

Machine Learning Project-3

Kuan-Lin Chen

Department of Applied Mathematics

National Chung Hsing University

Taichung, Taiwan

04131045@gm.scu.edu.tw

I. INTRODUCTION

這是第三次報告，我們在這裡主要會展示實作的結果，在第一篇paper中，我們的實驗結果高達0.943與原paper的0.96與0.98，皆相差在0.1的準確度之內，在第二篇paper，因為我們lab主要是在做跟人臉相關的研究，所以第二篇的data從原本的鳥類，我們換成第一篇paper的data，使用的模型架構則是第二篇的模型架構，使用第二篇的模型架構，我們得到了0.95的準確度，以下會開始介紹實驗細節。

II. DETAIL FOR MODEL

在這裡我們會分別介紹兩篇論文裡面的實作細節，包含參數的設定使用的環境...等等。

A. FaceNet2ExpNet: Regularizing a Deep Face Recognition Net for Expression Recognition - 2017 CVPR

在第一篇論文(FaceNet2ExpNet)中，會先把VGG 做fine tune，用這個fine tune好的VGG16中pool5的這一層當作我初始CNN中conv.的label先去訓練我要建構的CNN中裡面的conv.，這個階段完成後接下來就是把這個上個階段訓練好的CNN當作我的初始，先接上一層的1*1 conv. 去將低表情網路與人臉網路之間的差距，爾後在接上一層256維的fully connect，這個階段總共會訓練50個epoch，訓練影像及測試影像則是彩圖，且每張大小為224*224，output label總共有8個(原paper)，在收集資料時發現資料中的label只有7類，所以我在分類時做預測時雖然也是輸出8類，但是出來的結果就是少的那一類一直為0；這裡使用的loss function則是cross entropy。

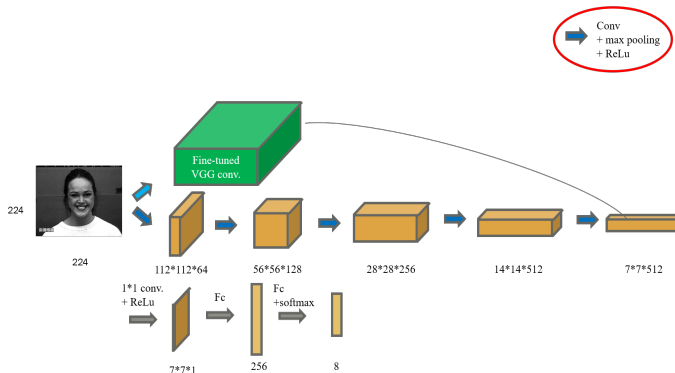


Fig. 1. FaceNet2ExpNet - model

接下來我會放上這篇paper的實作結果圖，原paper有做這個方法與其他方法的比較，並且有做6類及8類的預測準確度，我實驗到這預測的準確度有維持在0.943，然後放上我跟原本paper所做confuse matrix的圖。

Method	Average Accuracy	#Exp. Classes
CSPL [16]	89.9%	Six Classes
AdaGabor [35]	93.3%	
LBPSVM [36]	95.1%	
3DCNN-DAP [21]	92.4%	
BDBN [19]	96.7%	
STM-ExpLet [20]	94.2%	
DTAGN [22]	97.3%	
Inception [23]	93.2%	
LOMo [37]	95.1%	
PPDN [9]	97.3%	
FN2EN	98.6%	
AUDN [18]	92.1%	Eight Classes
Train From Scratch (BN)	88.7%	
VGG Fine-Tune (baseline)	89.9%	
FN2EN	96.8%	

Fig. 2. FaceNet2ExpNet - 原論文準確率結果比較

True Label \ Predicted Label	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	0.752	0.222	0.000	0.004	0.022	0.000
Disgust	0.073	0.873	0.019	0.015	0.020	0.000
Fear	0.000	0.005	0.949	0.000	0.046	0.000
Happy	0.000	0.000	0.092	0.908	0.000	0.000
Sad	0.088	0.000	0.028	0.000	0.884	0.000
Surprise	0.000	0.003	0.077	0.000	0.000	0.920

Fig. 3. FaceNet2ExpNet - 六類confuse matrix

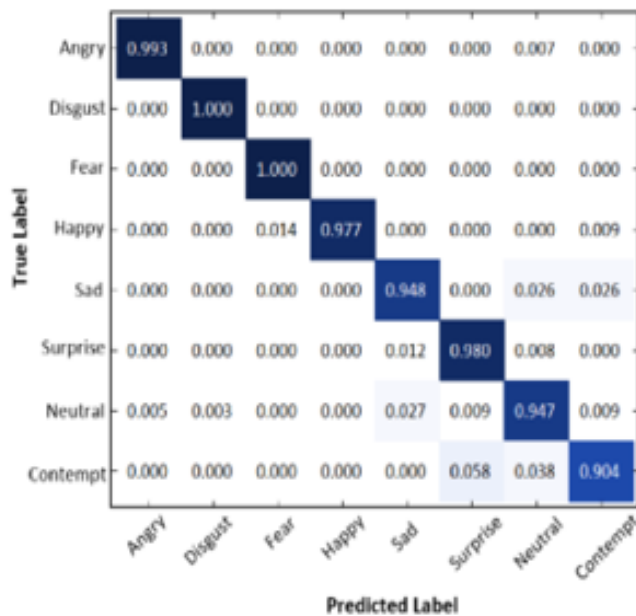
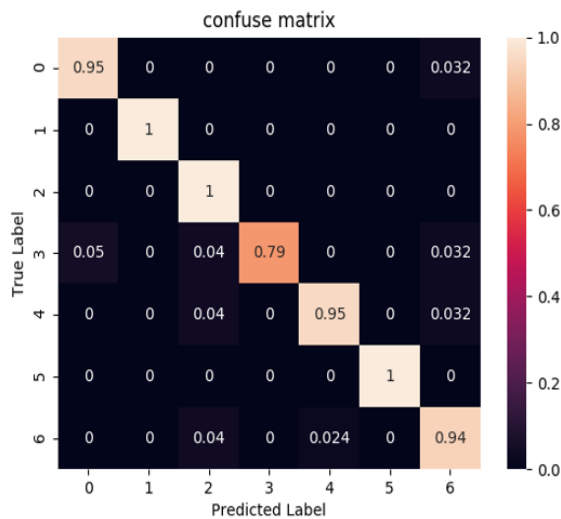


Fig. 4. FaceNet2ExpNet - 八類confuse matrix



accuracy: 0.946666579246521

Fig. 5. FaceNet2ExpNet - 實驗後的七類confuse matrix

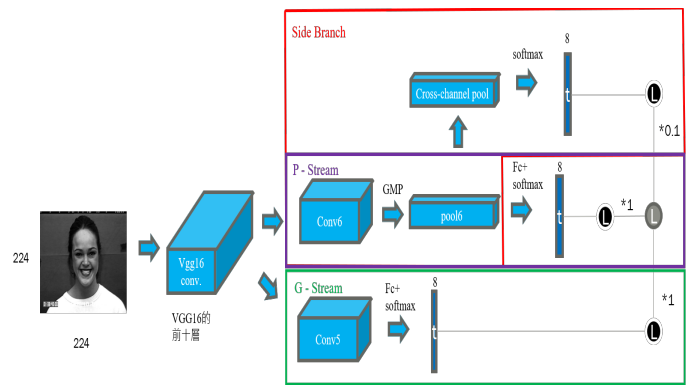
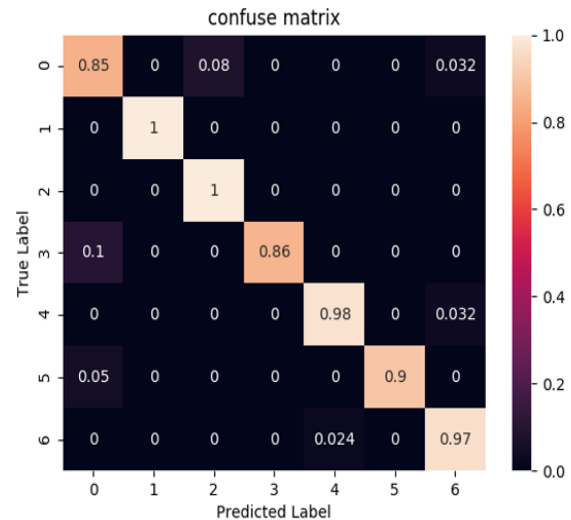


Fig. 6. Fine-grained model

接下來這張圖是所用這個模型所做出來的結果圖。



accuracy = 0.9466666666666667

Fig. 7. Fine-grained mode - confuse matrix

B. Learning a Discriminative Filter Bank within a CNN for Fine-grained Recognition -2018 ECCV

接下來介紹這篇論文的實驗成果，在這裡我們有更換dataset，將原論文所使用的鳥類data換成上篇論文的dataset，所以所預測的東西更上一篇一樣，使用的模型則是這篇論文的模型，參數上的設定也跟這篇論文所設定的大致相同，不同的只有total epoch我們設定50，照片大小設定224*224，其他架構設定都跟這篇論文一樣：optimizer is SGD,and batch size is 32,且VGG16是不用fine tune。