

Deep learning Lab 02

109610025 陳品妍

Task 1

Epoch 39	Train Loss: 0.0044	Val Loss: 0.0043	PSNR: 23.66 dB
model saved			
Epoch 40	Train Loss: 0.0044	Val Loss: 0.0032	PSNR: 25.01 dB
Epoch 41	Train Loss: 0.0042	Val Loss: 0.0049	PSNR: 23.12 dB
Epoch 42	Train Loss: 0.0043	Val Loss: 0.0041	PSNR: 23.83 dB
Epoch 43	Train Loss: 0.0045	Val Loss: 0.0043	PSNR: 23.70 dB
Epoch 44	Train Loss: 0.0042	Val Loss: 0.0042	PSNR: 23.80 dB
model saved			
Epoch 45	Train Loss: 0.0042	Val Loss: 0.0030	PSNR: 25.23 dB
Epoch 46	Train Loss: 0.0040	Val Loss: 0.0032	PSNR: 24.98 dB
Epoch 47	Train Loss: 0.0042	Val Loss: 0.0032	PSNR: 25.01 dB
Epoch 48	Train Loss: 0.0042	Val Loss: 0.0032	PSNR: 24.90 dB
Epoch 49	Train Loss: 0.0040	Val Loss: 0.0038	PSNR: 24.15 dB
Epoch 50	Train Loss: 0.0040	Val Loss: 0.0053	PSNR: 22.79 dB

Best PSNR: 25.23dB

Task 2

使用 ESPCN model

Epoch 25	Train Loss: 0.0040	Val Loss: 0.0056	PSNR: 22.53 dB"
Epoch 26	Train Loss: 0.0038	Val Loss: 0.0039	PSNR: 24.08 dB"
Epoch 27	Train Loss: 0.0038	Val Loss: 0.0044	PSNR: 23.53 dB"
Epoch 28	Train Loss: 0.0037	Val Loss: 0.0045	PSNR: 23.48 dB"
model saved			
Epoch 29	Train Loss: 0.0037	Val Loss: 0.0022	PSNR: 26.55 dB"
Epoch 30	Train Loss: 0.0037	Val Loss: 0.0025	PSNR: 26.08 dB"
Epoch 31	Train Loss: 0.0037	Val Loss: 0.0023	PSNR: 26.48 dB"
Epoch 32	Train Loss: 0.0037	Val Loss: 0.0030	PSNR: 25.21 dB"
Epoch 33	Train Loss: 0.0036	Val Loss: 0.0032	PSNR: 24.97 dB"
Epoch 34	Train Loss: 0.0037	Val Loss: 0.0049	PSNR: 23.08 dB"
Epoch 35	Train Loss: 0.0037	Val Loss: 0.0050	PSNR: 23.00 dB"
Epoch 36	Train Loss: 0.0036	Val Loss: 0.0030	PSNR: 25.20 dB"
Epoch 37	Train Loss: 0.0036	Val Loss: 0.0037	PSNR: 24.32 dB"
Epoch 38	Train Loss: 0.0036	Val Loss: 0.0029	PSNR: 25.35 dB"
Epoch 39	Train Loss: 0.0035	Val Loss: 0.0033	PSNR: 24.84 dB"
Epoch 40	Train Loss: 0.0036	Val Loss: 0.0039	PSNR: 24.06 dB"
Epoch 41	Train Loss: 0.0035	Val Loss: 0.0028	PSNR: 25.47 dB"
Epoch 42	Train Loss: 0.0036	Val Loss: 0.0031	PSNR: 25.15 dB"
Epoch 43	Train Loss: 0.0035	Val Loss: 0.0039	PSNR: 24.13 dB"
Epoch 44	Train Loss: 0.0035	Val Loss: 0.0027	PSNR: 25.72 dB"
Epoch 45	Train Loss: 0.0034	Val Loss: 0.0041	PSNR: 23.90 dB"
Epoch 46	Train Loss: 0.0035	Val Loss: 0.0035	PSNR: 24.61 dB"
Epoch 47	Train Loss: 0.0034	Val Loss: 0.0028	PSNR: 25.50 dB"
Epoch 48	Train Loss: 0.0035	Val Loss: 0.0037	PSNR: 24.26 dB"
Epoch 49	Train Loss: 0.0034	Val Loss: 0.0031	PSNR: 25.03 dB"
model saved			
Epoch 50	Train Loss: 0.0034	Val Loss: 0.0017	PSNR: 27.74 dB"

最佳: 27.74 Db

使用的是論文中提到的 ESPCN MODEL with upscale factor = 4 來改善 PSNR。

優點: 使用 sub-pixel convolution 進行圖像的 reconstruction，比 deconvolution 減少計算複雜度。因為 sub-pixel convolution 是在 low resolution 的 feature map 進行 convolution，而 deconvolution 則需要計算每個 pixel-pixel 的 convolution 值。因此，sub-pixel convolution 比 deconvolution 處理更少計算。另外，在高放大倍率(upscale factor)可以具有良好的視覺效果，透過 CNN 先 extract 圖像的特徵，再透過 sub-pixel convolution 生成 high resolution image，對於細節或結構性線條有較好表現。實驗過程當中，也發現 ESPCN 訓練的速度比 SRResNet 快，因為模型結構比較輕量，也適合用在有限計算資源的設備上使用。

缺點: 因為模型較為簡單，對於處理複雜圖像，其表現可能不如其他模型佳，例如:SRGAM 或者是 VDSR。另外，期效過也依賴數據，如果訓練不足以涵蓋各種圖像的細節的話，在處理新圖像時可能會表現得比較差。最後，ESPCN 可能會生成 Artifact，會使人看出該部分為人為處理過的痕跡、區域以及瑕疵等等。

其原因主要來自於 sub-pixel convolution 重排過程中是一種 interpolation 的方式，有時會在高頻細節的地方生成不自然的過渡現象，導致 artifacts 的產生。另外一個是因為 ESPCN 採用較淺的 CNN 使 receptive field 不夠大去抓取局部特徵，所以會產生紋理失真，邊緣不連貫等細節。

使用 resnet 50 with pretrained model

From: <https://pytorch.org/vision/master/models.html>

Epoch 40	Train Loss: 0.0243	Val Loss: 0.0277	PSNR: 15.58 dB"
Epoch 41	Train Loss: 0.0239	Val Loss: 0.0386	PSNR: 14.13 dB"
Epoch 42	Train Loss: 0.0240	Val Loss: 0.0305	PSNR: 15.16 dB"
Epoch 43	Train Loss: 0.0232	Val Loss: 0.0217	PSNR: 16.64 dB"
Epoch 44	Train Loss: 0.0225	Val Loss: 0.0260	PSNR: 15.84 dB"
Epoch 45	Train Loss: 0.0220	Val Loss: 0.0249	PSNR: 16.04 dB"
Epoch 46	Train Loss: 0.0216	Val Loss: 0.0225	PSNR: 16.49 dB"
model saved			
Epoch 47	Train Loss: 0.0212	Val Loss: 0.0132	PSNR: 18.79 dB"
Epoch 48	Train Loss: 0.0214	Val Loss: 0.0205	PSNR: 16.88 dB"
Epoch 49	Train Loss: 0.0223	Val Loss: 0.0257	PSNR: 15.90 dB"
Epoch 50	Train Loss: 0.0221	Val Loss: 0.0227	PSNR: 16.43 dB"

最佳:18.79dB

初步研判 resnet 50 效果這麼差的原因在於他主要是特徵擷取用來分類的。且 superresolution 需要細節恢復，然而這些細小邊緣紋理並不是 resnet-50 的主要任務重點。其核心在於提取全面結構的特徵，會忽略掉細節。

使用 basic CNN

```

model saved
Epoch 1 Train Loss: 0.0506 | Val Loss: 0.0188 | PSNR: 17.26 dB"
model saved
Epoch 2 Train Loss: 0.0112 | Val Loss: 0.0081 | PSNR: 20.94 dB"
model saved
Epoch 3 Train Loss: 0.0062 | Val Loss: 0.0047 | PSNR: 23.24 dB"
model saved
Epoch 4 Train Loss: 0.0049 | Val Loss: 0.0037 | PSNR: 24.33 dB"
Epoch 5 Train Loss: 0.0044 | Val Loss: 0.0041 | PSNR: 23.82 dB"
Epoch 6 Train Loss: 0.0041 | Val Loss: 0.0039 | PSNR: 24.09 dB"
Epoch 7 Train Loss: 0.0040 | Val Loss: 0.0050 | PSNR: 23.02 dB"
model saved
Epoch 8 Train Loss: 0.0039 | Val Loss: 0.0033 | PSNR: 24.84 dB"
Epoch 9 Train Loss: 0.0038 | Val Loss: 0.0042 | PSNR: 23.74 dB"
Epoch 10 Train Loss: 0.0037 | Val Loss: 0.0037 | PSNR: 24.34 dB"
Epoch 11 Train Loss: 0.0037 | Val Loss: 0.0040 | PSNR: 23.95 dB"
model saved
Epoch 12 Train Loss: 0.0036 | Val Loss: 0.0030 | PSNR: 25.24 dB"
model saved
Epoch 13 Train Loss: 0.0036 | Val Loss: 0.0024 | PSNR: 26.26 dB"
Epoch 14 Train Loss: 0.0036 | Val Loss: 0.0048 | PSNR: 23.15 dB"
Epoch 15 Train Loss: 0.0035 | Val Loss: 0.0035 | PSNR: 24.53 dB"

```

最佳 26.26dB

使用 SRResNet pretrained model

From: <https://github.com/twtygqyy/pytorch-SRResNet>

```

model saved
Epoch 20 Train Loss: 0.0037 | Val Loss: 0.0025 | PSNR: 26.01 dB"
Epoch 21 Train Loss: 0.0037 | Val Loss: 0.0039 | PSNR: 24.12 dB"
model saved
Epoch 22 Train Loss: 0.0036 | Val Loss: 0.0021 | PSNR: 26.68 dB"
Epoch 23 Train Loss: 0.0036 | Val Loss: 0.0032 | PSNR: 24.96 dB"
Epoch 24 Train Loss: 0.0036 | Val Loss: 0.0040 | PSNR: 23.99 dB"
Epoch 25 Train Loss: 0.0036 | Val Loss: 0.0027 | PSNR: 25.68 dB"
Epoch 26 Train Loss: 0.0036 | Val Loss: 0.0035 | PSNR: 24.61 dB"
Epoch 27 Train Loss: 0.0035 | Val Loss: 0.0031 | PSNR: 25.13 dB"
Epoch 28 Train Loss: 0.0035 | Val Loss: 0.0036 | PSNR: 24.45 dB"
Epoch 29 Train Loss: 0.0035 | Val Loss: 0.0037 | PSNR: 24.28 dB"
Epoch 30 Train Loss: 0.0035 | Val Loss: 0.0041 | PSNR: 23.89 dB"
Epoch 31 Train Loss: 0.0035 | Val Loss: 0.0031 | PSNR: 25.11 dB"
Epoch 32 Train Loss: 0.0034 | Val Loss: 0.0043 | PSNR: 23.67 dB"
Epoch 33 Train Loss: 0.0034 | Val Loss: 0.0024 | PSNR: 26.28 dB"
Epoch 34 Train Loss: 0.0034 | Val Loss: 0.0022 | PSNR: 26.49 dB"
Epoch 35 Train Loss: 0.0034 | Val Loss: 0.0043 | PSNR: 23.62 dB"

```

最佳 26.68dB

討論:

1. What is pixel shuffle?

是一種將 low resolution 的特徵圖轉為 high resolution 的技術。其原理是把輸入進行透過 convolution layer 進行重排，並將多個通道的特徵合併，再重新組合成一個 superresolution 的圖像。正常情況下因為 convolution layer 會使 feature map 的高度以及寬度變小，因此如果我們使用 stride = 1/r < 1 時，可以使 feature map 的高和寬變大，這樣就可以達到增加分辨率的效果。這種技術可以透過 interpolation, perforate or un-pooling 來實現，但是會增加運算成本。

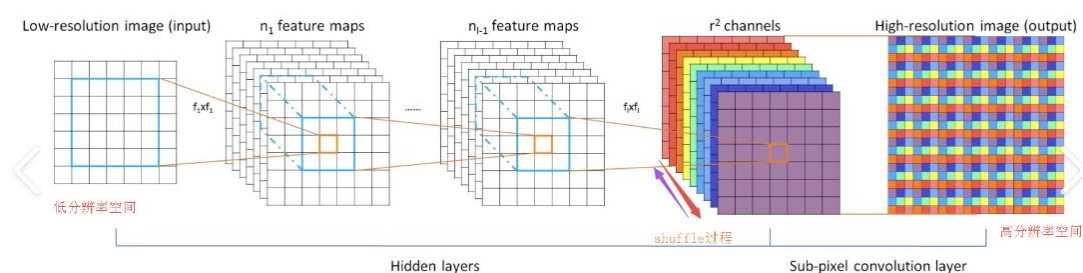


Figure 1. The proposed efficient sub-pixel convolutional neural network (ESPCN), with two convolution layers for feature maps extraction, and a sub-pixel convolution layer that aggregates the feature maps from LR space and builds the SR image in a single step.

圖片來源: <https://arxiv.org/pdf/1609.05158.pdf>

透過上張圖可以知道，原本一個 $h \times w$ 的 low resolution 圖 convolution layer 先得到 r^2 個通道的 feature map(每個大小 $h \times w$)在通過 periodic shuffling 的方法變成 $rh \times rw$ 的 high resolution 的圖片，其中 r 為 upscaling factor 為圖像的放大倍率。

2. What is PSNR and discuss why it is not the only metric used for evaluating super-resolution. And give some other metrics that provide different perspectives on image quality

PSNR 是用來評估 output 質量的一種指標，特別會被用來使用在圖像壓所或者是 superresolution 領域。其公式主要由圖像最大像素質以及原始圖像與 output 圖像的像素差異的均方誤差值。計算出來的分貝數越高，代表生成圖片越接近原本圖片。通常在 30dB 以上表示叫好的圖像質量，40 以上則難用肉眼區分。

$$PSNR(dB) = 10 \log_{10} \frac{c * MAX^2}{MSE}$$

MAX 為圖像最大像素質 (ex: 8 位圖像, $MAX = 255$)

MSE 為原始圖像與重建圖像的像速差異平方平均值

c 為色彩通道數量

PSNR 不是唯一的指標，其原因在於人眼對於圖像的品質感知不一定相關。代表較高的 psnr 值不代表人覺得畫素很高，因為 psnr 主要與像素差異有關。另外，

psnr 將圖像視為像素的集合(只考慮單個像素之間的數值差異)，不考慮其結構和模式。然而，圖像裡的邊界、形狀以及紋理等等，都會影響人感知圖像的因素。有可能在一幅明顯邊界的圖像當中，即使數值有所偏差，但其整體仍然看起來很清晰。而其他也用於評估圖像質量的指標包括 SSIM, MSE, NRMSE, FSIM 等。

指標	公式	優點	缺點
SSIM(structural similarity)	$SSIM(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$ <p> μ_x 和 μ_y 是兩幅圖像的平均值。 σ_x^2 和 σ_y^2 是兩幅圖像的方差。 σ_{xy} 是兩幅圖像的協方差。 C_1 和 C_2 是為了穩定計算而添加的常數。 </p>	考慮圖像的結構訊息 範圍介於-1 至 1 間	對於細小變化不靈敏
MSE	$MSE = \frac{1}{N} \sum_i^N (x_i - y_i)^2$ <p>N:像素總數, xi:輸入, yi: 輸出</p>	計算簡單	不考慮結構，不直接表示人類感知的結果 小誤差評估不敏感
NRMSE	$NRMSE = \sqrt{\frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} y[m, n] - x[m, n] ^2}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x[m, n] ^2}}$ <p>兩個單色影像 $x[m,n]$和 $y[m,n]$, N:像素總數</p>	易於比較 反映圖像值範圍對誤差影響	受極端值影響 不考慮結構以及特徵
FSIM	$FSIM(I_{original}, I_{reconstructed}) = \frac{\sum_i \phi(I_{original}(i), I_{reconstructed}(i))}{\sum_i \phi(I_{original}(i))}$ <p>其中 ϕ 是特徵相似性的計算函數。 l:像素質</p>	考慮整體結構 對於邊緣與細節捕捉能力強	計算複雜度較高

Reference:

[\[電腦視覺\]圖像衡量指標 PSNR、SSIM 介紹 - Adam He - Medium](#)

[图像质量评估指标：MSE，PSNR，SSIM 图像质量评估指标:mse,psnr,ssim-CSDN 博客](#)

[正規化均方根誤差 - 維基百科，自由的百科全書 \(wikipedia.org\)](#)

[图像质量评价（三）：FSIM - 知乎 \(zhihu.com\)](#)

[图像质量评估指标（3） 特征相似度 FSIM-CSDN 博客](#)

[1609.05158 \(arxiv.org\)](#)