

# Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network

- <https://arxiv.org/pdf/1609.05158>

## 概要

- pixelshuffleを初めて提案した論文

## architecture

- CNNからsub-pixel convolution layer

$$\begin{aligned} f^l(I^{LR}; W_{1:l}, b_{1:l}) &= \phi(W_l * f^{l-1}(I^{LR}) + b_l) \\ I^{SR} &= f^L(I^{LR}) = \mathcal{PS}(W_L * f^{L-1}(I^{LR}) + b_L) \end{aligned} \quad (1)$$

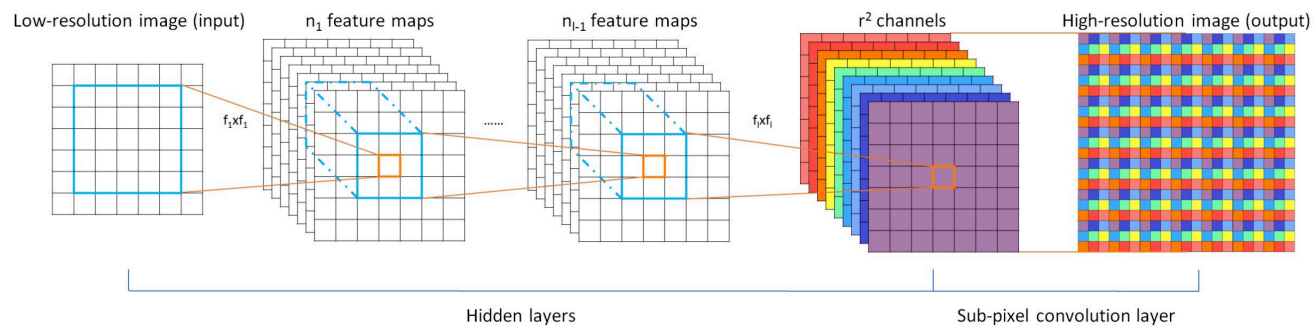
- $L$ 層のCNNを通して  $I^{LR} \in \mathbb{R}^{H \times W \times C}$  から feature map  $f \in \mathbb{R}^{H \times W \times r^2 C}$  を得る
- これをsub pixel convolution layer(pixel shuffle)に入れると  $I^{HR} \in \mathbb{R}^{rH \times rW \times C}$  となる

## downscale

- $I^{HR}$  から  $I^{LR}$  を作成する方法
- Gaussian filterを使ったあとにdownsample

## sub-pixel convolution layer

- pixel shuffle
- $\mathcal{PS} : \mathbb{R}^{H \times W \times r^2 C} \rightarrow \mathbb{R}^{rH \times rW \times C}$
- 配置方法については画像を見ればわかる



- 数学的には以下

$$\mathcal{PS}(T)_{x,y,c} = T_{\lfloor \frac{x}{r} \rfloor, \lfloor \frac{y}{r} \rfloor, C \cdot r \cdot \text{mod}(y,r) + C \cdot \text{mod}(x,r) + c}$$