

# Image Super-Resolution

Tuan Phan, Tung Pham  
AI For Everyone

Ngày 3 tháng 3 năm 2022

# Outline

---

## 1. Bài toán image super-resolution

## 2. Image Interpolation

- 2.1 Nearest-neighbor interpolation
- 2.2 Bilinear interpolation
- 2.3 Bicubic interpolation

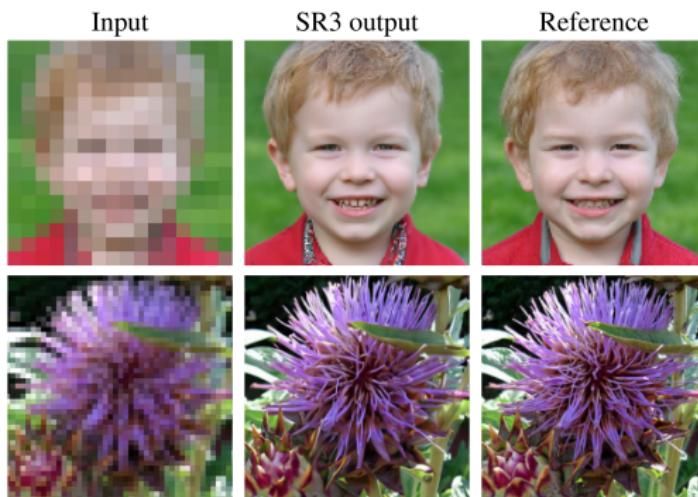
## 3. CNN for Super-Resolution

- 3.1 SRCNN
- 3.2 ResNet Super Resolution
- 3.3 Autoencoder Super Resolution

## 4. Tổng kết

# Bài toán image super-resolution

Cho ảnh  $X$  là một ảnh high-resolution (HR),  $Y$  là ảnh low-resolution (LR) tương ứng của  $X$ . Gọi hàm dùng để upscale ảnh LR thành HR là  $F$ . Bài toán image super-resolution hướng đến việc tìm  $F$  sao cho  $F(Y) \approx X$ . Hay ảnh upscale từ  $Y$  giống  $X$  nhất có thể



Hình: Ví dụ về bài toán image super-resolution

# Dataset

---

Bộ dữ liệu sử dụng: DIV2K

DIV2K bao gồm 1000 ảnh 2K trong đó:

- 800 ảnh cho training
- 100 ảnh cho validation
- 100 ảnh cho testing

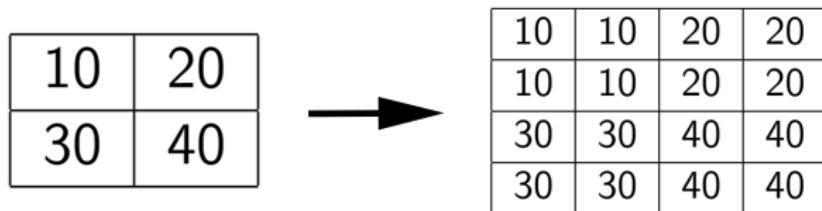
**Nguồn:** <https://data.vision.ee.ethz.ch/cvl/DIV2K/>

# Image Interpolation

# Image Interpolation

---

- Nearest-neighbor interpolation: là phương pháp đơn giản nhất. Trong phương pháp này, các pixel trong ảnh  $F(Y)$  sẽ dùng giá trị của pixel trong ảnh  $Y$  gần nó nhất.



Hình: Nearest-neighbor interpolation

# Image Interpolation

---

- Bilinear interpolation: phương pháp này sẽ nội suy giá trị của một pixel bằng cách tính trung bình có trọng số 4 (2x2) pixel lân cận.

10	20		
30	40		

→

10	12.5	17.5	20
15	17.5	22.5	25
25	27.5	32.5	35
30	32.5	37.5	40

Hình: Bilinear interpolation

# Image Interpolation

---

- Bicubic interpolation: phương pháp này, các pixel trong ảnh  $F(Y)$  sẽ dùng giá trị của pixel trong ảnh  $Y$  gần nó nhất.

10	20		
30	40		

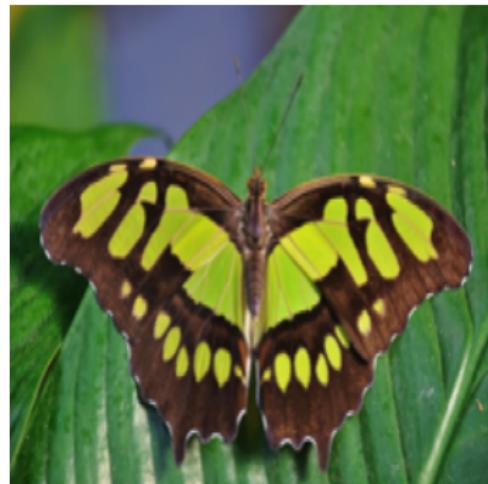
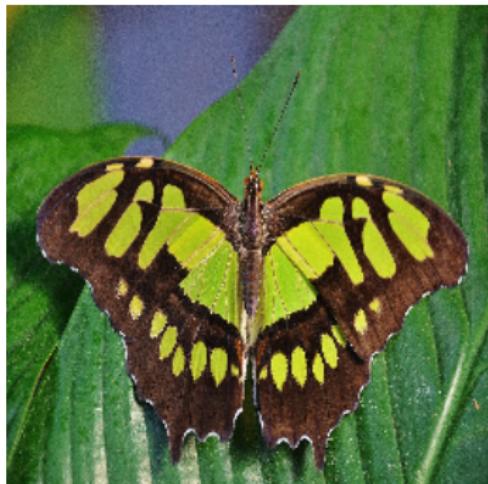
→

6.84	10.16	15.63	18.95
13.48	16.80	22.27	25.59
24.41	27.73	33.20	36.52
31.05	34.38	39.84	43.16

Hình: Bicubic interpolation

# Image Interpolation

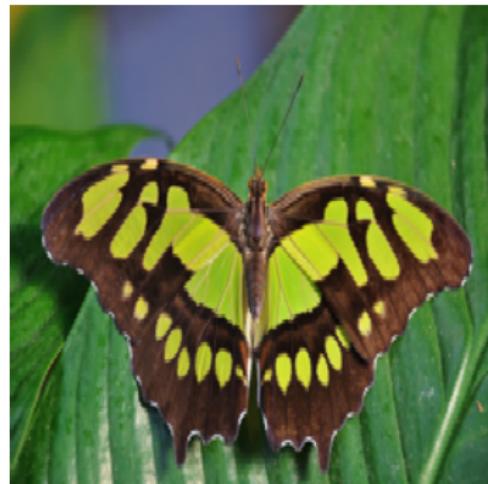
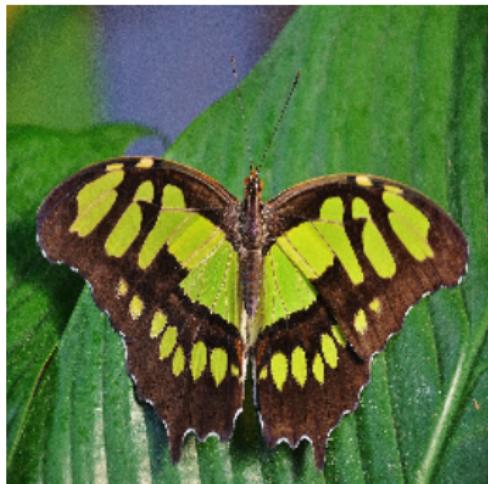
---



Hình: Ảnh high-resolution và low-resolution

# Image Interpolation

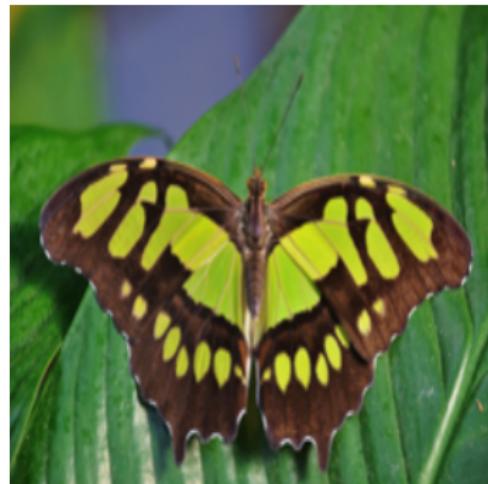
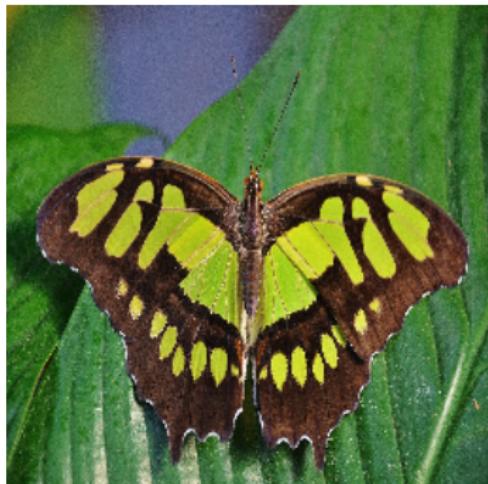
---



Hình: Ảnh high-resolution và Nearest-neighbor interpolation

# Image Interpolation

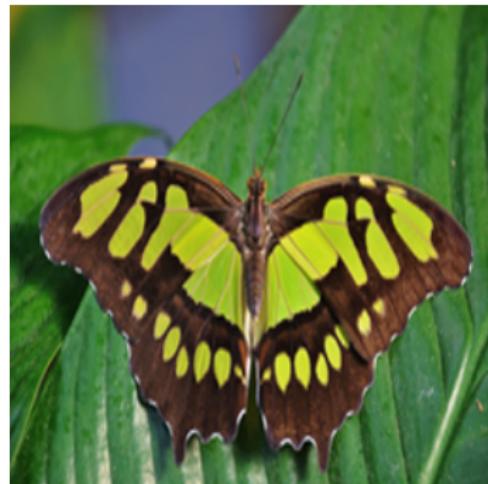
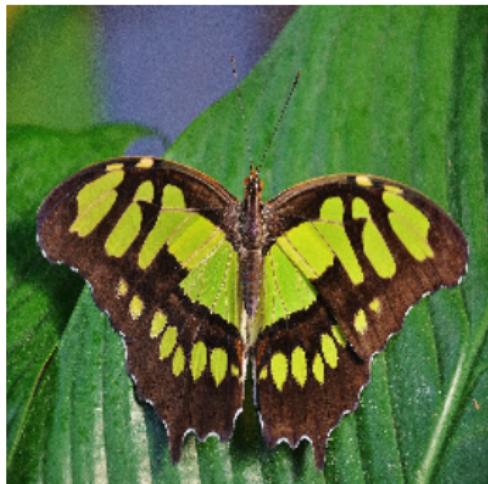
---



Hình: Ảnh high-resolution và Bilinear interpolation

# Image Interpolation

---



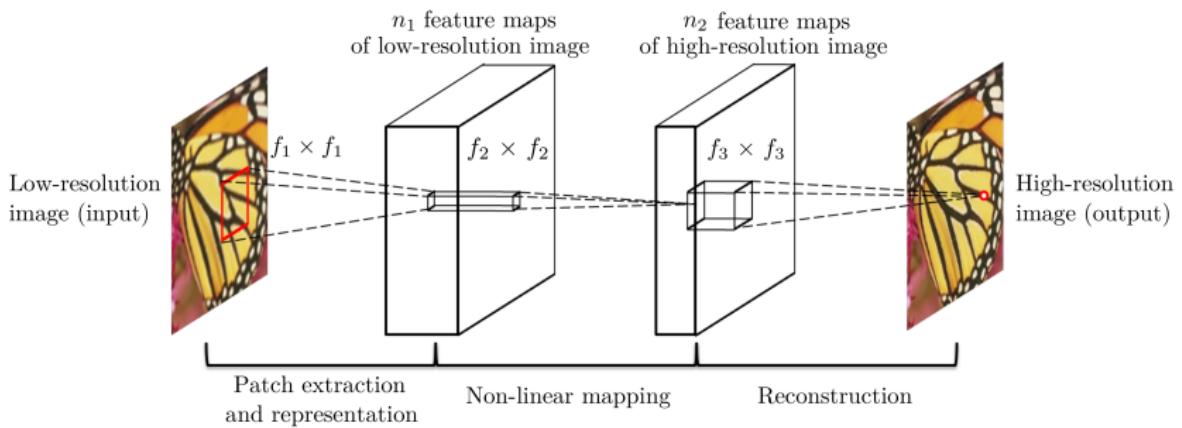
Hình: Ảnh high-resolution và Bicubic interpolation

# CNN for Super-Resolution

# SRCNN

---

# Kiến trúc mạng



Hình: Kiến trúc mạng SRCNN

# Patch extraction and representation

---

Lấy các patch overlap nhau ở input bằng cách cho một kernel trượt trên ảnh đầu vào. Số lượng feature map sẽ tương ứng với số chiều của vector. Phép toán được thực hiện ở layer đầu tiên là:

$$F_1(\mathbf{Y}) = \max(0, W_1 \circledast \mathbf{Y} + B_1)$$

Trong đó:

- $W_1$  là filter, kích thước  $f_1 \times f_1 \times c$ , có  $n_1$  filter như này.
- $B_1$  là bias
- $\circledast$  là phép conv

Output của layer này là một vector  $n_1$  chiều tương ứng với  $n_1$  feature map.

# Non-linear mapping

---

Ánh xạ vector  $n_1$  chiều sang vector  $n_2$  chiều

$$F_2(\mathbf{Y}) = \max(0, W_2 \circledast F_1(\mathbf{Y}) + B_2)$$

Trong đó:

- $W_2$  là filter, kích thước  $f_2 \times f_2 \times n_1$ , có  $n_2$  filter như này.
- $B_2$  là bias

**Câu hỏi đặt ra:** Có nên thêm nhiều layer này để tăng tính non-linear?

# Reconstruction

---

Phục hồi ảnh từ vector  $n_2$  chiều

$$F_3(\mathbf{Y}) = W_3 \circledast F_2(\mathbf{Y}) + B_3$$

Trong đó:

- $W_3$  là filter, kích thước  $n_2 \times f_3 \times n_3$ , có  $c$  filter như này.
- $B_3$  là bias

# Training

---

Loss function:

$$L(\Theta) = \frac{1}{n} \sum_{i=1}^n \|F(\mathbf{Y}_i; \Theta) - \mathbf{x}_i\|^2$$

Với n là số sample được dùng để train

# Kết quả

---



Hình: Ảnh high-resolution, low-resolution và output SRCNN

# ResNet Super Resolution

---

# Kết quả

---



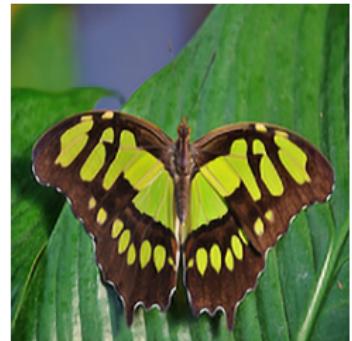
Hình: Ảnh high-resolution, low-resolution và output SRResNet

# Autoencoder Super Resolution

---

# Kết quả

---



Hình: Ảnh high-resolution, low-resolution và output Autoencoder

# Tổng kết

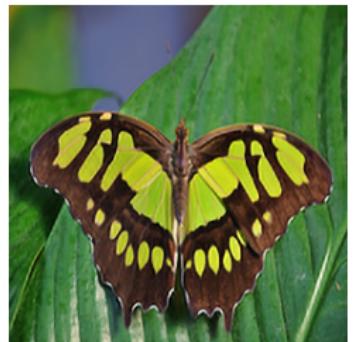
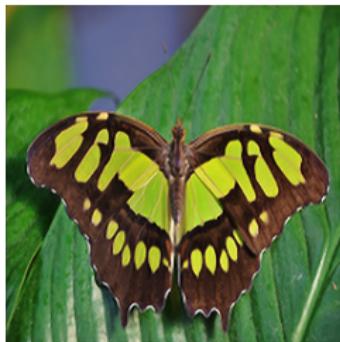
---

Kết quả đo được từ tập validation

Tên	Parameters	Loss (MSE)	PSNR	Thời gian
SRCNN	20,483	594.2112	21.0903	0.232658s
SRResnet	449,731	386.5698	23.4264	0.287240s
Autoencoder	1,848,067	400.6670	23.2711	0.718193s

# Tổng kết

---



Hình: SRCNN, SRResnet và Autoencoder

# Một số ảnh khác

---



Hình: Ảnh high-resolution và low-resolution

# Một số ảnh khác

---



Hình: SRCNN, SRResnet và Autoencoder

# Một số ảnh khác

---



Hình: Ảnh high-resolution và low-resolution

# Một số ảnh khác

---



Hình: SRCNN, SRResnet và Autoencoder

# Reference

---

-  Image Super-Resolution Using Deep Convolutional Networks
-  Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network
-  Super-Resolution Using Autoencoders and TF2.0