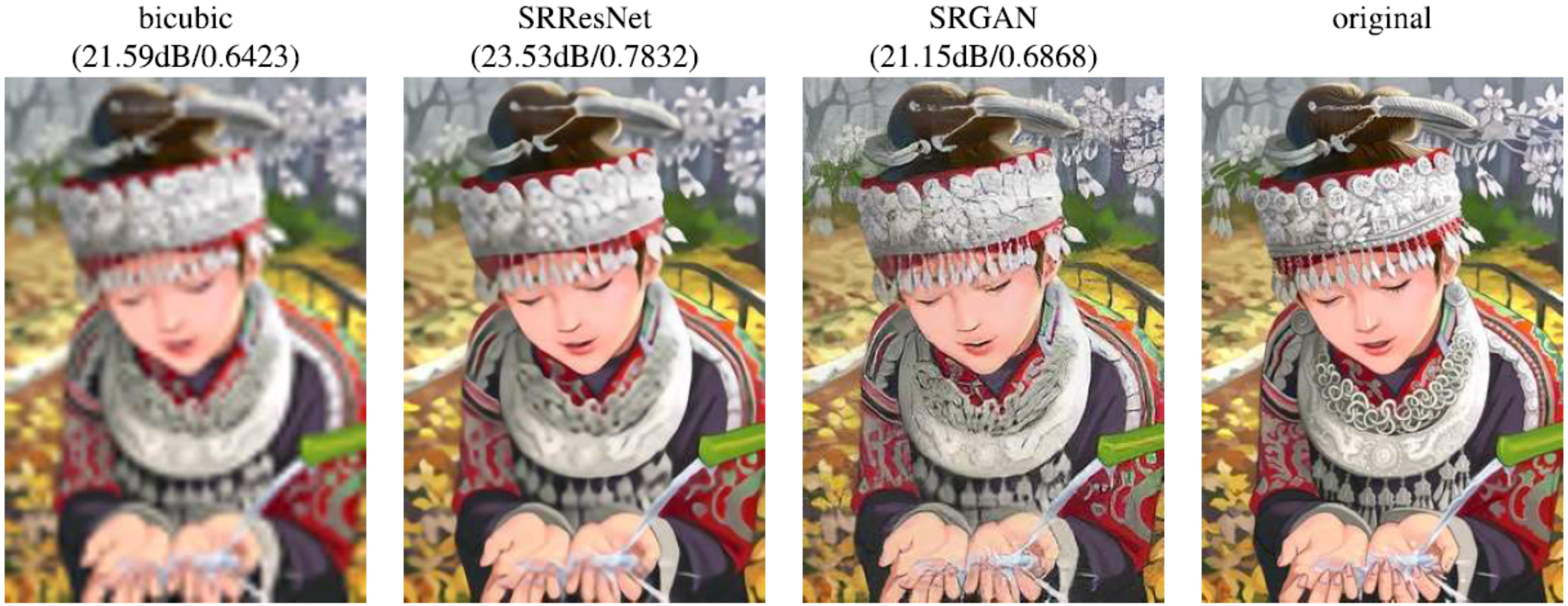
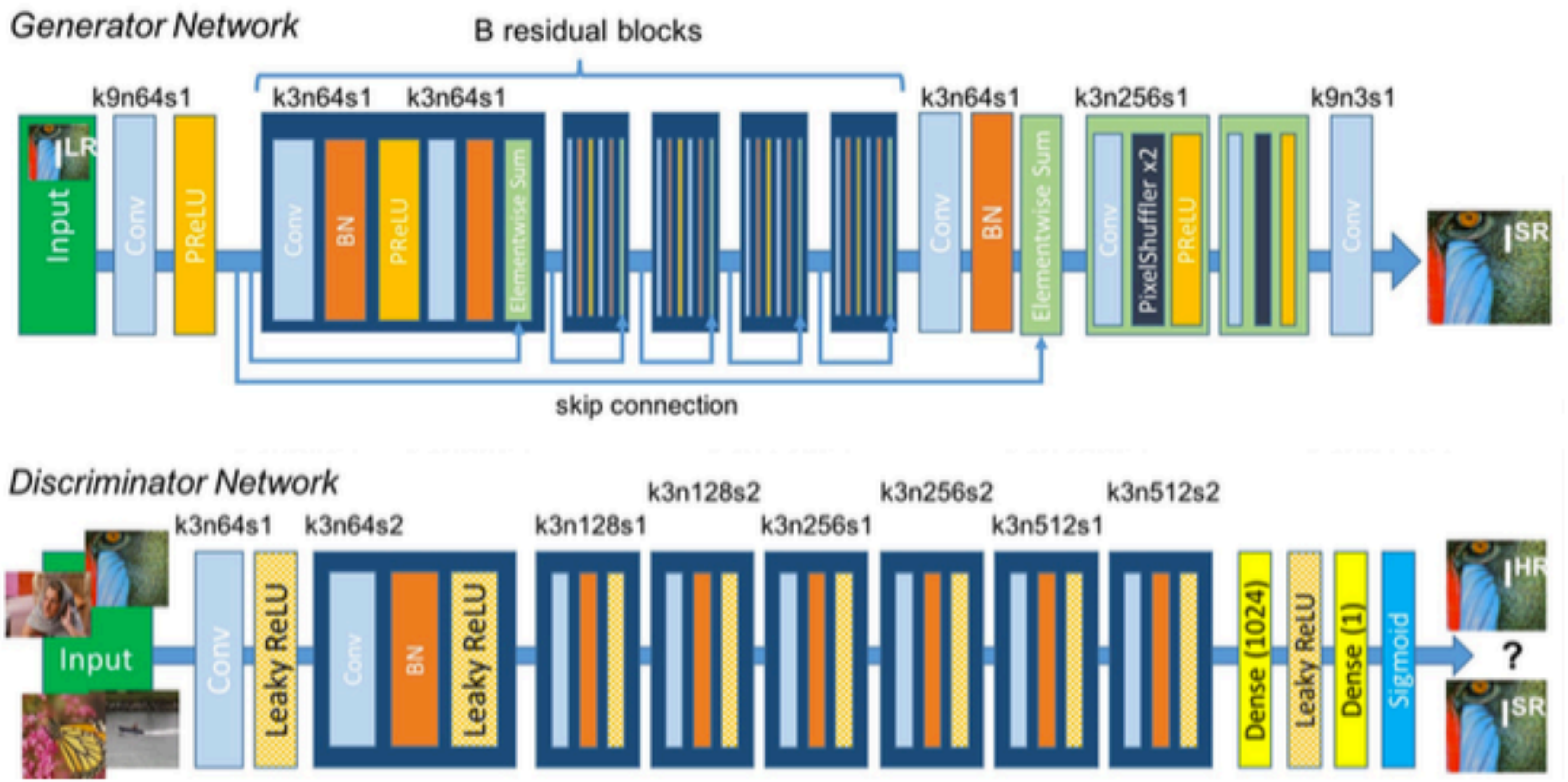


Figure 2: From left to right: bicubic interpolation, deep residual network optimized for MSE, deep residual generative adversarial network optimized for a loss more sensitive to human perception, original HR image. Corresponding PSNR and SSIM are shown in brackets. [4× upscaling]

Photo-realistic single image super-resolution using a generative adversarial network



Loss function

$$l^{SR} = \underbrace{\underbrace{l^{SR}_{X}}_{\text{content loss}} + \underbrace{10^{-3} l^{SR}_{Gen}}_{\text{adversarial loss}}}_{\text{perceptual loss (for VGG based content losses)}}$$

Content loss

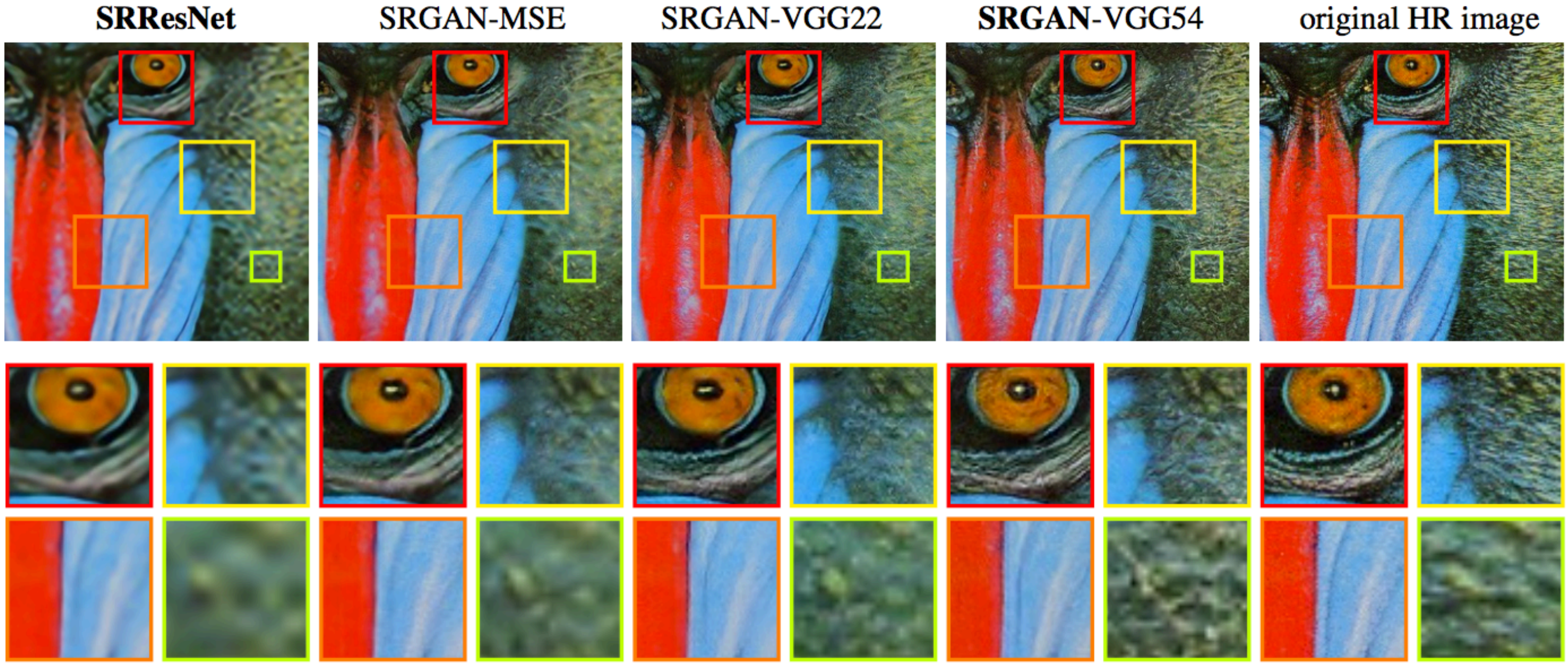
$$l^{SR}_{MSE} = \frac{1}{r^2WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I^{HR}_{x,y} - G_{\theta_G}(I^{LR})_{x,y})^2$$

$$l^{SR}_{VGG/i.j} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2$$

Adversarial loss

$$l^{SR}_{Gen} = \sum_{n=1}^{N} -\log D_{\theta_D}(G_{\theta_G}(I^{LR}))$$

Photo-realistic single image super-resolution using a generative adversarial network



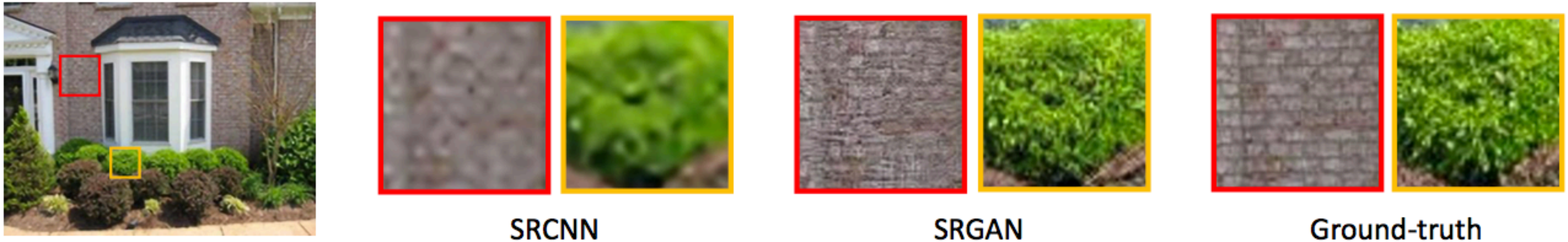
Recovering realistic texture in image super-resolution by deep spatial feature transform

Xintao Wang   Ke Yu   Chao Dong   Liang Lin   Chen Change Loy

http://ice.dlut.edu.cn/valse2018/ppt/2018ValseDChao.pdf

Previous work

- Contemporary SR algorithms are mostly CNN-based methods[1].
- Most of CNN-based methods use pixel-wise loss function. (MSE-based model)
  - ✓ good at recovering edges and smooth areas
  - ❑ not good at texture recovery
- Adversarial loss is introduced in SRGAN[2] and EnhanceNet[3]. (GAN-based model)
  - ✓ encourage the network to favor solutions that look more like natural images
  - ✓ visual quality of reconstruction is significantly improved



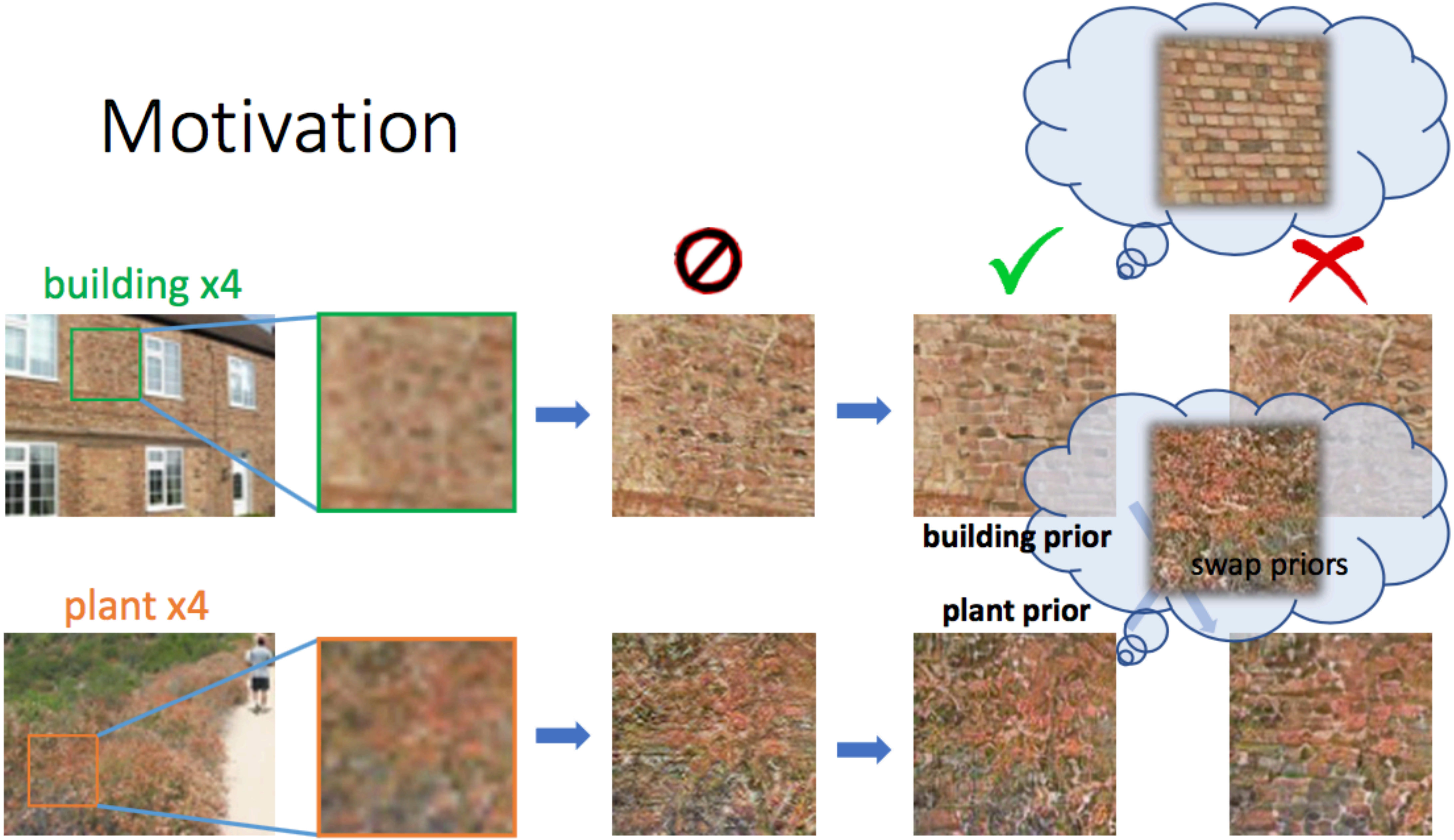
SRCNN SRGAN Ground-truth

[1] C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In ECCV, 2014.

[2] C. Ledig, L. Theis, F. Husz´ar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, et al. Photo-realistic single image super-resolution using a generative adversarial network. In CVPR, 2017.

[3] M. S. Sajjadi, B. Sch¨olkopf, and M. Hirsch. EnhanceNet: Single image super-resolution through automated texture synthesis. In ICCV, 2017.

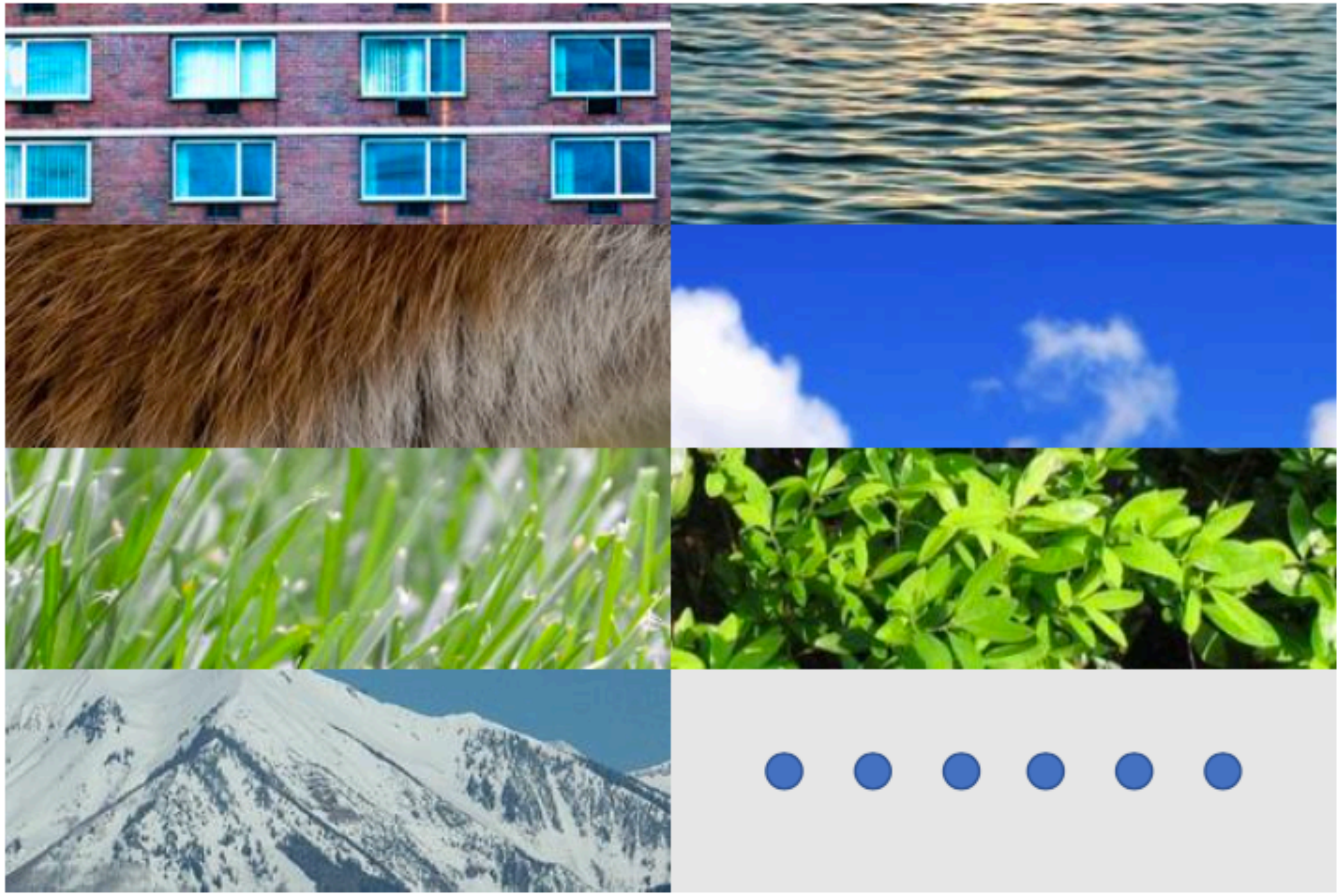
Motivation



http://ice.dlut.edu.cn/valse2018/ppt/2018ValseDChao.pdf

vision@ouc

Semantic categorical prior



http://ice.dlut.edu.cn/valse2018/ppt/2018ValseDChao.pdf

vision@ouc

Issues

1. How to represent the semantic categorical prior?

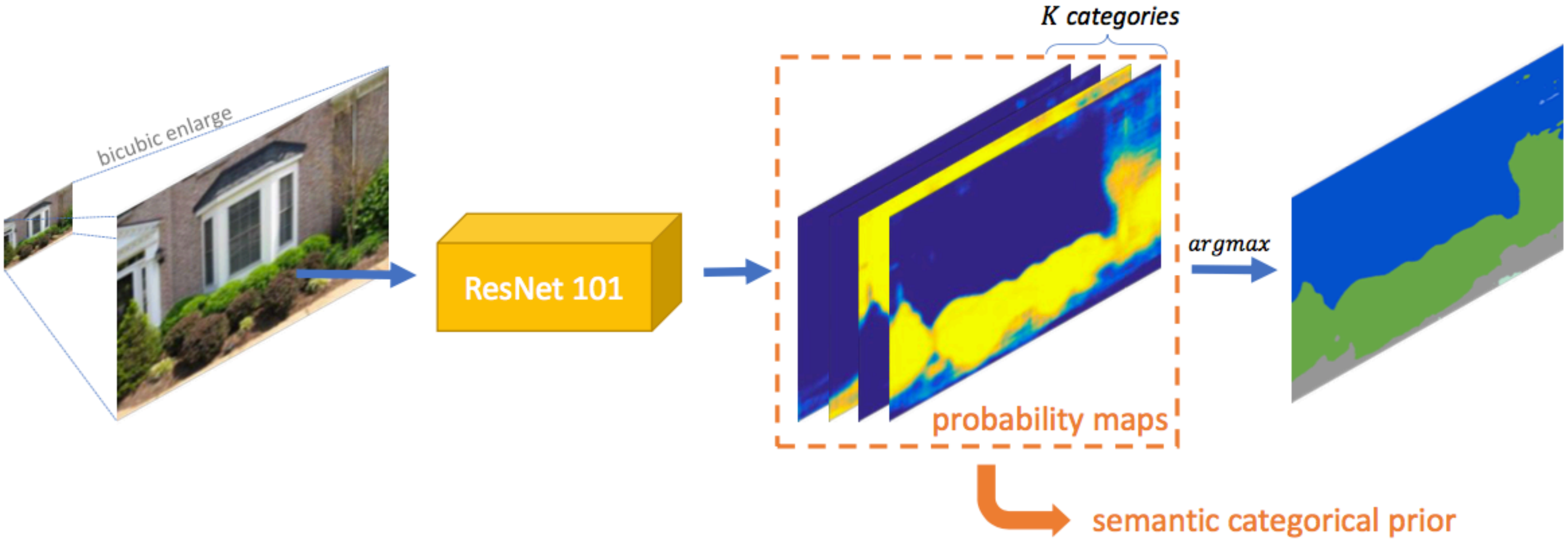
- ❑ Our approach:
  explore semantic segmentation probability maps as the categorical prior up to pixel level.

2. How categorical prior can be incorporated into the reconstruction process effectively?

- ❑ Our approach:
  propose a novel **S**patial **F**eature **T**ransform that is capable of altering the network behavior conditioned on other information.

Represent categorical prior

- **Contemporary CNN segmentation network[1]**
  - fine-tuned on LR images

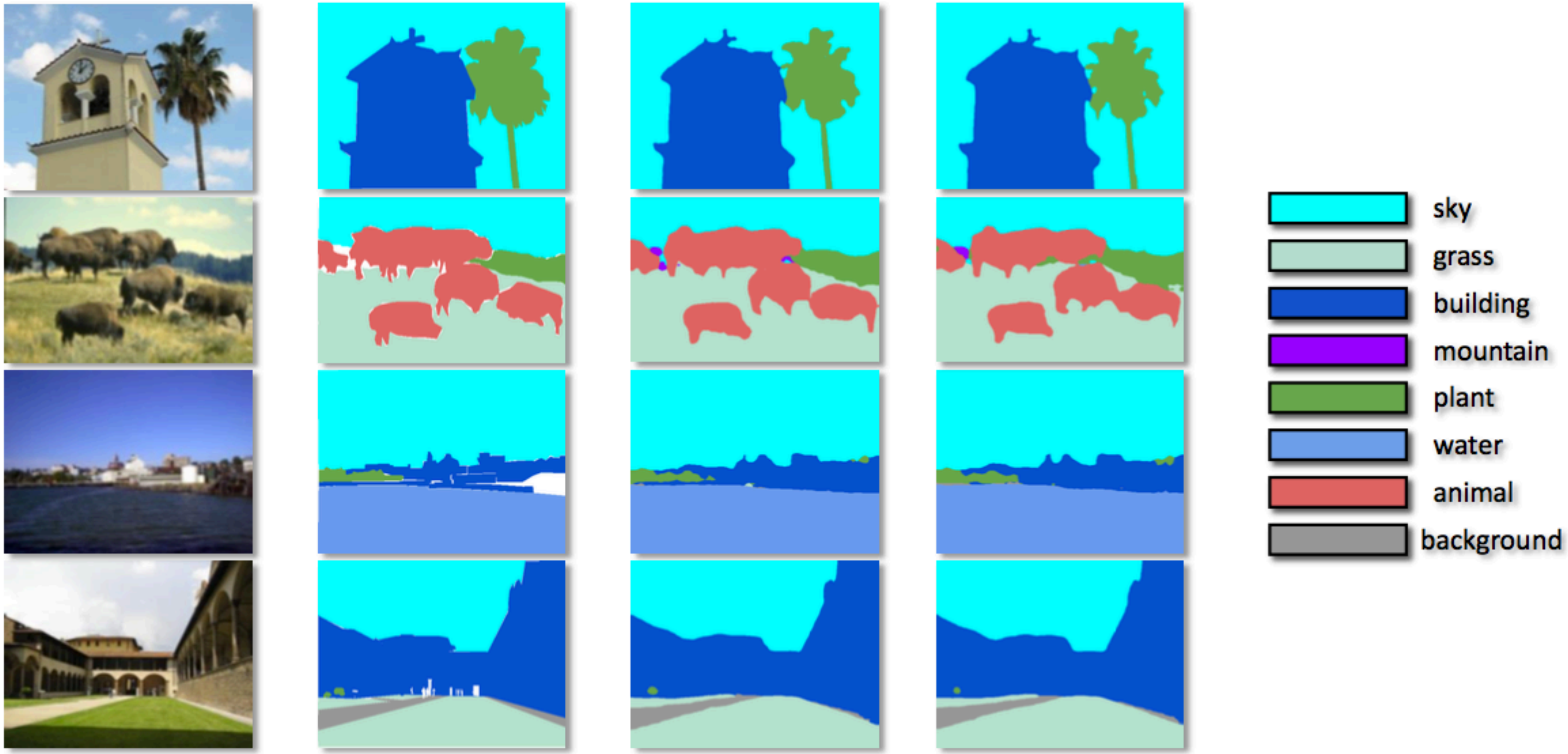


[1]Z. Liu, X. Li, P. Luo, C.-C. Loy, and X. Tang. Semantic image segmentation via deep parsing network. In ICCV, 2015

http://ice.dlut.edu.cn/valse2018/ppt/2018ValseDChao.pdf

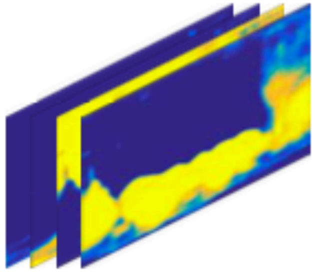
Examples on segmentation

Input LR images | **Ground-truth** | **Segments on HR images** | **Segments on LR images**



Incorporate conditions

Categorical prior



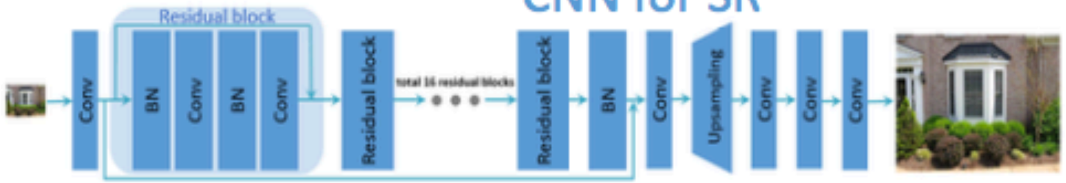
prior Ψ

probability maps $P_1, P_2, \ldots, P_K$

$$\Psi = (P_1, P_2, \ldots, P_K)$$

$$\hat{\boldsymbol{y}} = G_\theta(\boldsymbol{x}|\Psi)?$$

CNN for SR



input LR image $\boldsymbol{x}$

net G *parametrized by* θ

restored image $\hat{\boldsymbol{y}}$

$$\hat{\boldsymbol{y}} = G_\theta(\boldsymbol{x})$$

Spatial Feature Transform

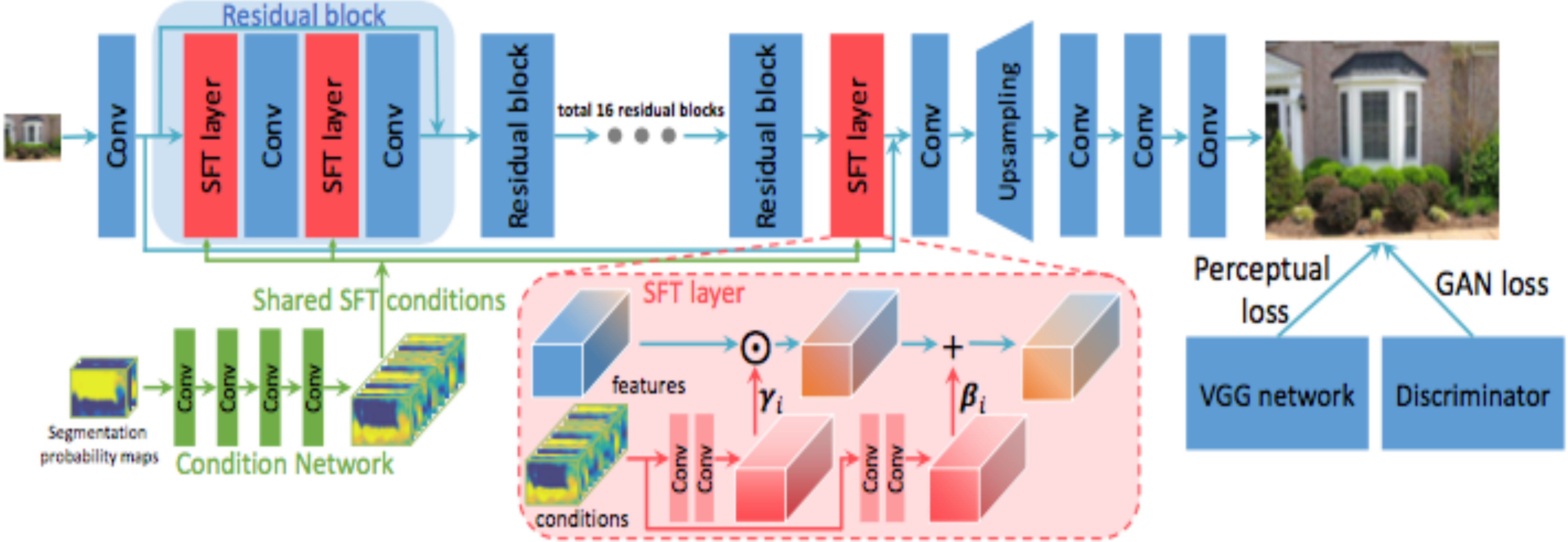
- By learning a mapping function $\mathcal{M}$, the prior Ψ is modeled by a pair of affine transformation parameters (γ, β) .

$$\mathcal{M}: \Psi \mapsto (\gamma, \beta)$$

- The modulation is then carried out by an affine transformation on feature maps $\boldsymbol{F}$. $\mathrm{SFT}(\boldsymbol{F}|\gamma, \beta) = \gamma \odot \boldsymbol{F} + \beta$

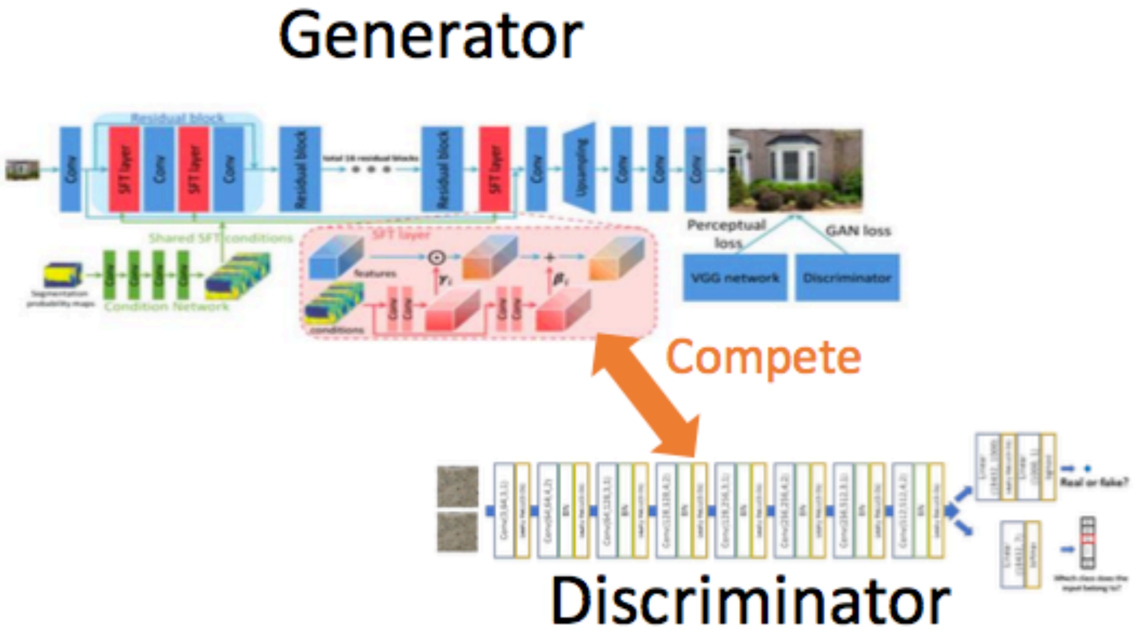
$$\hat{\boldsymbol{y}} = G_\theta(\boldsymbol{x}|\Psi) \xrightarrow[\mathrm{SFT}(\boldsymbol{F}|\gamma, \beta) = \gamma \odot \boldsymbol{F} + \beta]{\mathcal{M}: \Psi \mapsto (\gamma, \beta)} \hat{\boldsymbol{y}} = G_\theta(\boldsymbol{x}|\gamma, \beta)$$

Spatial Feature Transform



http://ice.dlut.edu.cn/valse2018/ppt/2018ValseDChao.pdf

loss function



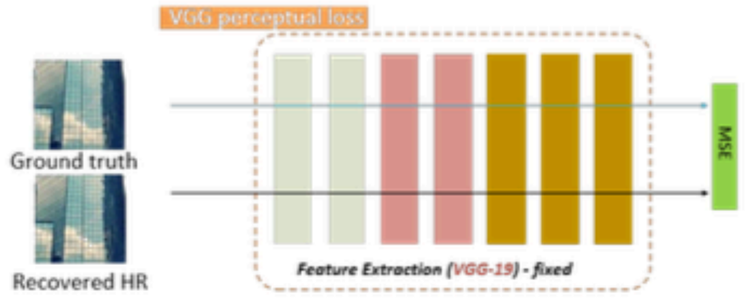
- Adversarial loss[1]
  - ✓ encourage the network to generate images that reside on the manifold of natural images

$$\min_{\theta} \max_{\eta} \mathrm{E}_{y\sim p\mathrm{HR}}\, logD_\eta(y) + \mathrm{E}_{x\sim p\mathrm{LR}} \log(1 - D_\eta(G_\theta(x)))$$

- Perceptual loss[2]

  use a pre-trained 19-layer VGG network (features before conv54)
  - ✓ optimize a super-resolution model in a feature space

$$\|\phi_{VGG}(\hat{y}) - \phi_{VGG}(y)\|_2^2$$



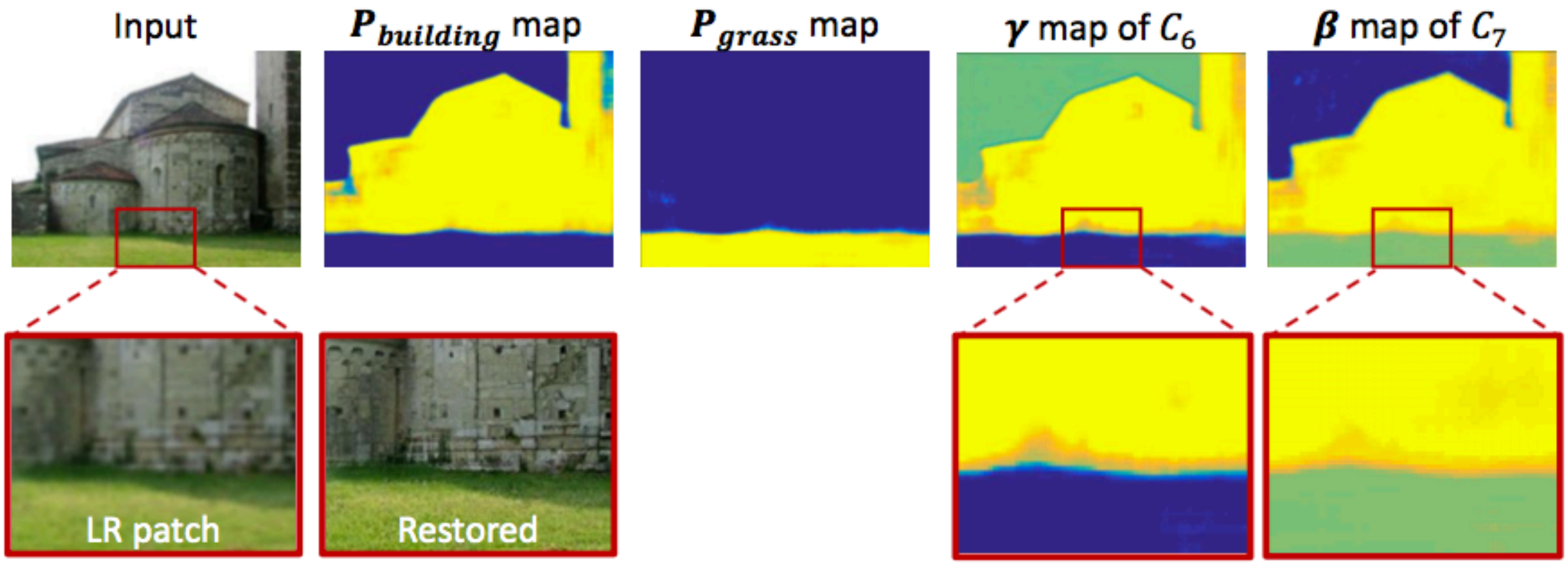
[1]Goodfelow, lan, et al. Generative adversarial nets. In NIPS. 2014

[2] J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In ECCV,2016

http://ice.dlut.edu.cn/valse2018/ppt/2018ValseDChao.pdf

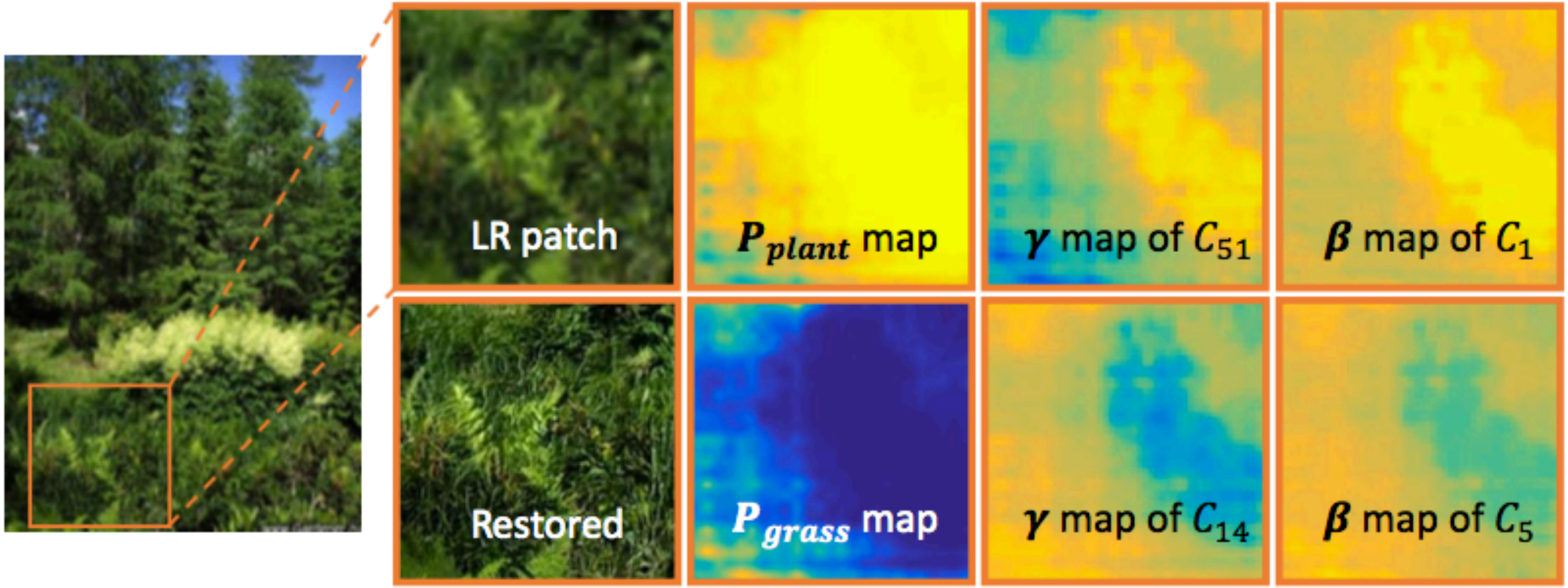
Spatial condition

- The modulation parameters (γ, β) have a close relationship with probability maps $\boldsymbol{P}$ and contain spatial information.



http://ice.dlut.edu.cn/valse2018/ppt/2018ValseDChao.pdf

Delicate modulation

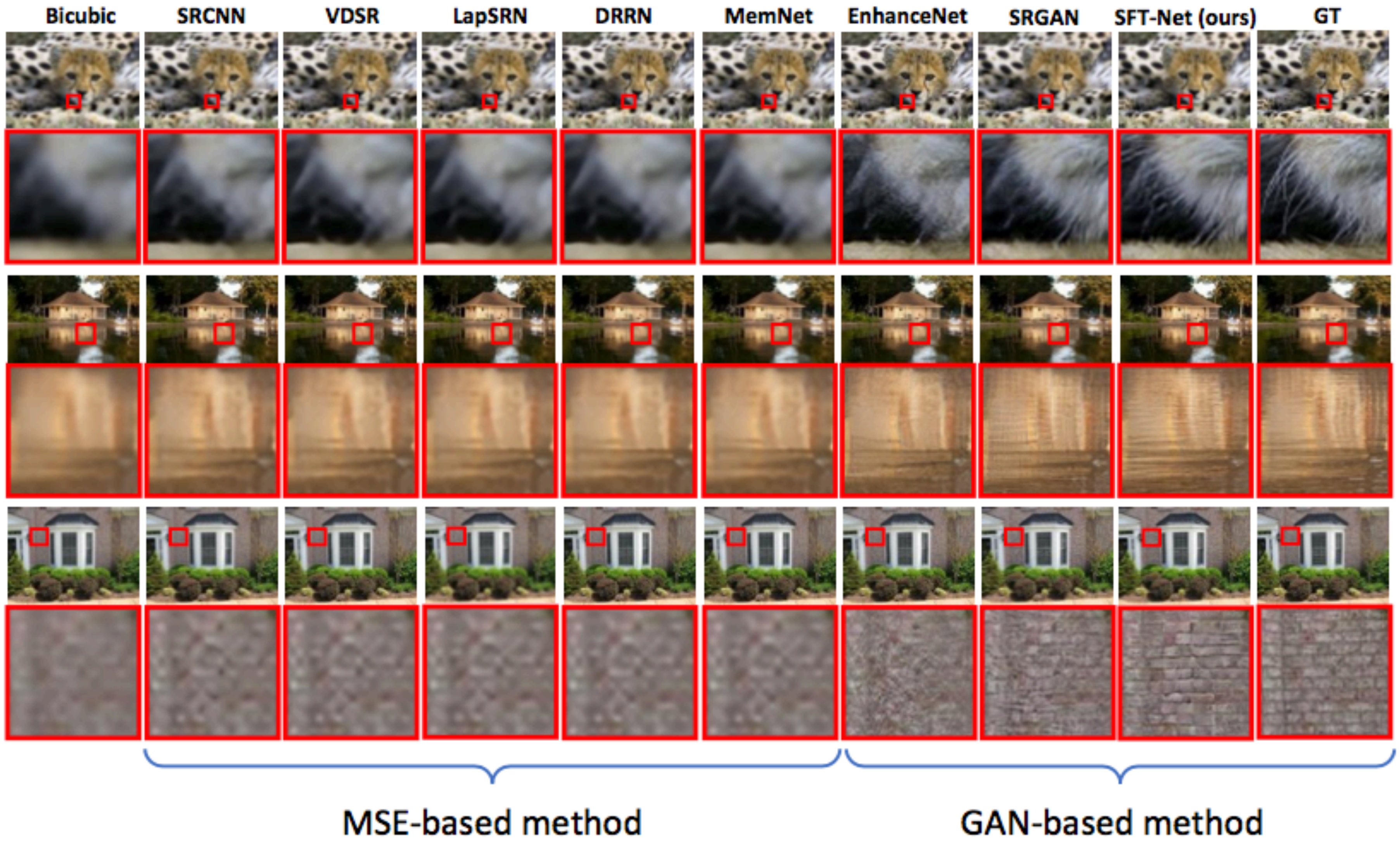


http://ice.dlut.edu.cn/valse2018/ppt/2018ValseDChao.pdf

Results

http://ice.dlut.edu.cn/valse2018/ppt/2018ValseDChao.pdf

Results



http://ice.dlut.edu.cn/valse2018/ppt/2018ValseDChao.pdf

Conclusion

- Explore semantic segmentation maps as categorical prior for realistic texture recovery.
- Propose a novel Spatial Feature Transform layer to efficiently incorporate the categorical conditions into a CNN-based SR network.
- Extensive comparisons and a user study demonstrate the capability of SFT-Net in generating realistic and visually pleasing textures.

05

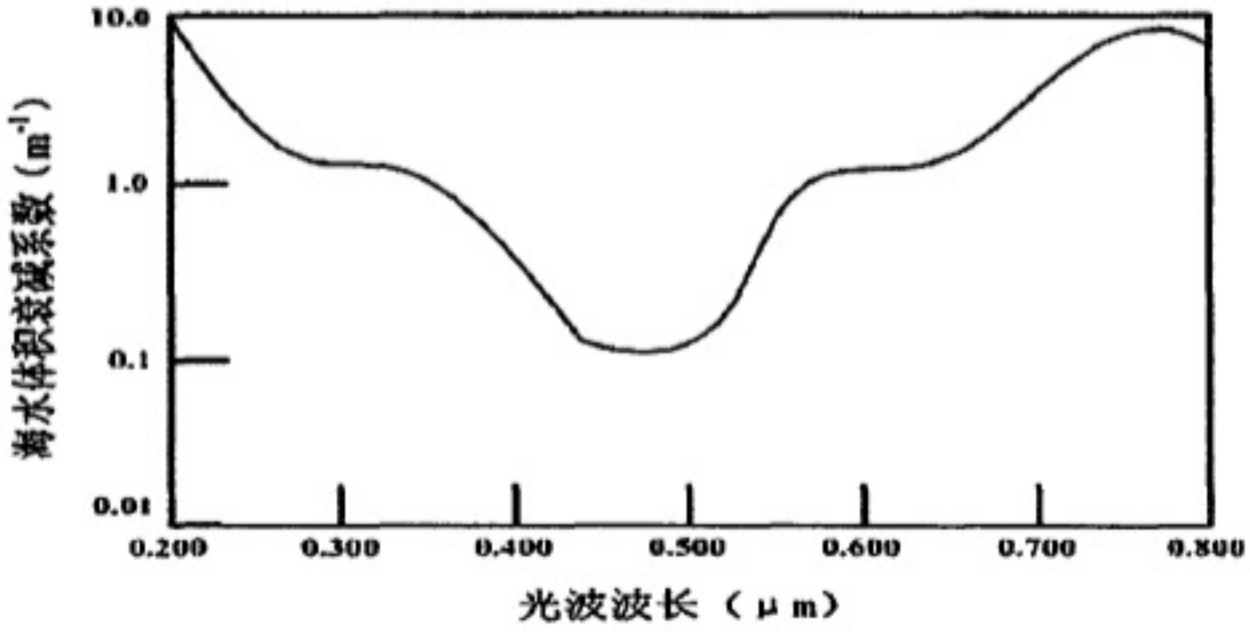
Underwater Image Enhancement and Restoration

Underwater Image Enhancement and Restoration

Underwater Image Enhancement and Restoration

- **Underwater imaging system**

Optical properties of light transmission under water

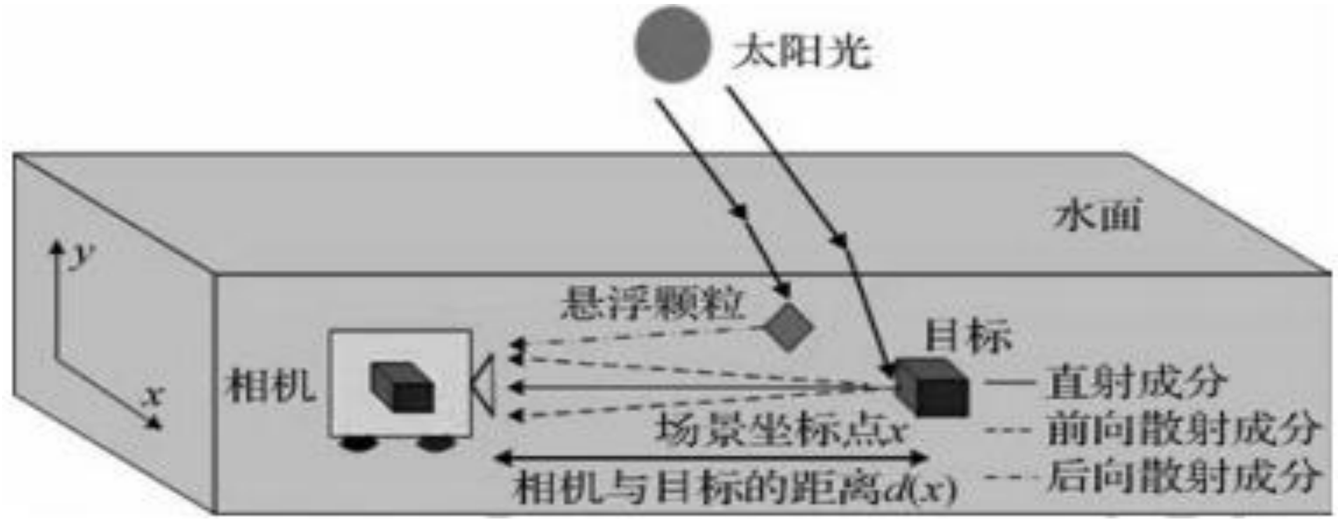


《水下图像复原处理方法的研究》陈琳

Underwater Image Enhancement and Restoration

- **Underwater imaging system**

light scattering： forward scattering
backward scattering



《水下图像增强和复原方法研究进展》
郭继昌、李重仪、郭春乐、陈善继

Underwater Image Enhancement and Restoration

- **The relationship between the quality of underwater imaging and the optical properties of water**

$$\boldsymbol{I = I_0 e^{-\alpha l}}$$

$$\alpha = \mathrm{a} + \mathrm{b}$$

Underwater Image Enhancement and Restoration

- **Use traditional image enhancement methods**

  Spatial domain： Histogram equalization 、White balance

  Frequency domain : Wavelet based color correction

- **Use special underwater image enhancement methods**

  "Underwater color constancy : enhancement of automatic live fish recognition.”

  “Underwater image enhancement by dehazing and color correction. ”

Emerging From Water: Underwater Image Color Correction Based on Weakly Supervised Color Transfer

Chongyi Li , Student Member, IEEE, Jichang Guo, and Chunle Guo

Contribution:

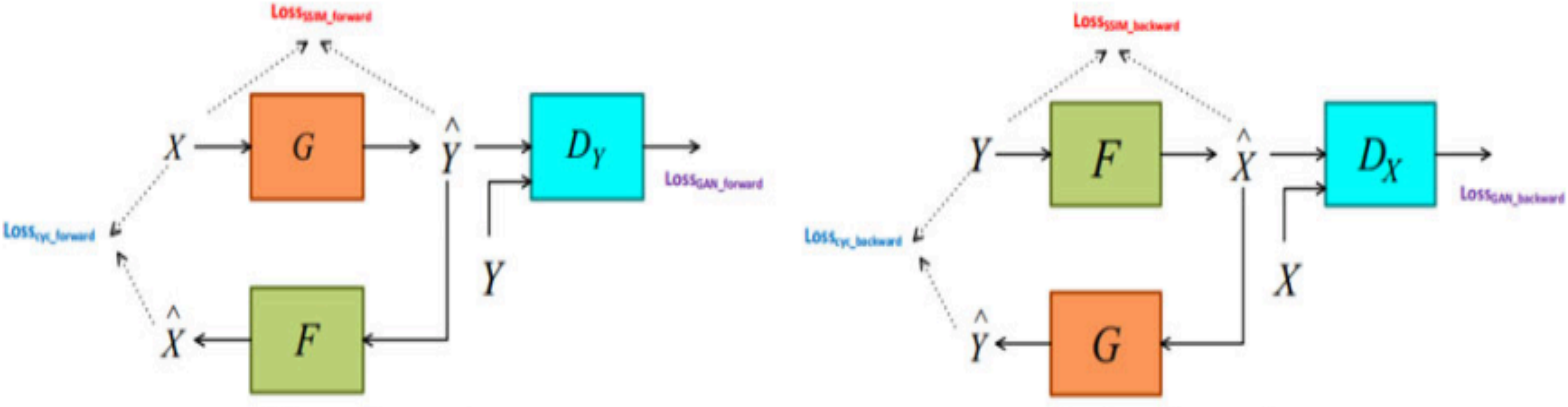
1) Compared with the existing priors/assumptions based and semisupervised methods, to our best knowledge, this is the first attempt to build a weakly supervised model for underwater image color correction. Here, “weak supervision” means that our model relaxes the need for paired underwater images for training and allows the underwater images being taken in unknown locations.

2) We tackle underwater image color correction problem from a new angle, which learns a cross domain mapping function between underwater images and air images.

3) A multiterm loss function allows our model capturing context and semantic information, which makes the content and structure of the outputs same as the inputs, meanwhile the color is similar to the images that were taken without the water.

Underwater Image Enhancement and Restoration

- **Network Architecture and Training Details**



Forward networks Backward networks

- **Loss function**

Adversarial Loss

$$L_{GAN}(G, D_Y, X, Y) = E_{y\sim P_{data}(y)}[(D_Y(y) - 1)^2] + E_{x\sim P_{data}(x)}[D_Y(G(x))^2]$$

Cycle Consistency loss

$$L_{\mathrm{cyc}}(G, F) = E_{x\sim P_{data}(x)}\left[\left\|F\big(G(x)\big) - x\right\|_1\right] + E_{y\sim P_{data}(y)}\left[\left\|G\big(F(y)\big) - y\right\|_1\right]$$

SSIM loss

$$SSIM(p) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \cdot \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

$$L_{SSIM}\big(x, G(x)\big) = 1 - \frac{1}{N}\sum_{p=1}^{N}(SSIM(p))$$

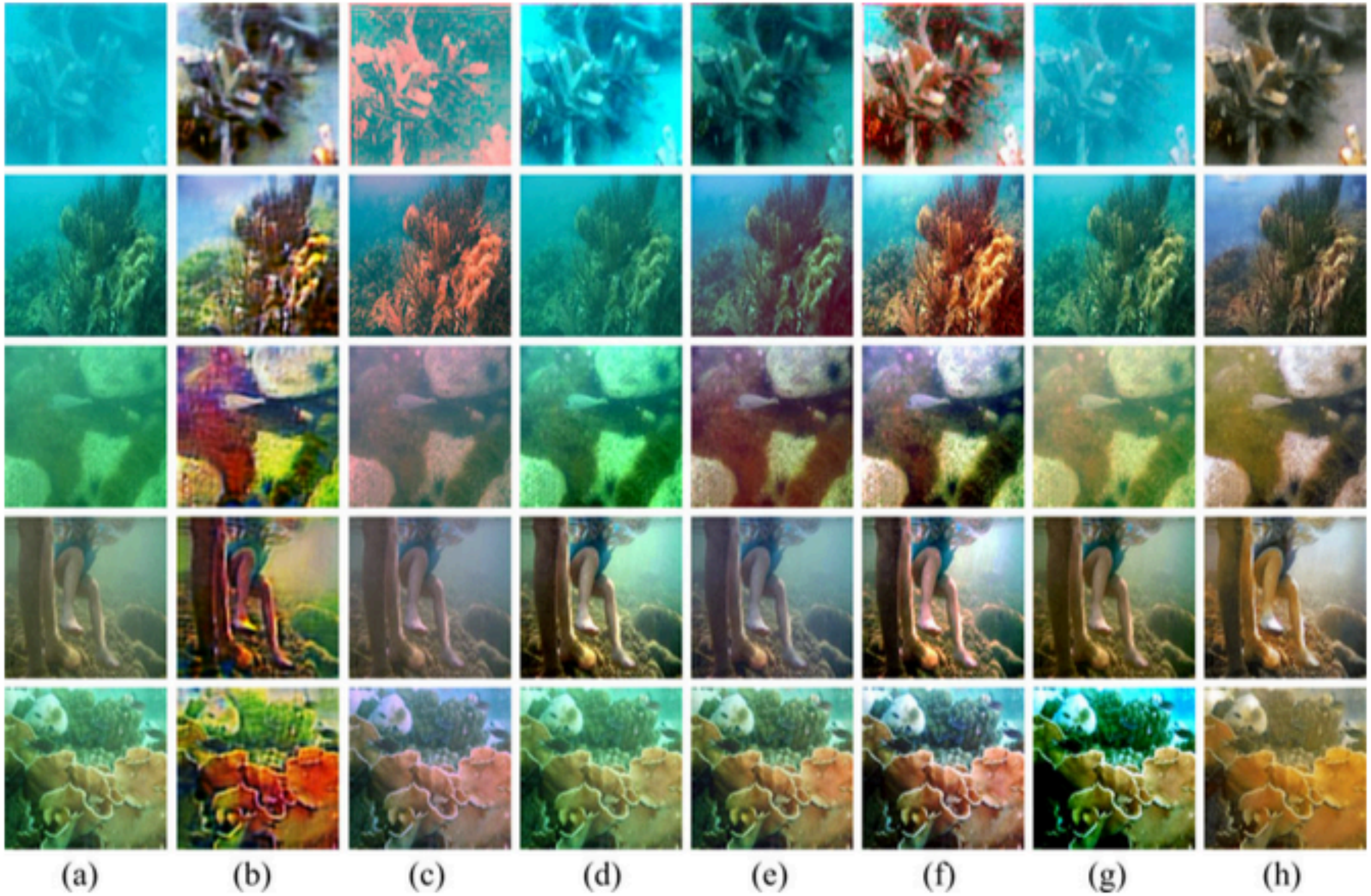


Fig. 4. Visual quality comparisons on varying underwater scenes. From top to bottom are image I1–I5. (a) Raws. (b) CycleGAN. (c) GW. (d) INT. (e) RED. (f) UWID. (g) UWIB. (h) Ours.

Thanks