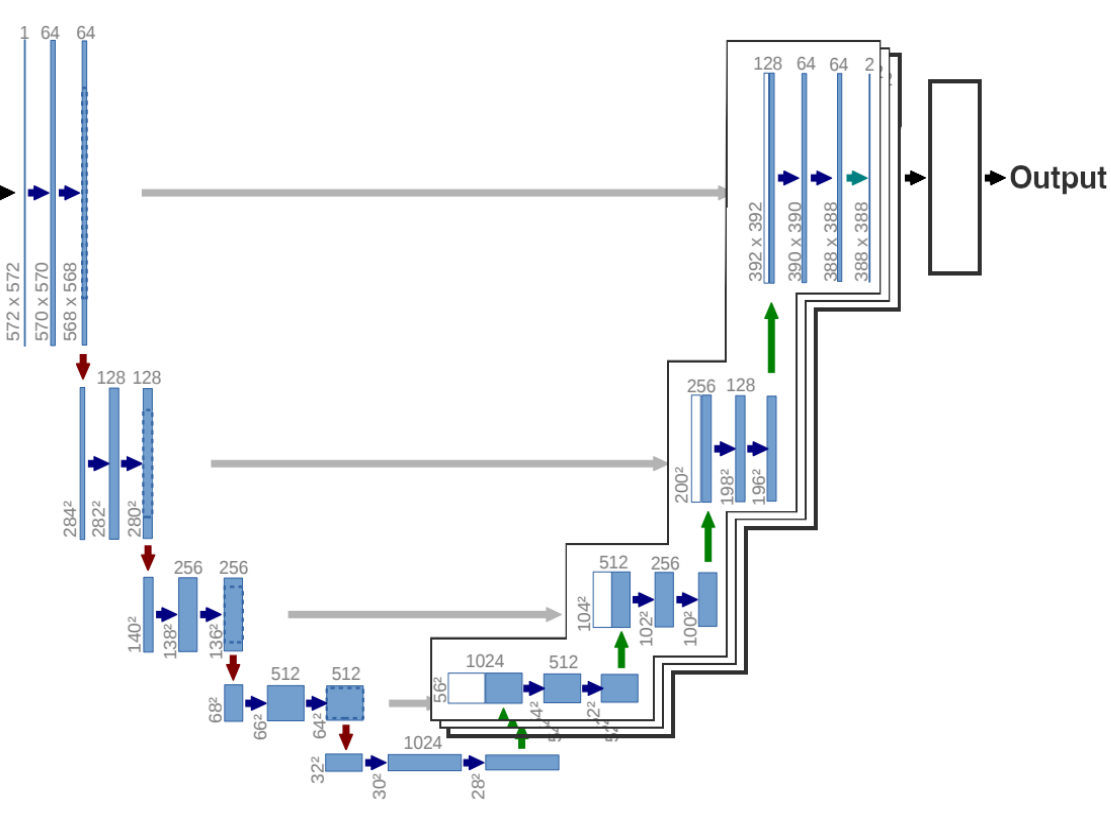
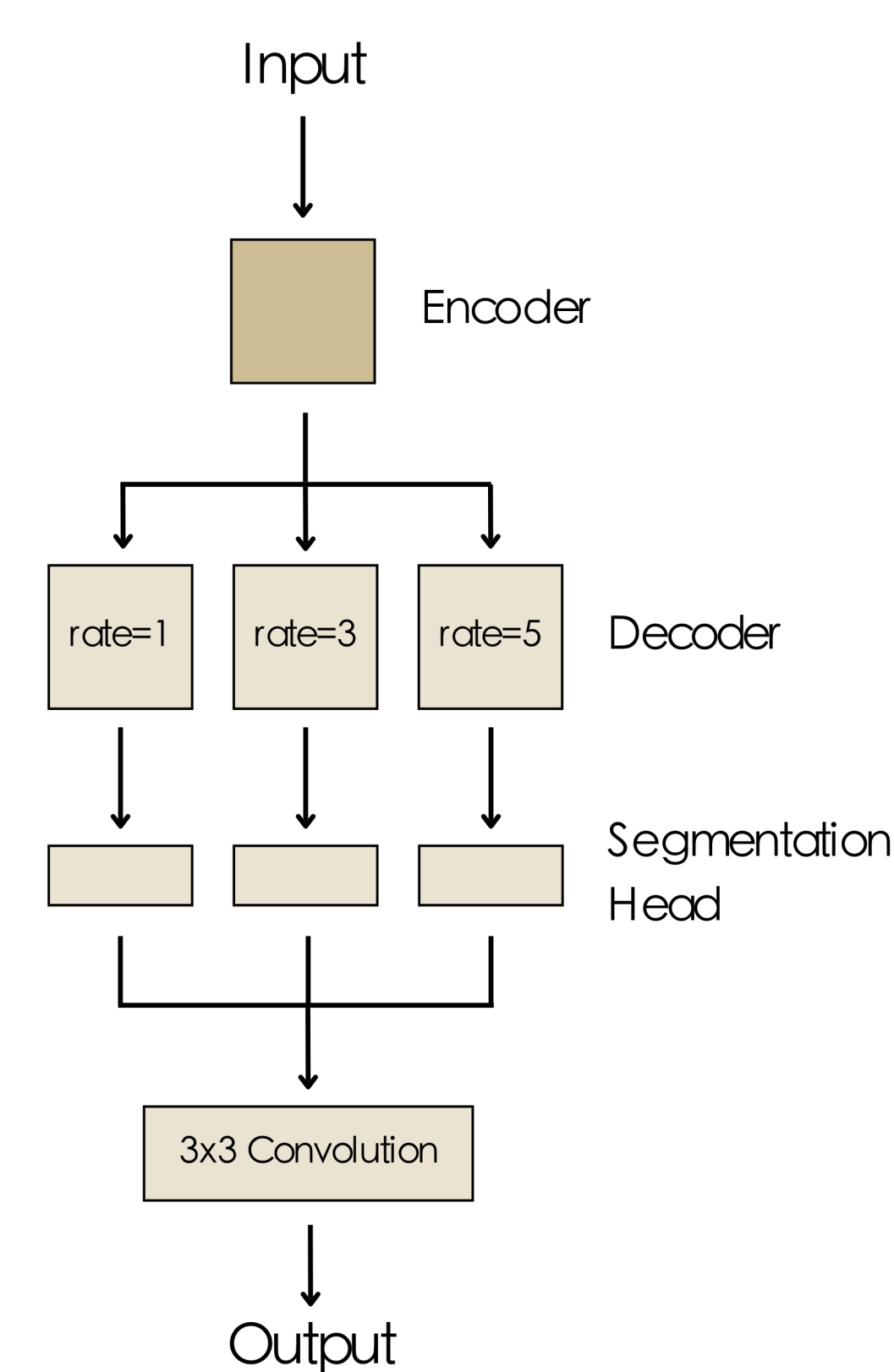


Architecture



修改自 arXiv:1505.04597

新式影像分割模型於建物外牆破壞偵測之研發

利用空洞卷積 (Dilated convolution) 增進 U-Net 分割模型效能之研究

指導老師 吳日騰 / 專案經理 王海威
實習生 周秉宇

Introduction

磁磚材質的建築外牆常有剝落之風險，在都市地區易造成危險使過路人受到傷害，因此成為研究之焦點。本專案試圖以深度學習技術為基礎，開發偵測建物外牆磁磚剝落區塊之影像分割模型 (Image segmentation model)，並採用目前影像分割領域之主流，編碼器/解碼器 (Encoder-Decoder) 設計的 U-Net 架構開發。

由於磁磚剝落區塊大小差異很大，我們嘗試藉由空洞卷積 (Dilated convolution) 來增加模型之感受視野 (Receptive field)。本專案將藉由實作多分支的空洞卷積層，提升模型辨識磁磚剝落之表現。加入空洞卷積層將引入一個新的參數：擴張率 (Dilation rate)，會影響卷積時核 (kernel) 的大小。由於像素採樣的差異，最適當的擴張率和影像大小會有所關聯。本專案旨在研究如何根據縮放後影像大小 (長/寬) 搭配適當之空洞卷積層擴張率組合，以最佳化訓練後模型用於影像分割之表現。

Methods

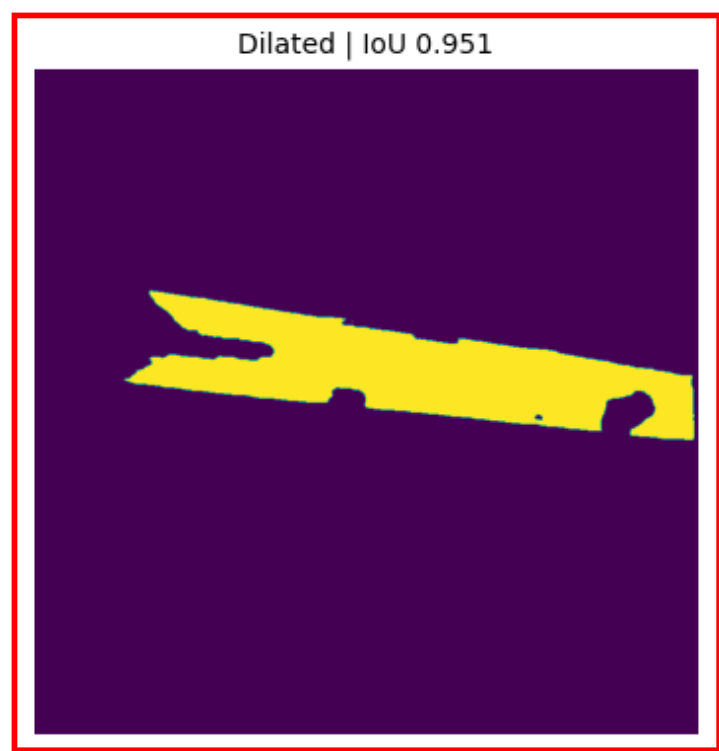
為了減少隨機誤差，我們針對不同組合各進行五次模型訓練實驗，過程將原始資料集按比例隨機拆分成訓練、驗證、測試集，五次實驗中採不同測試集，最終以測試集之平均 IoU (Intersection-over-Union) 值作為判斷模型效能之依據。

Experiments

修改架構前與修改架構後對比 (目前最佳表現)：

mIoU Score	Run 1	Run 2	Run 3	Run 4	Run 5	Mean
Original U-Net	0.520	0.542	0.515	0.379	0.544	0.500
Dilated (rate = 1, 3, 5)	0.515	0.534	0.545	0.441	0.517	0.510

Image size: 800, Encoder: EfficientNet-B6，最大 Epoch 50。訓練時間 1hr10m、GPU RAM 28.4G。

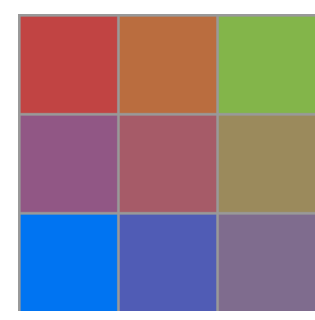


新架構下，Dilation rate 與 Image size 間組合對比：

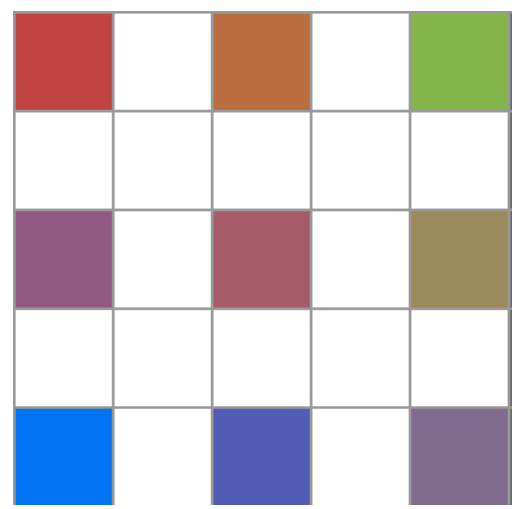
Mean mIoU Score	rate = 1, 3	rate = 1, 2, 3	rate = 1, 3, 5	rate = 3, 5, 7	rate = 1, 3, 5, 7
Image size = 384 訓練時間、GPU RAM 3.2G	0.243	0.244	0.223	0.173	0.228
Image size = 800 訓練時間 36min、GPU RAM 11.5G	0.339	0.339	0.353	0.307	0.383
Image size = 1600 訓練時間 1hr30m、GPU RAM 43.8G	0.367	0.372	0.369	0.371	GPU RAM不足

Encoder: MobileNet-V2，五次不同測試資料集訓練後平均，最大 Epoch 50。
(訓練時間及 GPU RAM 使用量以 rate = 1, 3, 5 為代表)

一般 kernel



Dilation rate = 2



空洞卷積 (Dilated convolution) Kernel 示意圖

Results & Discussion

比較結果後，可以看見新架構有明顯較佳之表現。而在下方實驗中，最適當的擴張率受影像大小影響主要顯現在影像大小 384 的該組結果，「擴張率 = 1, 3」及「= 1, 2, 3」組合較其他組合有更好的結果。在 1600 之影像大小較好的結果則是集中在「擴張率 = 1, 2, 3」及「= 3, 5, 7」上，因此最適擴張率會受影像大小微幅影響。

亦可觀察到，分支數較多的「擴張率 = 1, 3, 5, 7」在 800 之影像大小上表現較佳。因此推估若能將其套用在 1600 之影像大小上，可以有更好的表現。然而影像大小 1600 的訓練過程需佔據大量 GPU 記憶體，我們由於硬體限制無法完整測試。未來使用者需要在採用之影像大小選擇上權衡。

Conclusion

本專案利用多分支空洞卷積層的架構成功提升原模型之表現，此外若能依據圖像大小搭配適當的擴張率組合，將能使成果更進一步。