

高等數位訊號處理 期末報告

台科大 電機所碩一

游綾玉 M11407308

Dynamic U-Net: Adaptively Calibrate Features for Abdominal Multi-organ Segmentation

Jin Yang¹, Daniel S. Marcus¹, and Aristeidis Sotiras^{1,2}

¹ Department of Radiology, Washington University School of Medicine, St. Louis, MO, USA

² Institute for Informatics, Data Science and Biostatistics, Washington University School of Medicine, St. Louis, MO, USA
yang.jin@wustl.edu

arXiv:2403.07303v1 [eess.IV] 12 Mar 2024

U-Net & nnU-Net

U-Net

- 卷積神經網路架構
- 為生醫影像分割設計
- 「U 型」結構，包含編碼器、負責還原解析度的解碼器，以及將兩者特徵結合的跳躍連接

nnU-Net

- "No New-Net" (沒有新的網路)

Comparison

	U-Net	nnU-Net
定位	單純的「模型架構」	端到端 (End-to-End) 的「自動化訓練框架」
超參數調校	人工手動	全自動
架構選擇	固定	動態組合 (會自動訓練 2D, 3D 和 Cascade 3D U-Net 並從中挑選或融合最佳結果)
適應能力	換了一個資料集，通常需要重新修改程式碼與參數	幾乎不需要改程式碼，可以直接適應各種不同的醫學影像 (MRI、CT、顯微鏡影像等)

Dynamic U-Net

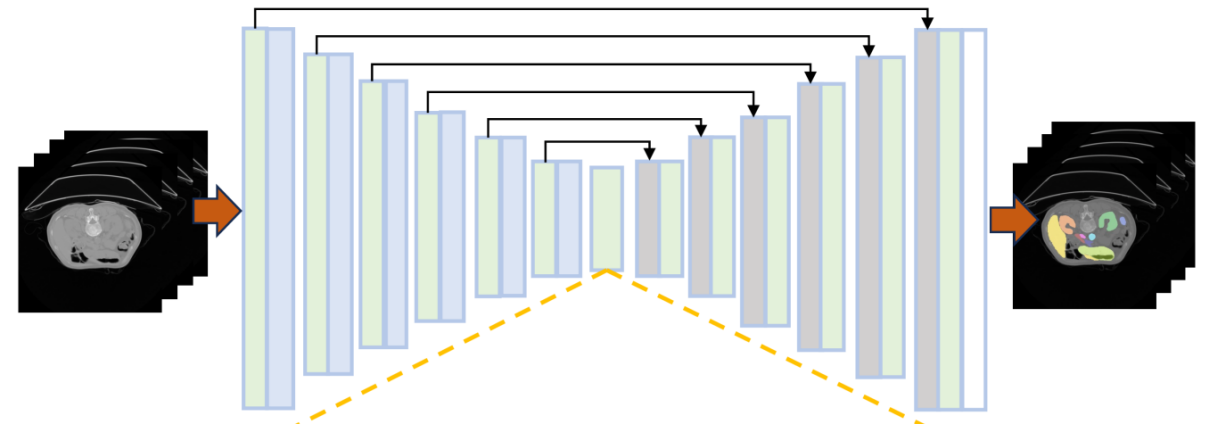
Encoder

萃取特徵並縮小影像尺寸

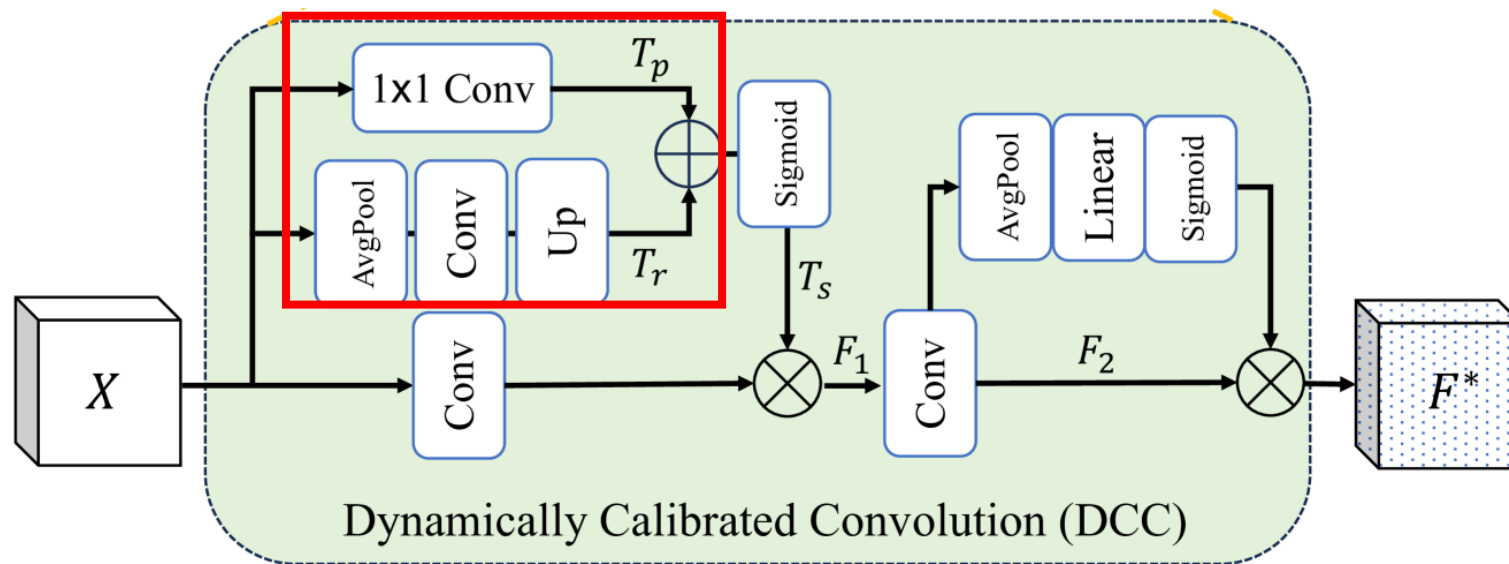
- DCC (特徵萃取與校準):
利用全局上下文資訊
校準空間與通道的權重
→ 取代傳統的卷積層
- DCD (降採樣/縮小尺寸):
自適應地決定該保留哪些
→ 取代傳統的 Max-Pooling

Decoder

放大特徵圖並還原出器官的位置



Spatial-wise Calibration



把輸入特徵 (X) 分成兩條平行的路徑來處理：

1. Pixel-wise :

逐一檢查每個獨立像素點的資訊，產生像素級的權重圖

2. Region-wise :

它先用平均池化 (4x4) 把圖片縮小，接著用 3x3 卷積捕捉鄰近區域的上下文資訊，最後再放大回原尺寸，產生區域級的權重圖 (T_r)。

Region-wise

0.1	0.1	0.1	0.1		0.2	0.2	0.2	0.2		0.1	0.1	0.1	0.1
0.1	0.1	0.1	0.1		0.2	0.2	0.2	0.2		0.1	0.1	0.1	0.1
0.1	0.1	0.1	0.1		0.2	0.2	0.2	0.2		0.1	0.1	0.1	0.1
0.1	0.1	0.1	0.1		0.2	0.2	0.2	0.2		0.1	0.1	0.1	0.1

0.0	0.0	0.0	0.0		0.3	0.3	0.3	0.3		0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0		0.3	0.3	0.3	0.3		0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0		0.3	0.3	0.3	0.3		0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0		0.3	0.3	0.3	0.3		0.0	0.0	0.0	0.0

0.1	0.1	0.1	0.1		0.2	0.2	0.2	0.2		0.1	0.1	0.1	0.1
0.1	0.1	0.1	0.1		0.2	0.2	0.2	0.2		0.1	0.1	0.1	0.1
0.1	0.1	0.1	0.1		0.2	0.2	0.2	0.2		0.1	0.1	0.1	0.1
0.1	0.1	0.1	0.1		0.2	0.2	0.2	0.2		0.1	0.1	0.1	0.1

$$P = \begin{bmatrix} 0.1 & 0.2 & 0.1 \\ 0.0 & 0.3 & 0.0 \\ 0.1 & 0.2 & 0.1 \end{bmatrix}$$

$$K = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$(0.2 \times 1) + (0.0 \times 1) + (0.0 \times 1) + (0.2 \times 1) + (0.3 \times 4) = 1.6$$

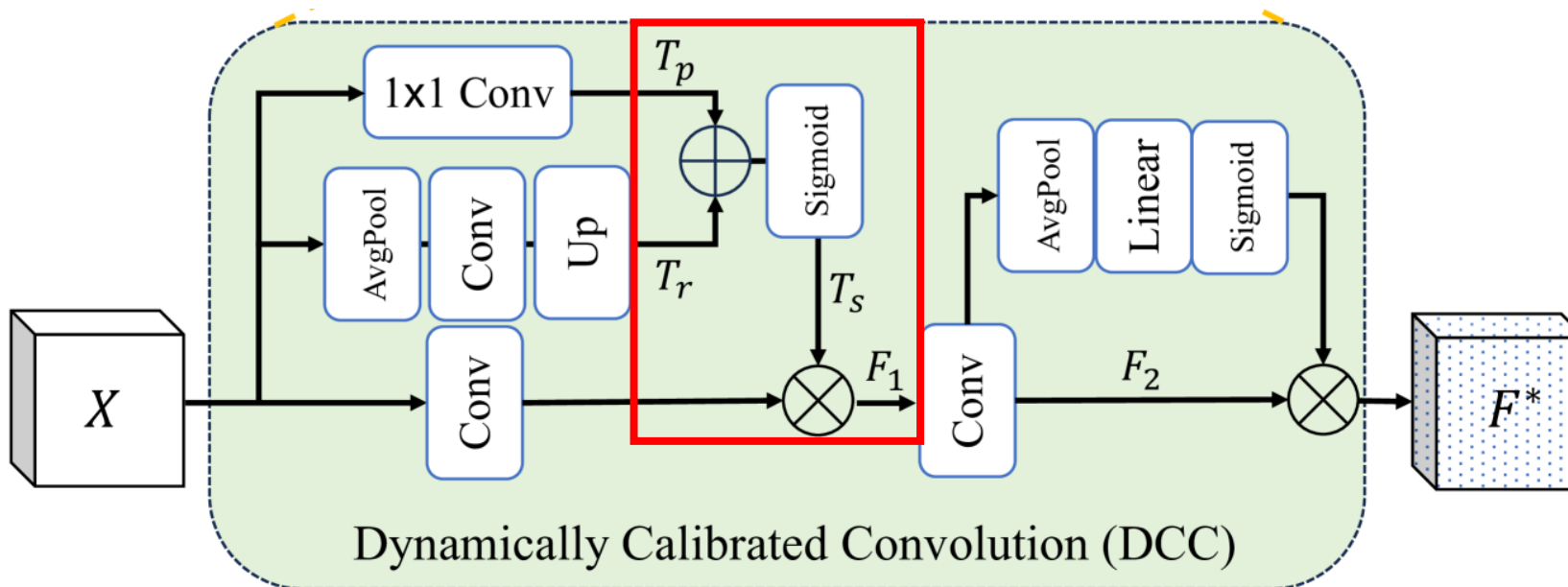
區域上下文分數圖：

$$O = \begin{bmatrix} 0.6 & 1.3 & 0.6 \\ 0.5 & 1.6 & 0.5 \\ 0.6 & 1.3 & 0.6 \end{bmatrix}$$

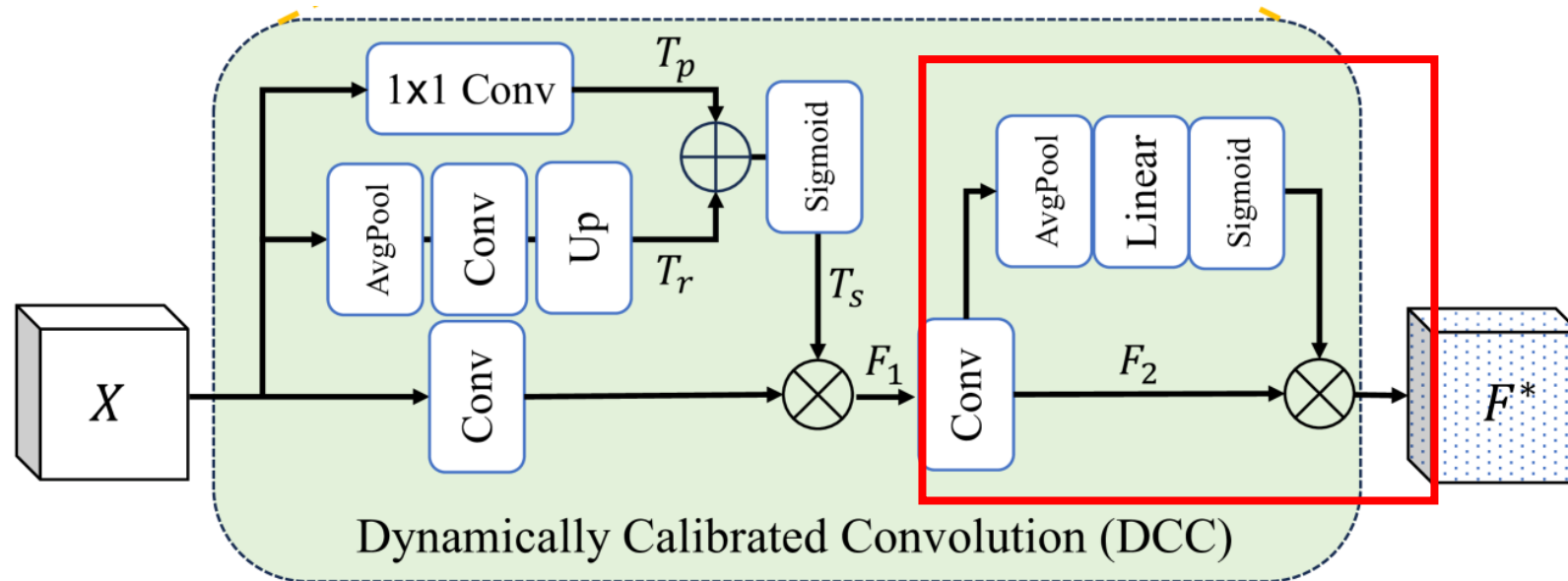
Spatial-wise Calibration

Calibration : 把這個 0~1 之間的權重矩陣乘上原本的特徵圖

$$p_i = \sigma(a_i) = \frac{1}{1+e^{-a_i}}$$



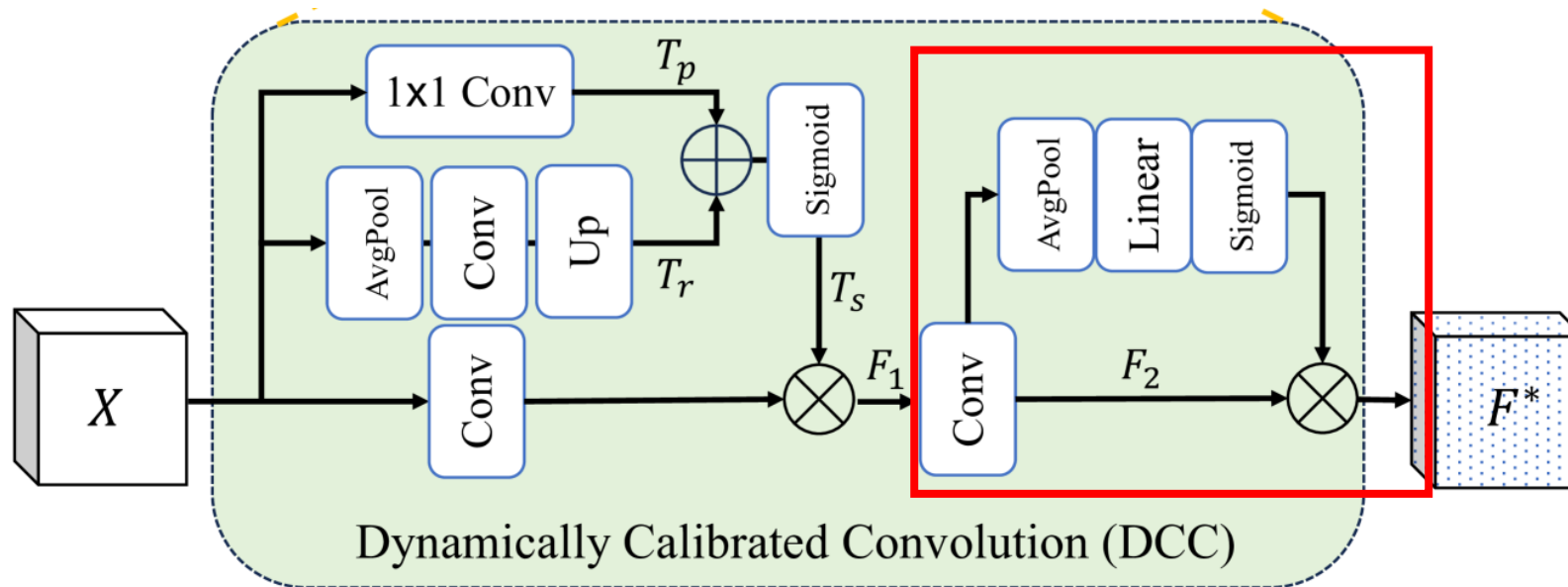
Channel-wise Calibration



F_2 : 萃取出特定器官的空間特徵

AvgPool : 把整張特徵圖濃縮成一個一維陣列

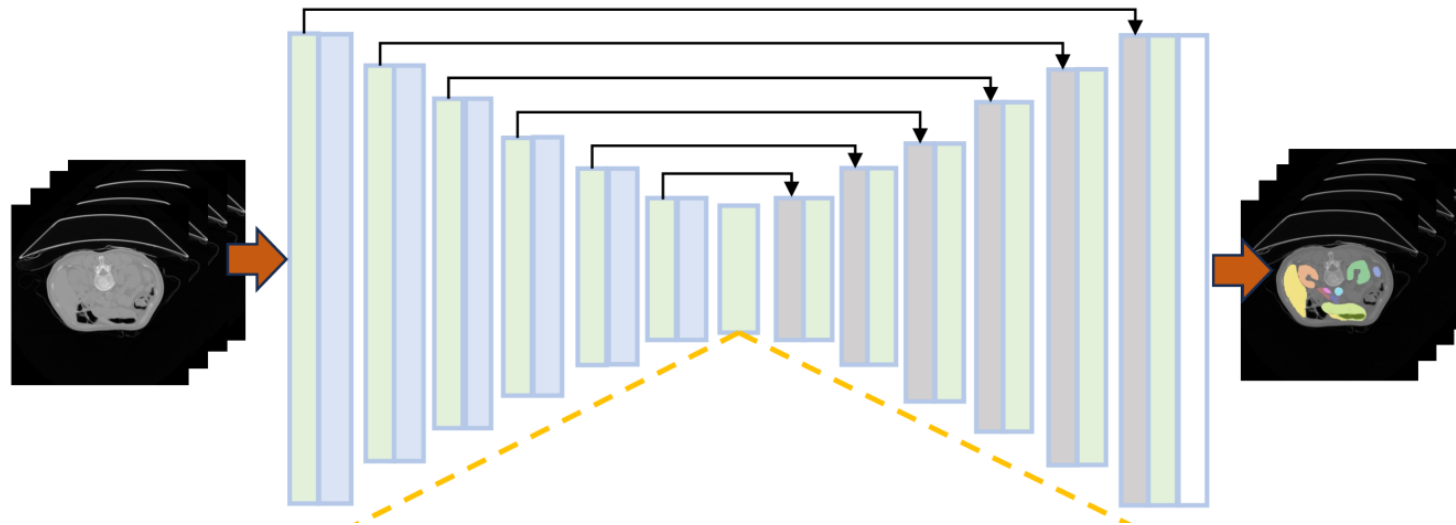
Channel-wise Calibration



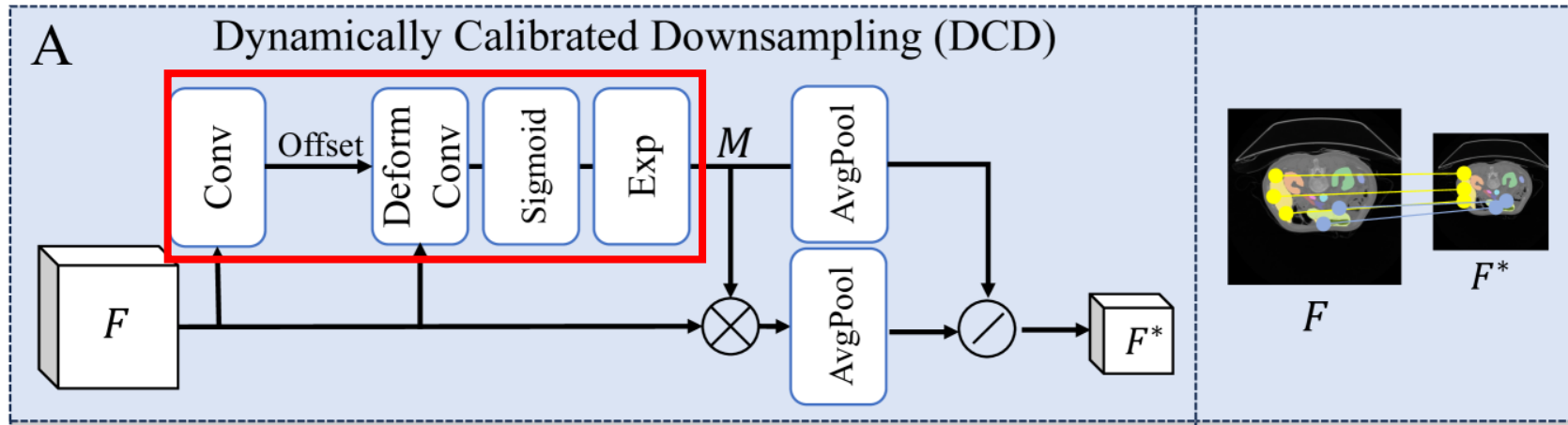
Linear : 學習並計算出『通道間的Inter-dependencies』

$$Y = WX + b$$

Architecture



Dynamically Calibrated Downsampling

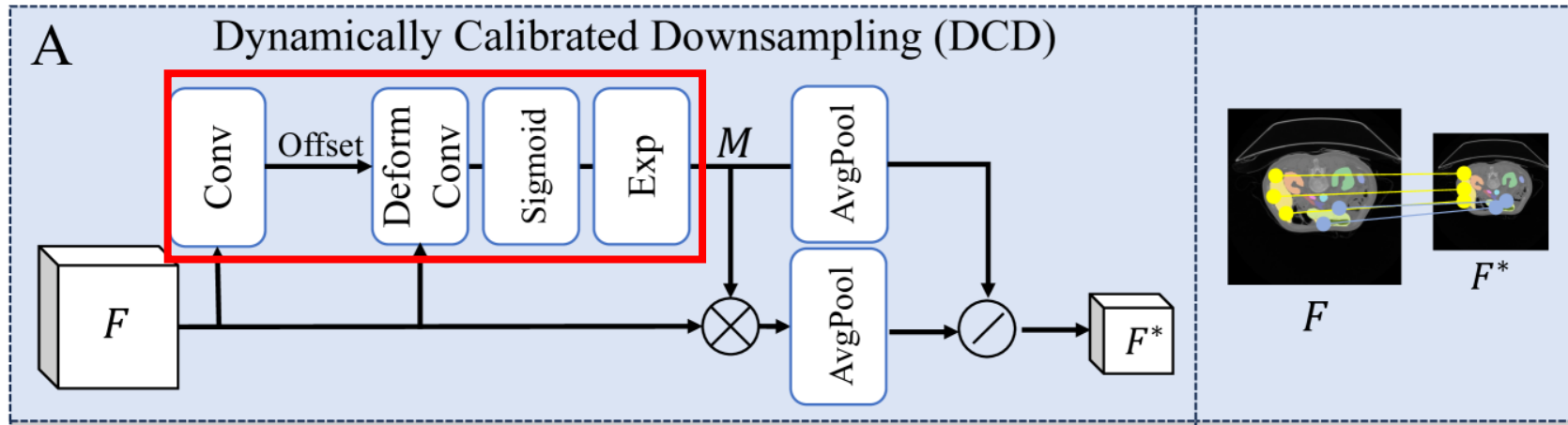


輸入的特徵訊號 F :

$$F = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0_{(1,1)} & 0 \\ 0 & 0 & 100_{(2,2)} \end{bmatrix}$$

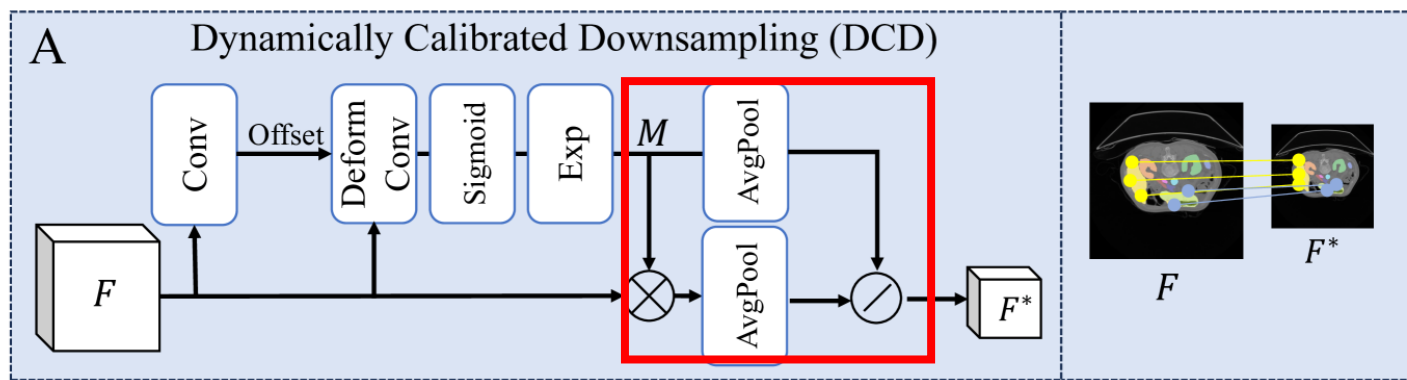
- 3×3 卷積: 找偏移量
→ 位移向量
 $\Delta x = +1, \Delta y = +1$

Dynamically Calibrated Downsampling



- 可變形卷積 \rightarrow 採樣座標 $(1+1, 1+1) = (2,2)$
- Exp: 數值呈現非線性的爆發，且結果絕對大於 0
- M: 標示了在縮小影像時，哪些空間位置的特徵必須被保留

Dynamically Calibrated Downsampling



把 2×2 的區域壓縮成 1×1 的單一數值

輸入特徵圖

$$F = \begin{bmatrix} 10 & 2 \\ 1 & 0 \end{bmatrix}$$

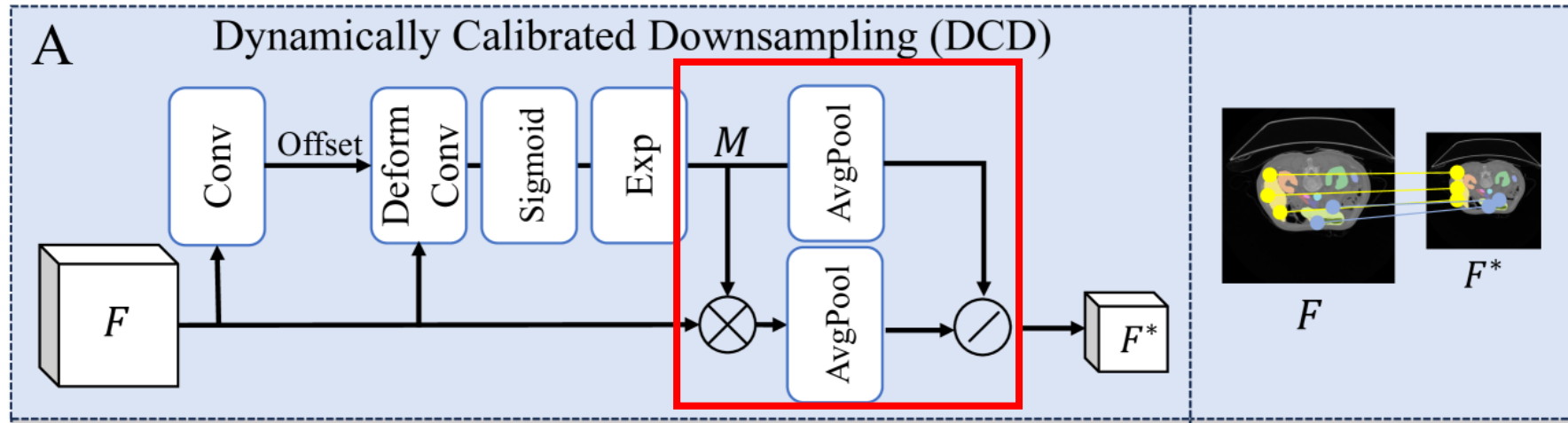
10：代表強烈的器官邊緣特徵

2, 1, 0：代表不重要的背景

- Max-Pooling：太極端，丟失周圍的上下文
- Average-Pooling：傳統降採樣會丟失鑑別性細節

$$(10 + 2 + 1 + 0) \div 4 = 3.25$$

Dynamically Calibrated Downsampling



$$F^* = \mathcal{F}_{AvgPool}(M \otimes F) / \mathcal{F}_{AvgPool}(M)$$

$M \otimes F$: Element-wise multiplication

$$\frac{\frac{1}{N} \sum (M_i \times F_i)}{\frac{1}{N} \sum M_i} = \frac{\sum (M_i \times F_i)}{\sum M_i}$$

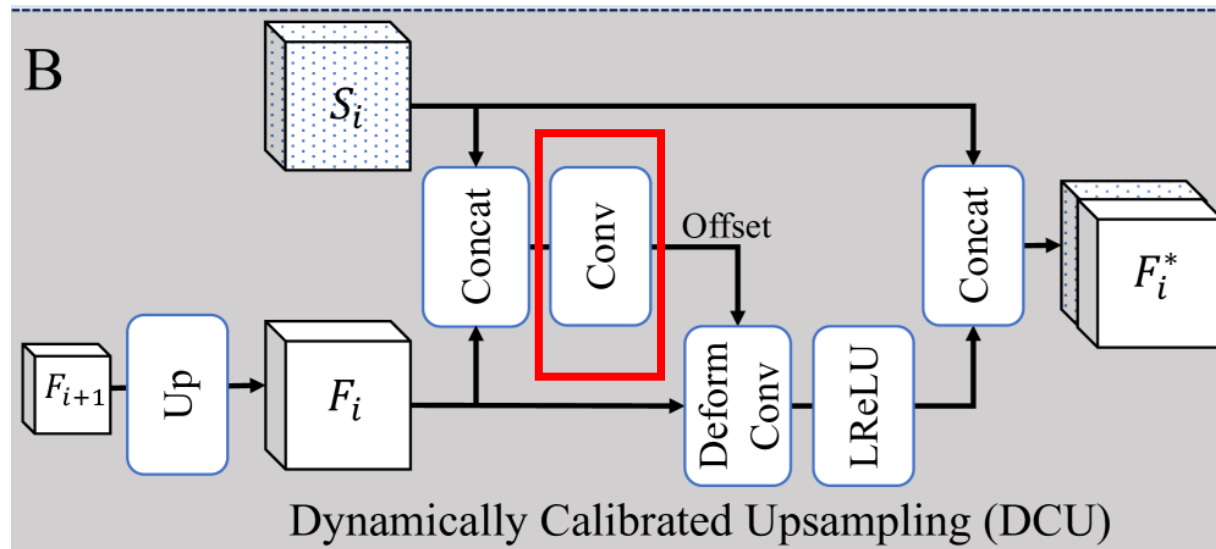
Dynamically Calibrated Upsampling

S_i : 器官位置是準確的

F_i : 從網路最底層一路放大上來的特徵。經歷了多次的降採樣與上採樣

Conv: 不是用來萃取器官特徵的，而是把它當作一個「誤差偵測器」

Offset: 記錄了 F_i 上面的每一個像素，距離它真正在 S_i 上的正確位置



Dynamically Calibrated Upsampling

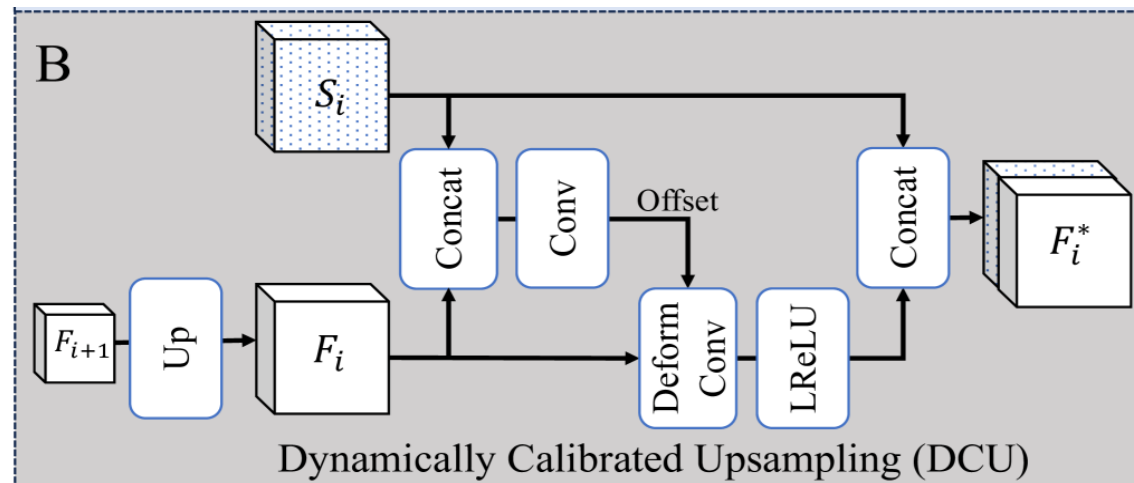
Leaky ReLU (像濾波整流器)

大於 0 的訊號：

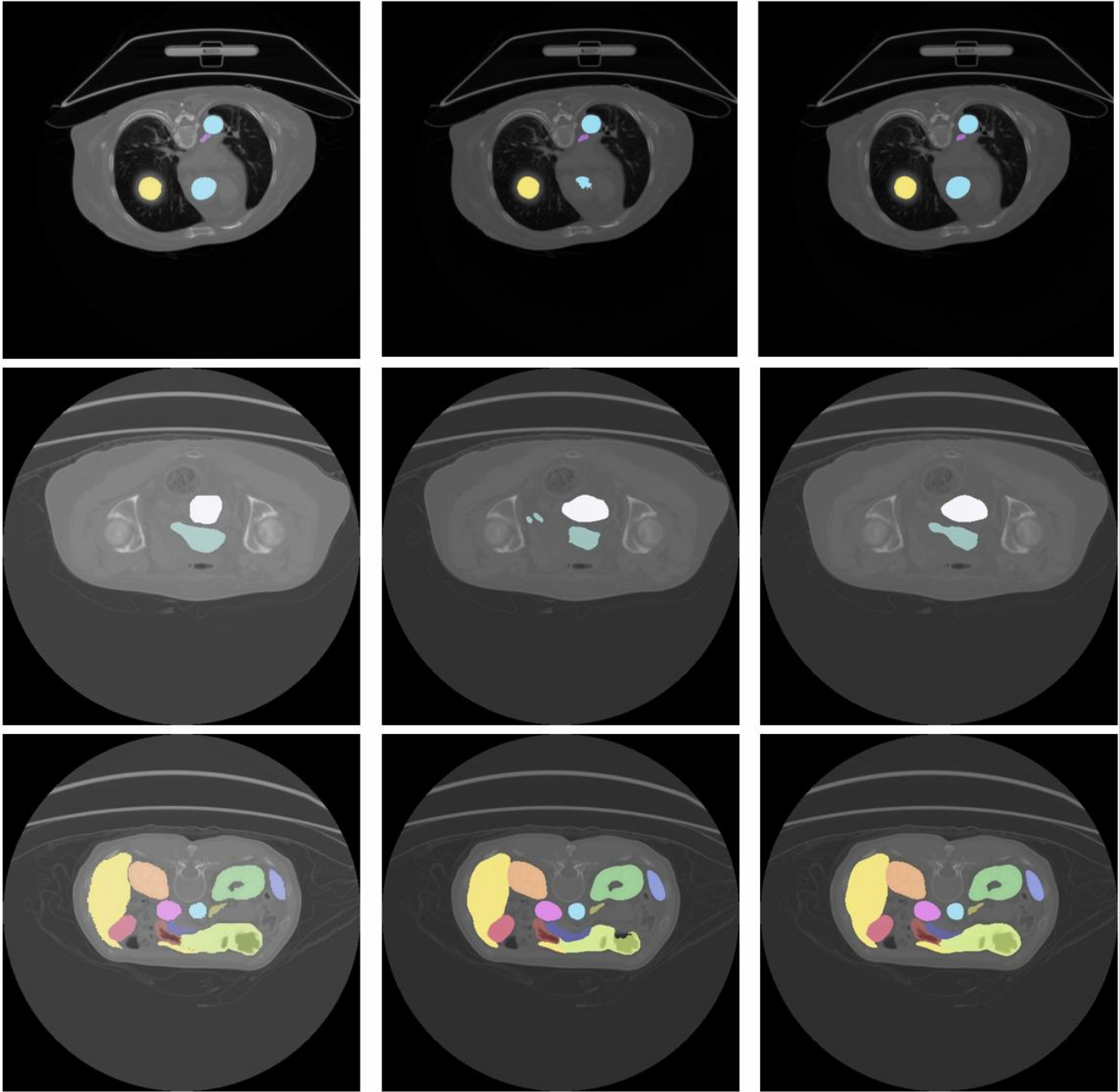
原封不動通過 ($y=x$)，代表保留強烈、自信的器官特徵。

小於 0 的訊號：

乘上一個很小的衰減係數。大幅抑制了負值雜訊，但沒有把它完全抹殺掉，保留了微弱的背景梯度讓網路還能繼續學習。



Results



Ground Truth

U-Net

Dynamic U-Net

Results

Table 1. Comparison of multi-organ segmentation between Dynamic U-Net and baselines on the FLARE 2021. (**Bold** represents the best segmentation results).

Tasks	U-Net [†]	SConv U-Net [†]	DCC U-Net	DCD U-Net	DCU U-Net	Dynamic U-Net
Liver	97.33	97.36	98.46	98.39	98.40	98.40
Kidney	95.50	95.58	96.66	96.70	96.62	97.22
Spleen	95.81	95.85	98.12	98.02	98.11	98.14
Pancreas	79.91	80.66	81.83	81.38	81.23	83.54
Average	92.14	92.36	93.77	93.62	93.59	94.33 *

[†]: we implemented nnUNet with different architectures here.

*: $p < 0.01$ with Wilcoxon signed-rank test between Dynamic U-Net, and U-Net, SConv U-Net.

Results

Table 2. Comparison of multi-organ segmentation between Dynamic U-Net and baselines on the AMOS 2022. (**Bold** represents the best segmentation results).

Tasks	U-Net [†]	SConv U-Net [†]	DCC U-Net	DCD U-Net	DCU U-Net	Dynamic U-Net
Spleen	96.31	96.60	96.83	97.05	97.12	97.64
R. kidney	96.45	96.15	96.61	97.18	97.03	97.70
L. kidney	96.36	96.23	96.68	96.81	96.99	96.33
Gall bladder	78.58	79.56	79.77	80.50	80.22	83.83
Esophagus	84.13	83.61	84.84	85.37	85.25	86.17
Liver	96.74	96.76	96.97	97.53	97.36	98.52
Stomach	87.39	88.65	89.44	89.32	89.62	90.29
Aorta	94.54	94.46	94.71	95.24	95.27	96.65
Postcava	89.02	88.78	89.90	90.25	90.06	90.36
Pancreas	83.93	84.50	85.52	85.67	86.38	87.06
R. adrenal gland	77.71	78.27	78.74	78.65	79.50	80.47
L. adrenal gland	79.65	79.58	81.29	81.32	81.16	81.44
Duodenum	77.49	77.87	79.04	78.79	80.22	81.11
Bladder	86.40	85.40	88.12	89.04	88.90	89.06
Prostate/uterus	80.56	80.97	83.30	84.52	84.49	83.81
Average	87.02	87.16	88.12	88.48	88.64	89.36*

[†]: we implemented nnUNet with different architectures here.

*: $p < 0.01$ with Wilcoxon signed-rank test between Dynamic U-Net, and U-Net, SConv U-Net.

T h a n k s f o r
Y o u r L i s t e n i n g