

# National Tsing Hua University

## Fall 2023 11210IPT 553000

### Deep Learning in Biomedical Optical Imaging

#### Homework 3

Student ID: 108011555

#### 1. Report

##### 1.1 Task A: Reduce Overfitting

Implemented technique : using dropout on fully connected layer  
Original CNN model test accuracy is 74.25%

Table 1. Accuracy of different Dropout rate

	Best validation Accuracy	Test accuracy
Dropout out rate		
0.9	93.25%	83.75%
0.8	95.25%	80.50%
0.5	96.75%	76.00%
0.3	96.75%	75.50%
0.1	96.25%	76.25%

#### Discussion :

隨著 dropout rate 減少，模型於 Test set 上的準確率接近原始模型的準確率。當將 90% 的 Neuron 停止運作時，Test accuracy 達到 83.75 %。Dropout 技術使 Neuron 有機率停止運作，而越大的 dropout rate，模型訓練時，減少 Neural Network 中 Neuron 之間共同效應程度越大，從而緩解 Overfitting 的問題。模型不會過度的依賴某些 Neuron Weight。模型複雜度也降低。

#### Implementation Visualization :

```
self.fc1 = nn.Linear(flattened_dim, 32)
self.dropout1 = nn.Dropout(0.9)
self.fc2 = nn.Linear(32, 1)
```

Fig. 1. the modifications made to the original code.

## 1.2 Task B: Performance Comparison between CNN and ANN

Table 2. Compare with ANN and CNN

	ANN	CNN
Feature extraction capabilities	Less powerful	More powerful
Training speed	Less powerful	More powerful
Model performance	Less powerful	Higher

### Discussion :

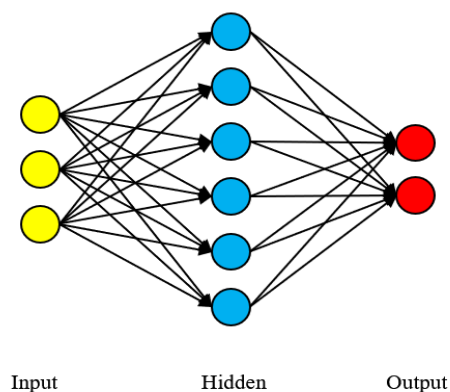
ANN 在解決圖像分類問題時，會先將 2 維圖像轉換成 1 維向量後透過神經元訓練，當圖像尺寸增加，可訓練參數會隨之增加。且 ANN 模型訓練時會失去圖像中像素排列的空間特徵。當網路層數變深，也會導致梯度消失、爆炸的問題。

CNN 在解決圖像分類問題時是透過核函數(Kernel)使用捲積運算從輸入提取相關特徵，捕獲圖像中的空間特徵、圖像中像素的排列及其之間的關係。

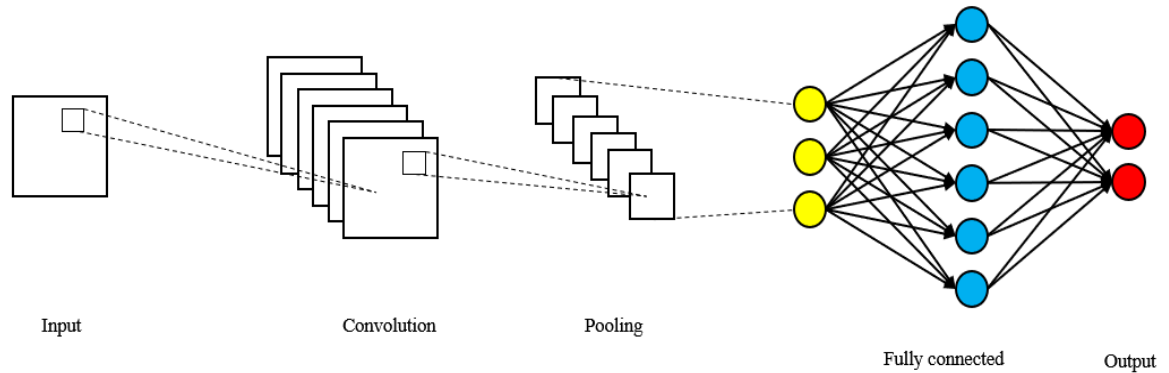
### Architecture Description :

#### Architecture of ANN

神經網路由輸入層、隱藏層、輸出層組成。輸入層將訊息傳送到隱藏層的神經元，隱藏層對輸入資料進行權重計算並將輸出傳送到輸出層。神經元之間存在連結權重，這些權重決定神經元的學習能力。



## Architecture of ANN



CNN 主要的架構有輸入層、卷積層、池化層、全連接層、輸出層。

**卷積層(Convolutional layer):**此層為輸入影像中提取各種特徵的第一層。在這層中，有特定大小的核函數濾波器( $M \times M$ )在輸入影像及濾波器之間進行卷積數學運算。透過在輸入影像上滑動濾波器，可以得到濾波器和輸入影像的各部分相對於 $M \times M$ 大小的濾波器之點積運算，其輸出稱為特徵圖(feature maps)。特徵圖提供局部圖像的特徵空間信息。隨後，特徵圖將被送往其他層以學習輸入影像的其他幾個特徵。

**池化層(Pooling layer):**通常會在卷積層後加入池化層，此層主要目的是減少卷積特徵圖的大小來降低計算成本。池化層有多種類型，如 Max Pooling、Average Pooling...等等。

**全連接層(Fully Connected layer):**通常接在卷積層或池化層之後，對多類別影像分類問題而言，其將最後一層卷基層得到的 feature maps 拉直成向量，對此向量做乘法，最終降低維度後，輸入到 softmax 層中得到對應的每個類別的分數。全連接層存在著參數量過大，可能降低訓練速度，且容易過擬合的問題。

### 1.3 Task C : Global Average Pooling in CNNs

Explanation :

Global Average Pooling 捨棄傳統的全連接層，於最後一個卷積層中輸出數量等於分類數的 feature maps，再取每個 feature maps 的平均值輸出一個向量，且此向量維度為分類數，丟進 softmax 中進行分類。GAP 方法增強了 feature map 與類別之間的關聯，且 GAP 中沒有參數需要訓練優化，可避免 overfitting。

Increase performance :

首先嘗試增加 epoch 數，準確率並沒有顯著提升。

第二次嘗試調整 hyperparameters : kernel size，分別嘗試 16、64、128 個 kernel，準確率分別為 58.75%、73.75%、71.5%，當 kernel size 太少，準確率 75.75%、74.25%、75.25%，準確率沒有顯著提升。

最後調整模型結構，分別增加層數至 4 層、5 層，準確率分別為 76.00%、83.25%，增加卷積層數對提升準確率有幫助。