

CV Final Project

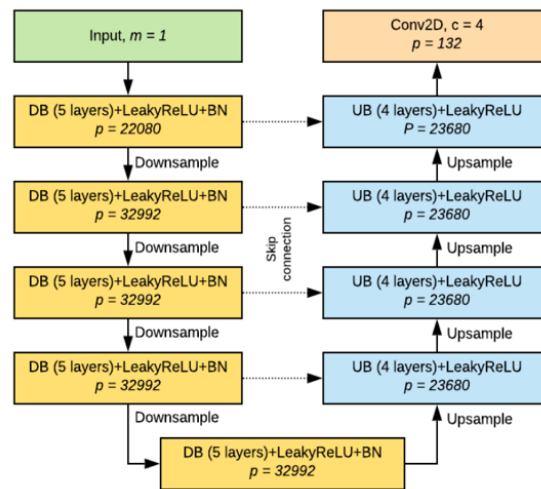
Pupil Tracking

李奕勳、陳昱仁、宋家齊
Department of Electrical Engineering
Media IC & System Lab

r10943019@ntu.edu.tw、r10943117@ntu.edu.tw、b07901187@ntu.edu.tw

1 Methodology or Model Architecture

這次的期末專題，pupil tracking，我們的解法直接將問題拆解成切割以及分類兩個獨立的問題，經過多個模型實驗，切割使用的是RITnet，分類使用的是Efficient net B1。RITnet融合Unet及DenseNet，如下圖所示，每個sub block皆會如同DenseNet交互連接，整體像Unet架構。



2 Implementation Details

2.1 Data preprocessing

這次的期末專題，我們總共做了兩個資料預處理，分別是伽瑪校正(gamma correction)和標準化(normalization)，下文將針對伽瑪校正來介紹。

伽瑪校正是針對圖像中的明度(luminance)和三刺激色值(tristimulus value)所進行的非線性運算，其轉換公式為

$$V_{out} = AV_{in}^{\gamma} \quad (1)$$

在這次的期末專題中，經過不同的實驗和嘗試，我們將A設為1， γ 設為0.4時，會得到最好的結果，以下是圖像經過不同 γ 值轉換後的結果。

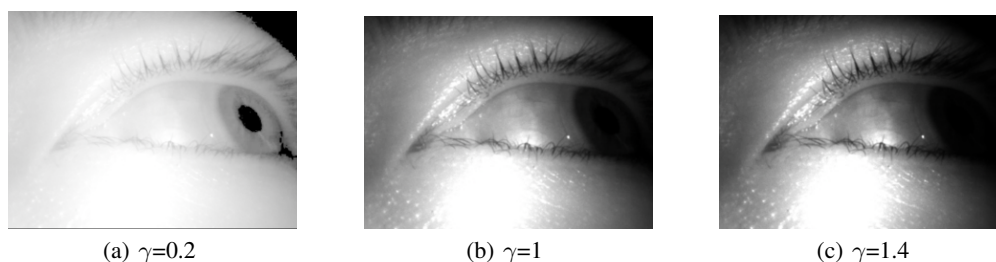


Figure 1: 圖像經過不同 γ 值轉換後的結果

由上圖可以發現，經過伽瑪校正後的圖像能夠凸顯瞳孔的位置和形狀，讓我們在訓練神經網路時更加得心應手。

2.2 Data augmentation

Data augmentation	Kernel size	Sigma
Gaussian Blur	Randomly choose between 3,5,7	Randomly choose between 1 to 5
Gaussian Noise		Randomly choose between 0.02 to 0.08
Rotation		
HorizontalFlip		
Translation		
Perspective transformation		

2.3 Data postprocessing

為了讓我們最終的輸出接近橢圓形，我們做了兩種後處理，其一是形態學(Morphology)中的閉合(closing)，另一個則為找到最符合圖形之橢圓形(fitEllipse)，兩者皆為opencv中提供的函式

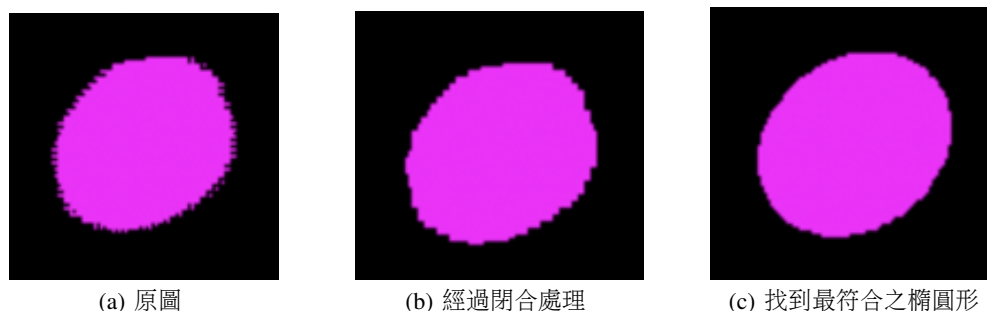


Figure 2: 圖像經過後處理的結果

由上圖可以發現，經過後處理的結果更加接近瞳孔的形狀，並且在圖形的邊緣更圓滑平整，不會有缺口或是孔洞的發生。

2.4 Hyperparameter Choices

在資料前處理，伽瑪校正使用 $\gamma = 0.4$ 較佳，並做歸一化，mean和std都設定0.5。切割模型訓練使用CrossEntropyLoss，lr = $2e-4$ ，optimizer為Adam，betas = (0.9, 0.999)，weight decay = 0.01，validation ration = 0.1，epoch= 200。分類模型訓練也是使用CrossEntropyLoss，但加上weight = [0.9, 0.1]，平衡開眼閉眼數量不均問題，epoch = 30，其他設定和切割模型相同。

另外值得注意的是，因已將問題拆成兩獨立問題，故為了使RITnet可以更專注在學

習切割出pupil位置，在訓練時將GT是閉眼全黑圖從Dataset中移除，只讓他學習可以切割出pupil位置的data。

3 Experiments

我們在測試了許多Segmentation Model，將confidence全部都給1的情況下將不同model測試在S5上。

Model	S5 score
Unet	0.648
Att unet	X
DeeplabV3 resnet101	0.638
DeeplabV3 hrnet	X
Segformer	0.630
RITnet	0.659

上述結果，X表示在validation就表現較差就無上傳，此結果可知，在其他benchmark上表現較佳的model不一定適合這次的pupil tracking。

4 Conclusion

這次期末專題，我們得到了Top 1的佳績，較為有效改善score如下列幾點

1. Gamma Correction
2. Data Augmantation
3. 訓練RITnet將閉眼全黑圖去除
4. 使用簡單而不複雜的神經網路

只要使用以上幾點skill，在S5 public就可以直接達到95的高分。

References

- [1] <https://ieeexplore.ieee.org/document/9055424>
- [2] <https://github.com/jaegal88/pupil-shape-prior>
- [3] <https://arxiv.org/abs/1812.07032>
- [4] https://github.com/yiskw713/boundary_loss_for_remote_sensing
- [5] <https://arxiv.org/pdf/2007.09600.pdf>
- [6] <https://github.com/RSKothari/ElSeg>
- [7] <https://github.com/AayushKrChaudhary/RITnet/tree/master>
- [8] <https://www.kaggle.com/code/bigironsphere/loss-function-library-keras-pytorch/notebook>
- [9] <https://jason-chen-1992.weebly.com/home/-gamma-correction>
- [10] https://blog.csdn.net/weixin_41693877/article/details/121370452
- [11] <https://zhuanlan.zhihu.com/p/72783363>