Paper Reading

—— Image Inpainting

赵燕

2018.03.07

Paper

Image Inpainting using Multi-Scale Feature Image Translation

Yuhang Song， Chao Yang， Zhe Lin， Hao Li， Qin Huang， C.-C.Jay Kuo

USC， Adobe Research， Pinscreen， USC Institute for Creative Technologies

—— from arXiv 2017

Abstract

We study the task of image inpainting, which is to fill in the missing region of an incomplete image with plausible contents. To this end, we propose a learning-based approach to generate visually coherent completion given a high-resolution image with missing components. In order to overcome the difficulty to directly learn the distribution of high-dimensional image data, we divide the task into initialization and texture-refinement as two separate steps and model each step with a deep neural network. We also use simple heuristics to guide transferring of textures from boundary to the hole. We show that, by using such techniques, inpainting reduces to the problem of learning two image-feature translation functions of much smaller dimensionality. We evaluate our method on several public datasets and show that we not only generate results of comparable or better visual quality, but are orders of magnitude faster than previous state-of-the-art methods.

Contributions

- High-resolution image with high-quality contents and textures
- Addresses the issue of noisy input and avoids underfitting
- The trained model can be directly used on other tasks like style transfer and achieve performance comparable with state-of-the-art

(a) Input image (b) Initialization (c) Final result (d) Neural inpainting's result

Figure 1. Our result comparing with neural inpainting [35].
(a)The input image with missing hole.
(b) Initialization given by the Image2Feature network.
(c) Final inpainting result using our approach.
(d) Inpainting result given by neural inpainting [35]. The size of images are 512x512.
.

Steps

Step 1. **Initialization**: Image2Feature network

Step 2. **Neural patch transfer**: patch-swap

Step 3. **Image reconstruction**: Feature2Image network

Network Architecture

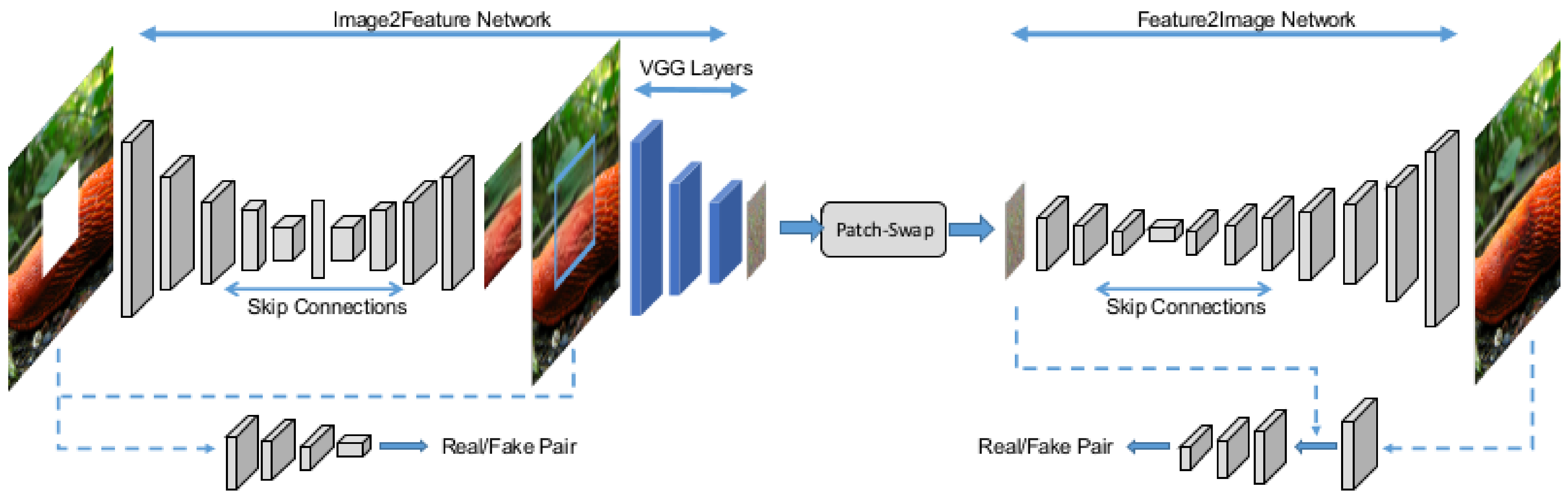


Figure 2. Overview of our network architecture. We use Image2Feature network as coarse initialization and use VGG network to extract a feature map. Then patch-swap transfers neural patches from boundary to the hole. Finally the Feature2Image network reconstructs a complete, high-resolution image.

Patch-swap

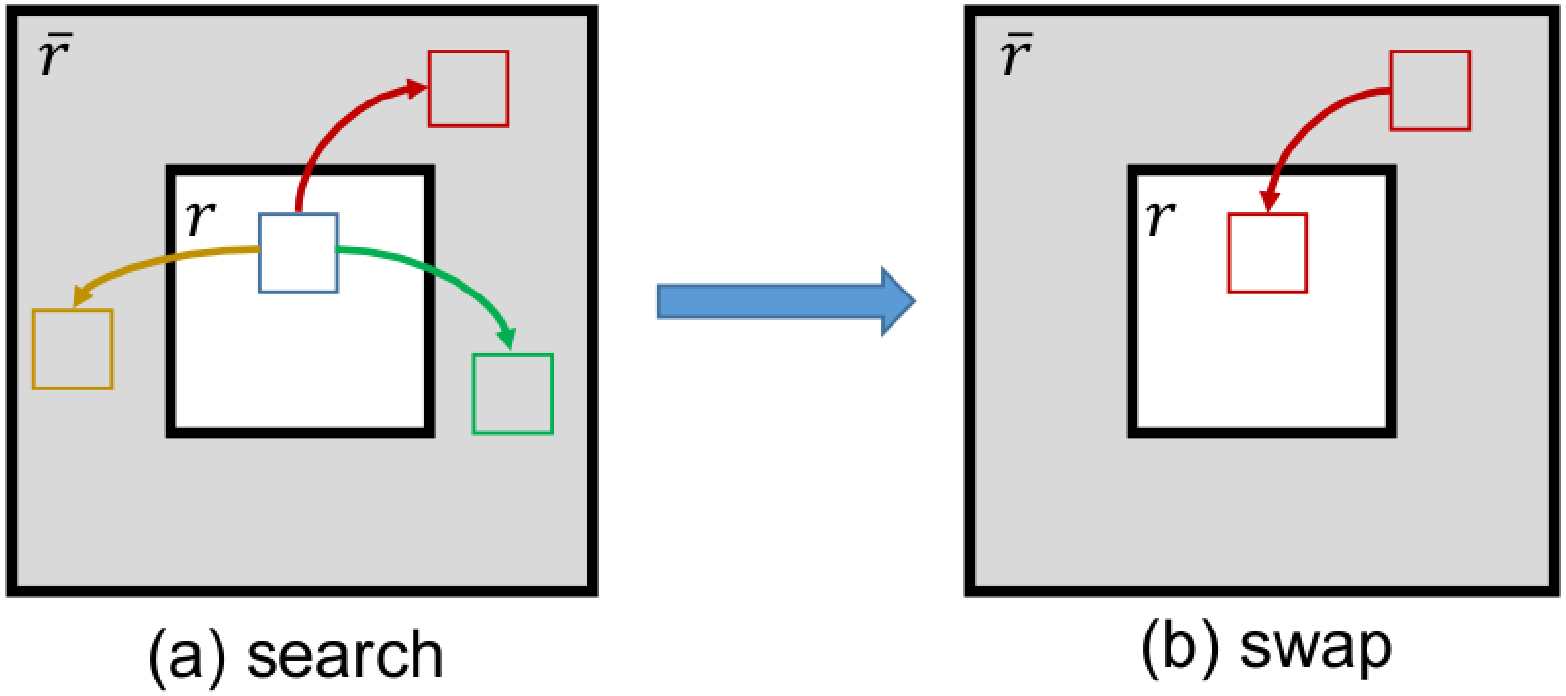


Figure 3. Illustration of patch-swap operation. Each neural patch in the hole r searches for the most similar neural patch on the boundary $\bar{r}$, and then swaps with that patch.

Multi-Scale

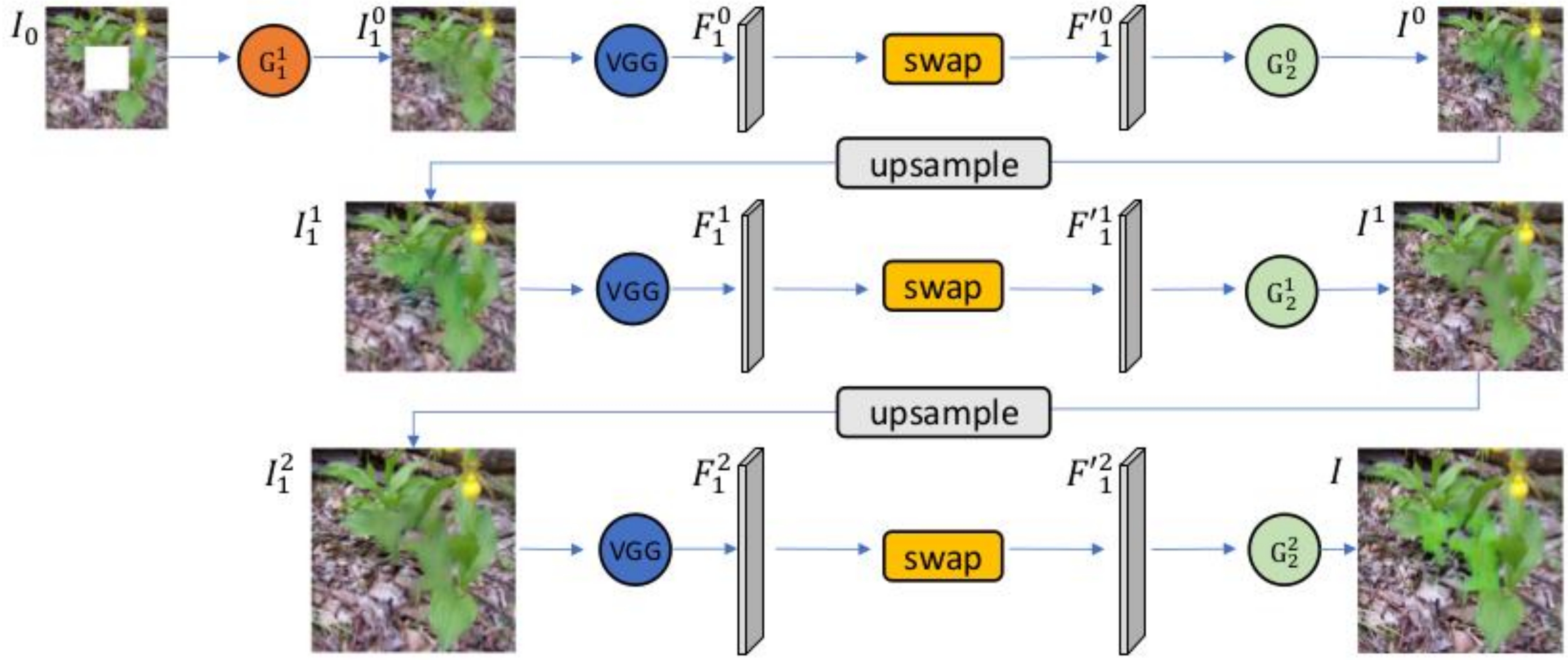


Figure 4. Illustration of multi-scale inference.

Quantitative Comparison

Method	Mean ℓ_1 Error	SSIM
Context Encoder [27]	15.46%	0.87
Neural inpainting [35]	**15.13%**	0.88
our approach	15.61%	**0.89**

Table 1. Numerical comparison on 200 test images of ImageNet.

Comparison

(a) (b) (c) (d)

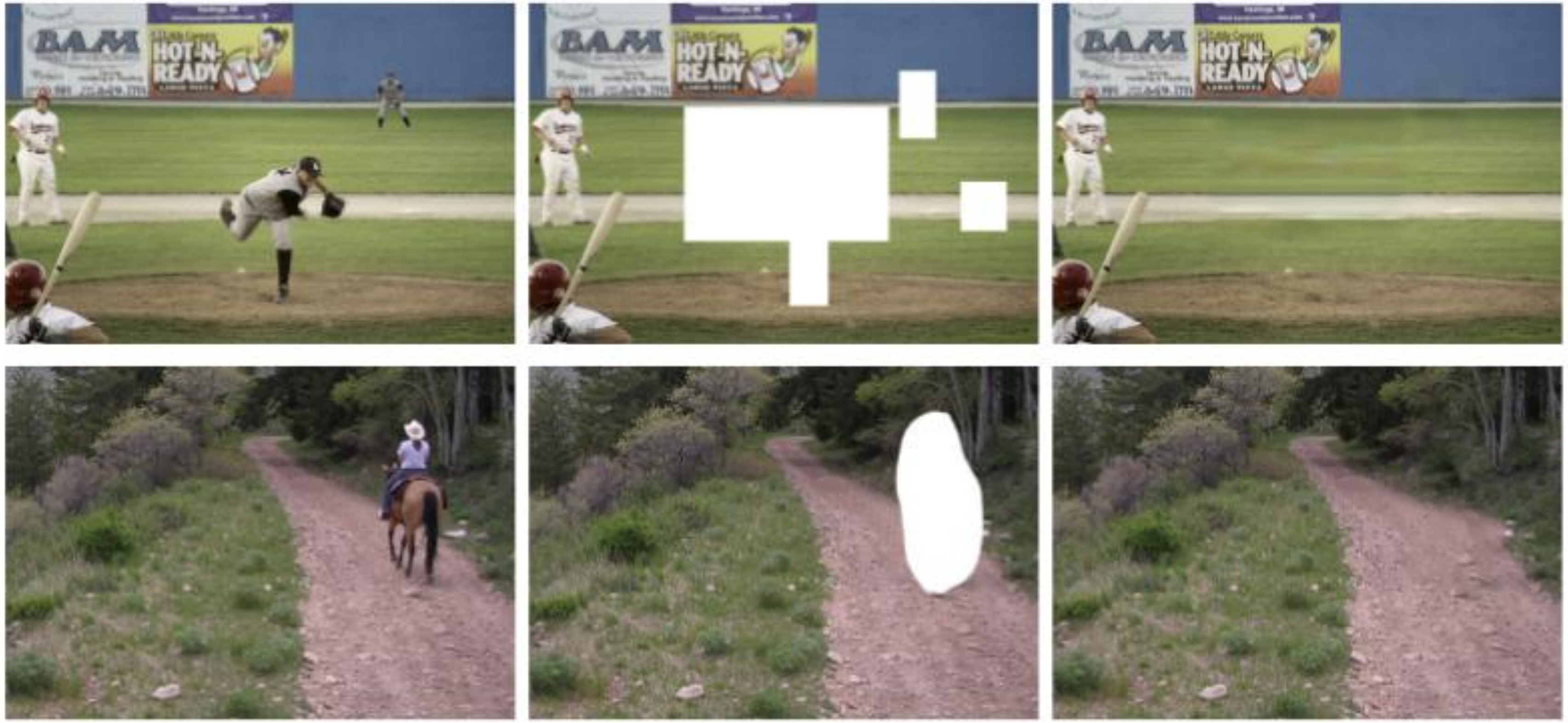
Figure 5. Our results ((b) & (d)) comparing with [35]'s results ((a) & (c)) on 256x256 images (input not shown). We can see that our inpainting is much sharper at this scale.

Comparison

(a) CE (b) ours (c) [16] (d) ours

Figure 6. Our results ((b) & (d)) comparing with context encoder (a) and [16] (c) (input not shown). Originally (a) & (b) are 128x128 and (c) & (d) are 256x256.

Comparison



(a) (b) (c)

Figure 9. Arbitrary shape inpainting of real-world photography. (a) Input. (b) Inpainting mask. (c) Output.

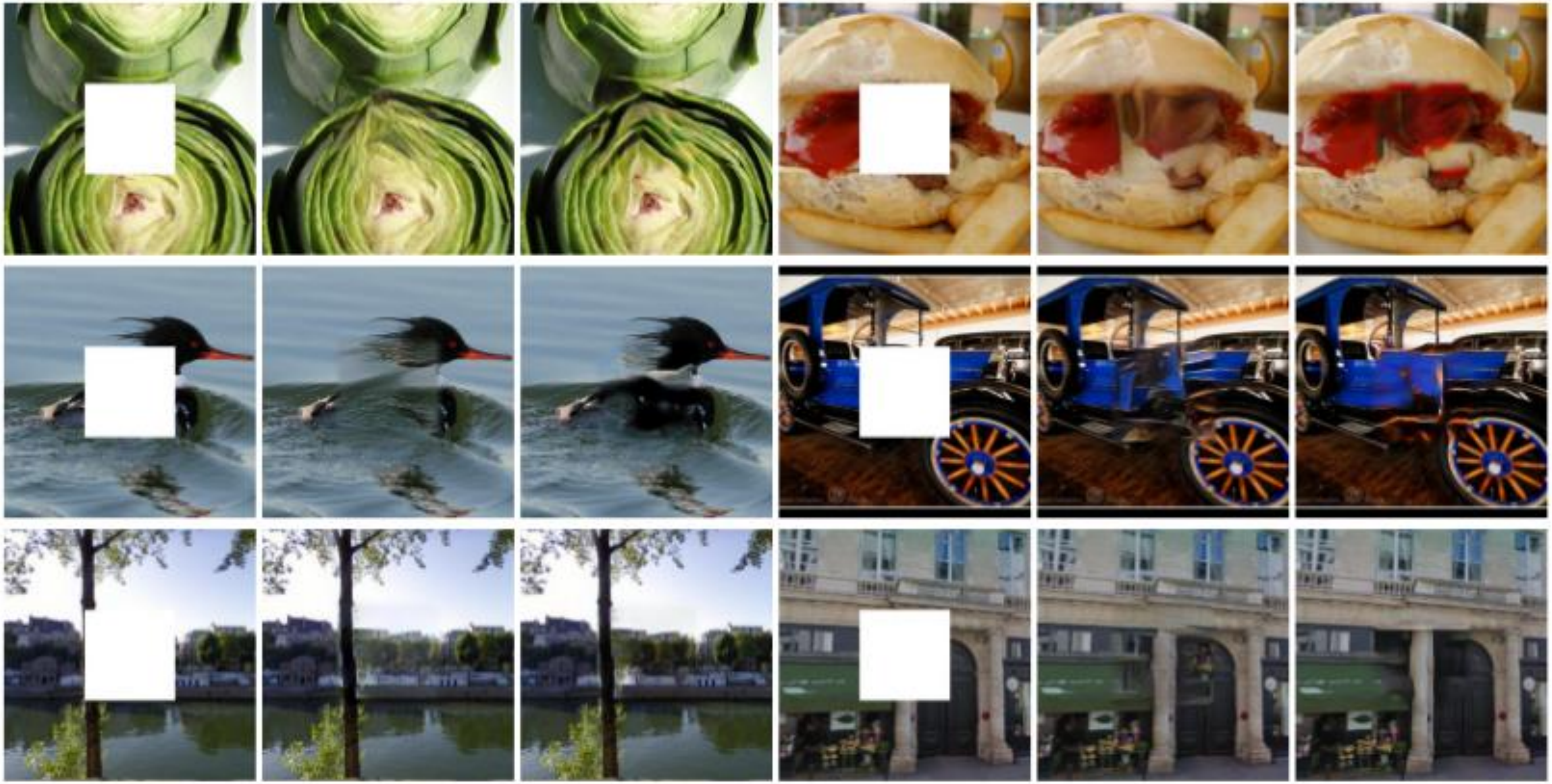


Figure 12. Visual comparisons of Paris StreetView and ImageNet results. Each example from left to right: input image, neural inpainting's result, our result. All images have size 512×512.

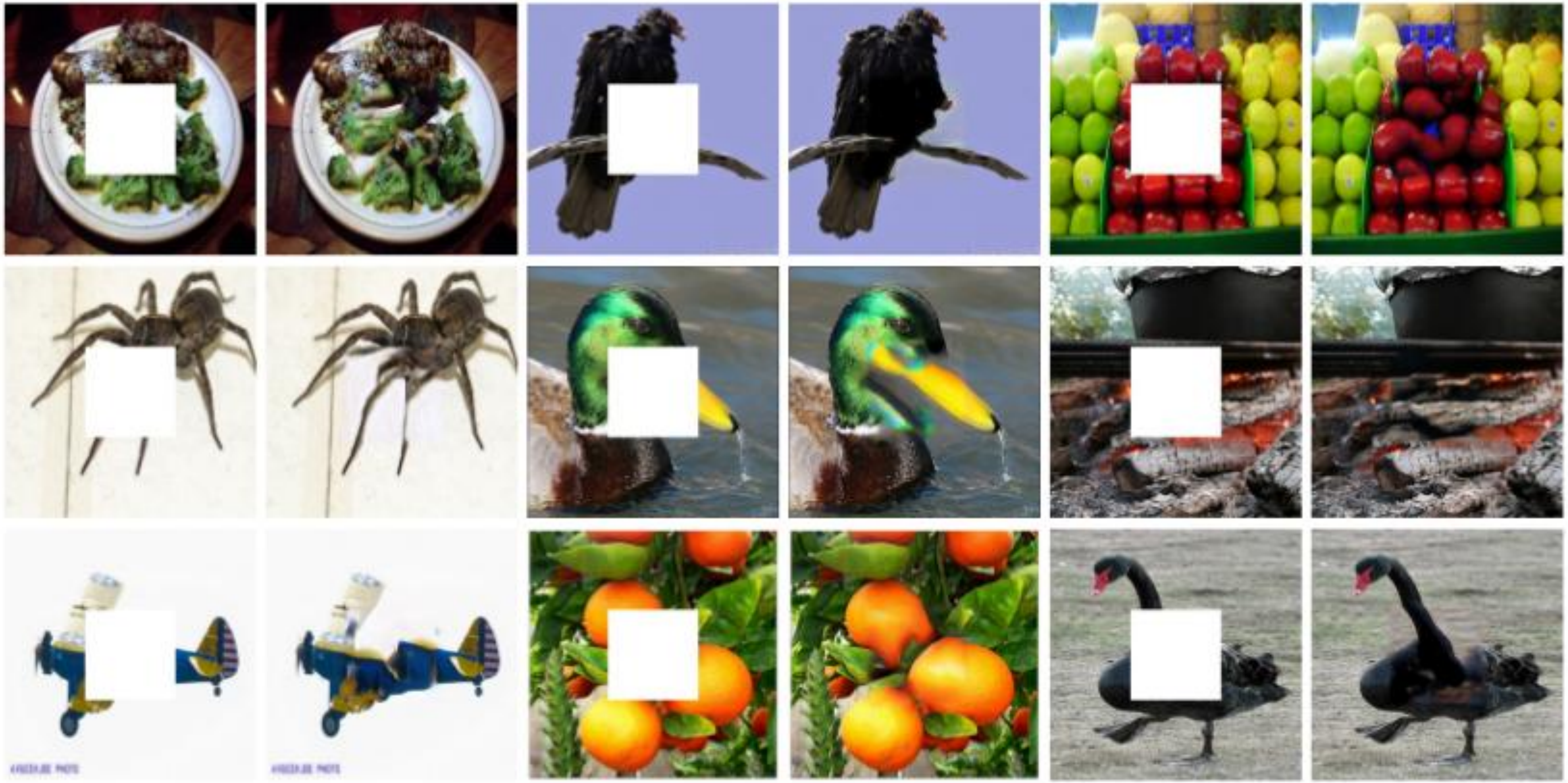


Figure 13. More ImageNet and COCO results. All images have size 512×512.

Comparison

(a) Input (b) Initialization (c) Inpainting

Figure 11. A failure case where the main part of an object is missing.

Thank you !