

## 背景与要求

人脸识别由于技术难度偏小，应用比较广泛。成为近年来大热且很成功的应用方向。本次大作业，我将从深度方法和传统方法着手，进行效果对比。

人脸识别主要在于鉴别人脸所属的身份。我将其理解为两个任务。一是性别识别。而是身份识别。采取的数据集为Face Recognition Data数据集。

## 神经网络方法--facenet

facenet 的主要思想是 直接学习图像到欧式空间的映射。两张图片对应特征的欧式空间上点的距离对应着两张图片是否相似。这很大程度上避免了因为姿态的变化造成的识别不准。虽然本数据集在姿态上变化不大。但也取得了较好的成果。

特征提取，我采用的神经网络是inceptionV1。将图片送入该网络学习后获取512维的特征。然后将特征进行L2归一化。也就是

$$f(x)^2 = 1, f(x) \text{为} 512 \text{维特征}$$

这样保证了所有特征都在一个超球面中。

接下来采用triplet loss 学习样本的可分性。这对于正确划分人脸是最重要的。输入是<a,p,n>三元组。

a: anchor是当前图像， p:positive 代表与a同一类的样本， n: negative 代表与a不同类的样本

优化目标为  $L = \max(d(a,p) - d(a,n) + \text{margin}, 0)$ 。

- 这意味着，当 $d(a,p) + \text{margin} < d(a,n)$ 时， $L=0$ ，分类正确，无需优化
- 当 $d(a,n) < d(a,p)$ 时，分类错误。需要优化
- 当 $d(a,p) < d(a,n) < d(a,p) + \text{margin}$ ，说明虽然已经被正确分类，但是类间距离依然较近，需进一步优化

可以看到，facenet的原则与感知机准则有较多相似之处。但它优化的是神经网络提取的特征。相较人类提取的特征更加高维，描述更准确。

## face94数据集，身份识别

数据集共3060张图片，153个人，其中训练集取80%，测试集取20%，验证集属于测试集，占总数据集的5%，大约每人会被测试一次。第一次实验结果验证集准确率99.06%。但后来发现测试错误的来自female/gotone文件夹，该文件夹并非仅一个女性。而是多个女性的混合。这说明我们的模型不仅可以正确识别身份还有抗干扰能力。对于少数标错的训练对象，也能较好的识别出来。

```
131, 'saduah': 132, 'saedwa': 133, 'sandm': 134, 'sbains': 135, 'sidick': 136, 'sjbeck': 137, 'skumar': 138,
'slbirc': 139, 'smrobb': 140, 'spacl': 141, 'spletc': 142, 'svkriz': 143, 'swewin': 144, 'swsmitt': 145, 'tony'
: 146, 'voudcx': 147, 'vpsavo': 148, 'vstros': 149, 'whussa': 150, 'wjalbe': 151, 'yfhisie': 152}
accuary: gotone res: astefa
accuary: gotone res: drbost
total test 152 imgs!
```

为保证免受该文件夹的干扰，我们再次测试。

在第11个epoch测试准确率达到100%，在13个epoch准确率为99.84%，逐个测试时没有识别错的人物。

实验截图如下：

```

Epoch 11/13
-----
Train |    76/76 | loss:    0.0078 | fps:  160.1023 | acc:    1.0000
Valid |    19/19 | loss:    0.0086 | fps:  781.4340 | acc:    1.0000

Epoch 12/13
-----
Train |    76/76 | loss:    0.0081 | fps:  159.0397 | acc:    1.0000
Valid |    19/19 | loss:    0.0091 | fps:  734.8524 | acc:    0.9984

Epoch 13/13
-----
Train |    76/76 | loss:    0.0077 | fps:  159.7012 | acc:    1.0000
Valid |    19/19 | loss:    0.0091 | fps:  675.5400 | acc:    0.9984
total test 152 imgs!
[ltt@gpu4 faceRecognize-IMGprocess]$ 

```

## face96数据集

共152个人，大约每人20张图片，80%训练集，20%验证集，5%逐个测试。在第8个epoch 验证集准确率达99.84%。在第13个epoch准确率为99.69%。测试的152张图片中只有一张出错。如图所示，左边为测试图片，右边为识别到的身份。一定程度上，两位小哥长得比较像，且背景一致。可以看出，facenet会受到背景干扰。



实验截图如下：

```

Epoch 8/13
-----
Train |    77/77    | loss:    0.0072 | fps:   173.4380 | acc:    1.0000
Valid |    20/20    | loss:    0.0181 | fps:   787.9216 | acc:    0.9984

Epoch 9/13
-----
Train |    77/77    | loss:    0.0071 | fps:   171.5717 | acc:    1.0000
Valid |    20/20    | loss:    3.5008 | fps:   788.9122 | acc:    0.2766

Epoch 10/13
-----
Train |    77/77    | loss:    0.0065 | fps:   172.0268 | acc:    1.0000
Valid |    20/20    | loss:    0.0204 | fps:   817.7003 | acc:    0.9969

Epoch 11/13
-----
Train |    77/77    | loss:    0.0062 | fps:   172.1174 | acc:    1.0000
Valid |    20/20    | loss:    0.0121 | fps:   768.2029 | acc:    0.9969

Epoch 12/13
-----
Train |    77/77    | loss:    0.0062 | fps:   171.0255 | acc:    1.0000
Valid |    20/20    | loss:    0.0103 | fps:   752.1503 | acc:    0.9969

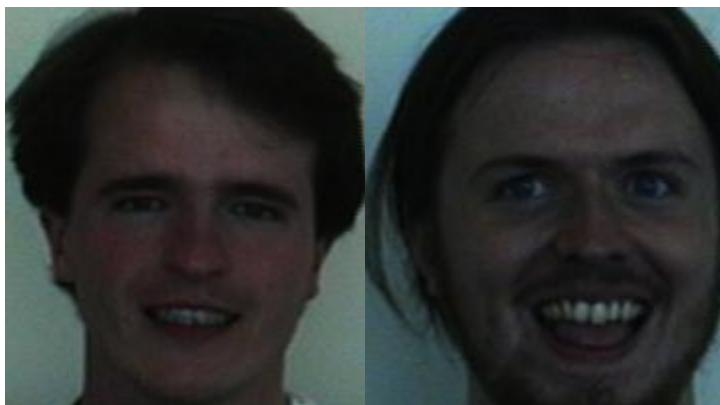
Epoch 13/13
-----
Train |    77/77    | loss:    0.0060 | fps:   172.3503 | acc:    1.0000
Valid |    20/20    | loss:    0.0116 | fps:   792.1564 | acc:    0.9969
accuracy: ldebn res: jlrum
total test 152 imgs!
[11:09:44] faceRecognize: IMGs processed: find / - name: jlrum

```

## grimace数据集

该数据集由于光线较暗且人物表情变化大，因此确实比较难。但facenet表现出了良好的性能。

在第一次尝试中，模型在epoch13测试准确率仅有90.62%。采用了18张图片逐张测试。将pat认作glen。



实验结果如图：

```

Epoch 11/13
-----
Train |    9/9    | loss:    0.0046 | fps:   151.2425 | acc:    1.0000
Valid |    3/3    | loss:    8.7072 | fps:   313.2836 | acc:    0.0833

Epoch 12/13
-----
Train |    9/9    | loss:    0.0036 | fps:   151.6997 | acc:    1.0000
Valid |    3/3    | loss:    1.7388 | fps:   322.8234 | acc:    0.5938

Epoch 13/13
-----
Train |    9/9    | loss:    0.0048 | fps:   152.0829 | acc:    1.0000
Valid |    3/3    | loss:    0.1502 | fps:   307.2332 | acc:    0.9062
accuary:  pat res:  glen
accuary:  pat res:  glen
total test 18 imgs!

```

很明显模型还未收敛，因此进行第二次试验。而此次模型在epoch13达到100%准确率，并在接下来两轮测试中都为100%准确。逐张测试也没有出错。实验截图如下：

```

Epoch 12/15
-----
Train |    9/9    | loss:    0.0069 | fps:   158.3916 | acc:    1.0000
Valid |    3/3    | loss:    0.1354 | fps:   319.3457 | acc:    0.9375

Epoch 13/15
-----
Train |    9/9    | loss:    0.0070 | fps:   151.3113 | acc:    1.0000
Valid |    3/3    | loss:    0.0026 | fps:   229.2427 | acc:    1.0000

Epoch 14/15
-----
Train |    9/9    | loss:    0.0053 | fps:   158.2228 | acc:    1.0000
Valid |    3/3    | loss:    0.0049 | fps:   326.4247 | acc:    1.0000

Epoch 15/15
-----
Train |    9/9    | loss:    0.0052 | fps:   156.7871 | acc:    1.0000
Valid |    3/3    | loss:    0.0028 | fps:   315.6908 | acc:    1.0000
total test 18 imgs!

```

## 性别识别：

将face94数据集的female 和 malestaff一起训练，从而保证男女数据集的均衡。该分类更加简单，毫无疑问模型的识别效率也是100%。在epoch8已经达到100%识别率

```

Epoch 7/15
-----
Train |   20/20   | loss:    0.0026 | fps:   182.7366 | acc:    1.0000
Valid |    5/5    | loss:    1.6341 | fps:   512.2762 | acc:    0.6438

Epoch 8/15
-----
Train |   20/20   | loss:    0.0014 | fps:   192.1058 | acc:    1.0000
Valid |    5/5    | loss:    0.0100 | fps:   382.3565 | acc:    1.0000

```

## 传统方法--pca + SVM判别法

### PCA 数据降维，提取特征

一张192X192图片的像素数高达36864，因此如何在尽可能减少信息丢失的基础上，降低数据分析的工作量是个很重要的问题。而尽可能地减少变量之间的相关性，也有助于尽可能多的提取信息。主成分分析就属于这种算法。

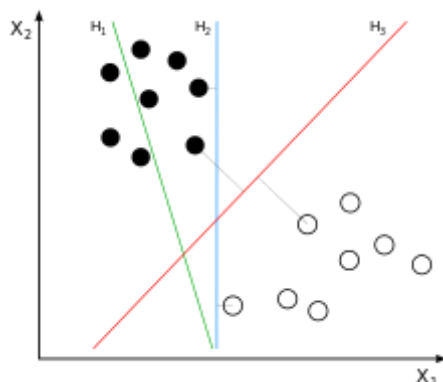
PCA通过将N维特征映射到K维实现降维。第一维选择原始数据中方差最大的方向，其他维依次选择与已确定维度正交且方差最大的方向。这样剩下的维度方差几乎为0，可以忽略不计。具体实现表现为取协方差矩阵中特征值最大的前K个向量。具体实现如下：

```
sklearn.decomposition.PCA(n_components=10).fit(dataset)
```

该对象保存了前10个方差最大的特征。

## SVM 数据分类

获得图像的前几个特征后，需要选取适当的分类超平面，我选择了SVM。一般我们认为更好的分类方法应该是两类之间的距离越远越好。类内的距离越近越好。SVM通过寻找特征空间上**间隔最大**的线性分类平面获得更好的分类结果。如图所示，红线保证了两类之间的间隔最大，因此是最好分类面



SVM拓展为高维后可以用于多分类。

我结合了PCA和SVM进行人脸的识别，并采用网格搜索 `GridSearchCV` 对模型简单优化。具体实现如下：

```
n_components = 10
pca = PCA(n_components=n_components)

param_grid = {'C': [1e3, 5e3, 1e4, 5e4, 1e5],
              'gamma': [0.0001, 0.0005, 0.001, 0.005, 0.01, 0.1], }
clf = GridSearchCV(SVC(kernel='rbf'), param_grid)
pca = pca.fit(training_data)

transformed = pca.transform(training_data)
clf.fit(transformed, train_lab)
```

## 性别识别

这是较为简单的二分类问题。我们采用face94数据集，主要使用了female和malestaff的图像。随机选取其中的90%为训练集，其余为测试集。最终模型在测试集上的准确率达到83.75%。

在调参的过程中，我还发现，并不是选取的主成分越多越容易分开。选取前20个主成分的最高准确率大约为68.75%。而降为10个主成分后，效果大大提升了。这说明，男女性别的差别较为明显，多余的成分只会干扰SVM的判断

```
40 images_labels_row_col_CNT = prepare_dataset(images_labels_directory)

DEBUG CONSOLE PROBLEMS OUTPUT TERMINAL
2: Python

Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
test all 80 acc: 0.8375 67
```

```
61 pca = pca.fit(training_data)
62
63 transformed = pca.transform(training_data)
64 clf.fit(transformed,train_lab)
65 #####train finish
66
67 def get_key(dict_, value):
68     for k, v in dict_.items():
69         if v == value:
70             return k
71 ri = 0
72 for i in range(len(test_img)):
```

```
DEBUG CONSOLE PROBLEMS OUTPUT TERMINAL
2: Python

Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
test all 80 acc: 0.6875 55
PS E:\2020-Spring\IMGProcecss\Face\Face-Recognition>
```

但是PCA对光照十分敏感，而图中由于某些任务的光照差别过大，因此我们采用了直方图均衡的办法对图像预处理。从而使PCA减少对光照的关注。更关注图片的结构。但是效果不升反降。准确率为68.75%。

下面四张图片是直方图均衡化后的结果，可以看到，人与人之间的肤色差别被大大缩小了。





因此我推测，男女性别的差异还表现在了肤色上。或者是由于数据集太小对男性女性的判断过分依赖光照变化，过拟合了。

```

55 test_img = np.array(np.array(images)[test_idx])
56 test_img = test_img.astype(int)

DEBUG CONSOLE  PROBLEMS  OUTPUT  TERMINAL
2: Python
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
Predicted By Classifier : 1.0 Actual : 0.0
test all 80 acc: 0.6875 55
PS E:\2020-Spring\IMGProcess\Face\Face-Recognition>

```

## 身份识别

我选取了face94中三位女性和两位男性的所有照片，使用类似的方法进行分类，但是结果并不理想，正确分类的概率只有38.1%，身份识别需要在有限的样本中学习多种类型，因此这个结果也比较正常。

```

"This module will be removed in 0.20.", DeprecationWarning)
D:\software4Code\anaconda3\lib\site-packages\sklearn\grid_search.p
_selection module into which all the refactored classes and functi
DeprecationWarning)
all person img 103
test all 21 acc: 0.38095238095238093 8
PS E:\2020-Spring\IMGProcess\Face\Face-Recognition>

```

我们再次采取直方图均衡的方法对图像预处理，在该问题上，直方图均衡同样表现的逊色，准确率仅达到23.8%

```
m that of this module. This module will be removed in 0.20.  
  "This module will be removed in 0.20.", DeprecationWarning)  
D:\software4Code\anaconda3\lib\site-packages\sklearn\grid_search  
_selection module into which all the refactored classes and fu  
  DeprecationWarning)  
all person img 103  
test all 21 acc: 0.23809523809523808 5  
PS E:\2020-Spring\IMGProcecss\Face\Face-Recognition> |
```

## 总结

---

我们在 Face Recognition Data 数据集上进行了一系列实验，尝试对性别或者姓名分类。受限于样本数量，测量性别的时候模型表现远高于确定身份。

此外，PCA对光照十分敏感。不同肤色会很大影响模型的判断，这不失为一个好办法，也许不必扼杀这样的表现。

PCA特征数量的选取并不是越多越好。适当的数量免于模型受其他因素干扰而判断失误。

支持向量机对于分类问题表现较好，但是如果类内样本少而种类较多，SVM的表现会大大下降