

基于多分类模型加权投票法的人脸微笑检测

冯泽安, 王 鹏

(西安工业大学 电子信息工程学院 陕西 西安 710021)

摘 要: 为了进一步提高人脸微笑检测率并解决微笑检测系统用于训练标签数据不足的问题, 结合人脸图像的纹理和几何特点, 应用了一种基于多分类模型加权投票法的微笑检测方法。在经预处理和直方图均衡化的面部图像上, 利用局部二进制模式(LBP)和Gabor小波变换提取局部的、抽象的特征, 同时以人脸特征点检测作为补充, 构建了四种不同的分类模型(UPN、GPS、AdaBoost和LDA), 分别对人脸图像进行分类检测, 同时结合各模型之间互补和各自对微笑检测的优势, 通过计算权值对各结果进行加权投票, 得到面部图像的最终检测结果。实验结果显示该方法的有效性, 在公开的GENKI-4K人脸数据集上获得了95.8%的微笑检测率, 比单个分类模型的平均检测率提高了10.3%, 与该数据集的最新的微笑检测率相等。

关键词: 微笑检测; 人脸图像; 纹理; 几何; 分类模型; 加权投票法

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2019)02-0081-06

doi: 10.3969/j.issn.1673-629X.2019.02.017

Facial Image Smile Detection Based on Multi-class Model Weighted Voting

FENG Ze-an, WANG Peng

(School of Electronic Information Engineering, Xi'an Technological University, Xi'an 710021, China)

Abstract: In order to further improve the face smile detection rate and solve the problem that the smile detection system is not enough for training tag data, we apply a smile detection method based on multi-classification model weighted voting in combination with the texture and geometric characteristics of human face image. On the facial image with pre-processing and histogram equalization, local and abstract features are extracted by the local binary model (LBP) and Gabor wavelet transform. At the same time, facial feature point detection is used as a supplement to construct four different classification models (UPN, GPS, AdaBoost and LDA). Combined with the complementarity between the models and their respective advantages on smile detection, the final detection results of facial images are obtained by calculating the weights and voting the results weighted. The experiment shows the effectiveness of this method. The detection rate of 95.8% is obtained on the open GENKI-4K face dataset, which is 10.3% higher than the average detection rate of a single classification model and the same as the latest smile detection rate.

Key words: smile detection; face image; texture; geometry; classification model; weighted voting

0 引 言

对无约束面部图像的微笑检测是一个具有挑战性的问题。作为最具信息含义的表达之一^[1], 微笑表达了一些基本情绪, 如快乐和满意度, 于是在不同领域有多种应用, 如人类行为分析^[2]、照片选择^[3]、患者监护^[4]等。近年来, 局部二进制模式^[5](LBP)常用于微笑检测, 在Viola-Jones级联分类器的微笑检测器^[6]的分类准确率可达到96.1%。虽然分类准确性很高, 但对所采用的图像样本要求同样非常高, 比如必须使用

面部正面照。还有使用AdaBoost来选择和组合弱分类器^[7], 并在GENKI-4K图像数据集上达到了88%的微笑检测率。一般用于微笑检测的面部信息^[8]可以分为两大类: (I) 几何信息和(II) 纹理(或外观)信息。几何信息一般包含面部特征或头骨的形状, 而纹理信息主要考虑面部图像像素强度或由像素构成的图案(像素图)。但在面部微笑检测的文献中, 通常将几何信息与纹理信息相结合, 而不是单独使用。最新的微笑检测算法是基于卷积神经网络(CNN)的算法, 该算

收稿日期: 2018-03-29

修回日期: 2018-07-31

网络出版时间: 2018-11-15

基金项目: 陕西省科技计划重大重点项目(2017ZDXM-GY-114)

作者简介: 冯泽安(1993-), 男, 硕士研究生, 研究方向为图像处理; 王 鹏, 副教授, 研究方向为图像处理和模式识别。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.tp.20181115.1050.072.html>

法可以对低级特征组成的高级特征的层次表示^[9]进行学习, CNN 实现的高分类精度使这种算法成为微笑检测领域中的先进技术, 但需要更多的样本来训练网络。例如, 对一个由 4 个卷积层和 1 个完全连接层组成的 CNN 架构^[10]进行微笑检测, 其准确率达到 94.6%, 这比先前在 GENKI-4K 数据集上使用的任何检测方法所获得的准确度都要高。文中应用了一种基于多分类模型加权投票的方法来进行人脸微笑检测。

1 多分类模型加权投票法

多分类模型加权投票法主要由 3 部分组成, 即预处理、四种分类模型和最后加权投票, 其总体结构如图 1 所示。其中 $R_i (i = 1, 2, 3, 4)$ 分别表示各模型的分类结果。

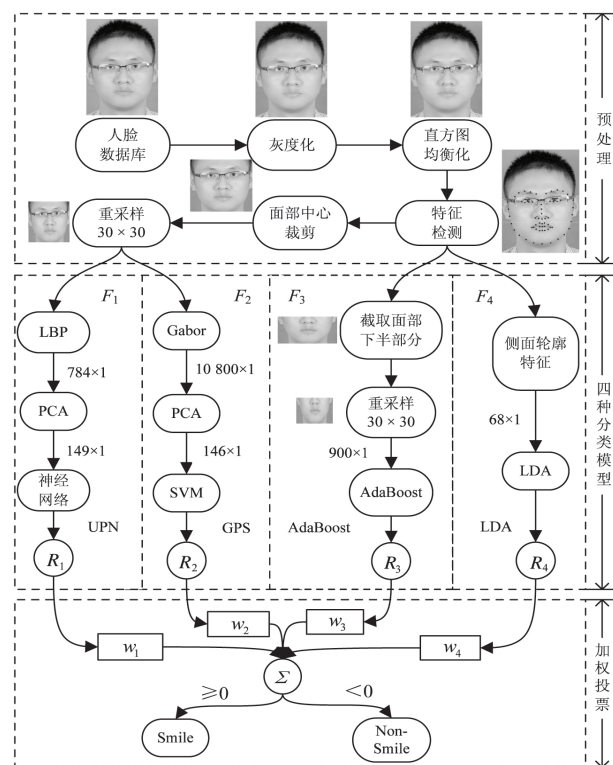


图 1 总体结构

1.1 预处理

预处理阶段包括对面部图像进行灰度缩放和直方图均衡化、特征检测, 并裁剪出面部的中心, 最后按照 30×30 的均匀网格 (1 格等于 1 像素点) 对其进行重采样。

(1) 灰度和直方图均衡化: 在图像处理算法中, 为了加快处理速度 (可减小后期的计算量), 需将彩色图像灰度化, 但为了获得更好的对比度^[11], 需将灰度图像进行直方图均衡化。这在此次检测中很重要, 因为该方法可获得更多的面部图像细节。

(2) 特征检测: 特征检测方法 ASM^[12] (active shape model) 或 CLNF^[13] (constrained local neural

field) 在人脸识别中应用广泛, 文中选择 CLNF 方法; 一张完整的面部图片上总共 68 个特征点 (17 个面部侧轮廓点, 嘴唇有 20 个标记点, 每只眼睛上 6 个点, 鼻子 9 个点和每个眉毛上 5 个点)。图 2 为在几个面部样本上检测到的面部特征点。

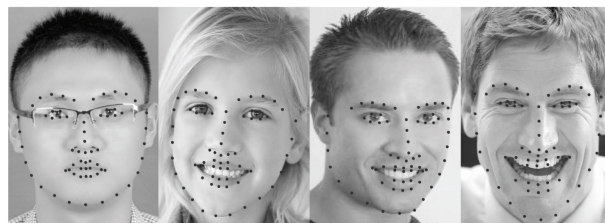


图 2 面部样本上的特征点

(3) 截取面部中心图和重采样: 文中检测的图像仅考虑脸部中心部位, 以防止额外的信息被利用, 从而影响检测结果。因此, 得对图像进行裁剪。现今, 对人脸进行检测和分割的方法有很多, 例如 Viola-Jones 人脸检测器^[14]; 同时还有利用肤色来对图像进行分割而得到脸部中心图。文中方法是根据面部侧面特征点形成的轮廓对人脸进行裁剪。

1.2 分类模型 F_i (UPN)

分类模型 F_i 包括 Uniform LBP、PCA 和神经网络。

(1) Uniform LBP 模式: LBP (local binary pattern, 局部二值模式) 是一种用来描述图像局部纹理特征的算子^[15], 具有旋转不变性和灰度不变性等显著的优点, 广泛用于图像局部纹理特征提取。其基本思想为: 在 3×3 的窗口 (以像素点为基准) 内, 以窗口中心像素为阈值, 将相邻的 8 个像素的灰度值与其进行比较, 若周围像素值大于中心像素值, 则该像素点的位置被标记为 1, 否则为 0。从而, 3×3 邻域内的 8 个点经比较可产生 8 位二进制数 (通常转换为十进制数即 LBP 码, 共 256 种), 即得到该窗口中心像素点的 LBP 值, 并用这个值来反映该区域的纹理信息。在像素 3×3 的邻域内, 对于中心像素 (强度为 i_c) 和相邻的 P 个元素 (强度为 $i_p, p = \{0, 1, \dots, P-1\}$), 计算公式为:

$$LBP = \sum_{p=0}^{P-1} 2^p s(i_p - i_c) \quad (1)$$

其中, $s(x)$ 为符号函数, 有:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (2)$$

但由于图像的旋转会产生不同的 LBP 值, 同时绝大多数 LBP 模式得到的二进制数大多只包含两次从 1 到 0 或从 0 到 1 的跳变, 为了尽可能降低图像特征的维数, 于是引入 LBP 的等价模式 (Uniform LBP), 计算方法为:

$$LBP^u = \begin{cases} \sum_{p=0}^{P-1} s(i_p - i_c), & \text{if } U(LBP) \leq 2 \\ P+1, & \text{otherwise} \end{cases} \quad (3)$$

其中, $U(\text{LBP})$ 表示二进制数从 0 到 1 或从 1 到 0 的跳变次数。

由于 Uniform LBP 比传统 LBP 效果更好^[16]; 因此, 在第一个分类模型中, 使用 Uniform LBP 方法在 30×30 的面部图像上进行特征提取。

(2) 主成分分析 (principal components analysis, PCA): PCA 是最重要的降维方法之一。在数据压缩消除冗余和数据噪音消除等领域都有广泛的应用, 顾名思义, 就是找出数据里最主要的方面代替原始数据。由于图像数据一般都具有很高的维度, 因此可以使用 PCA 进行特征提取和缩减。其降维方法为:

如果有 m 个面部样本, 且每个训练样本 i 对应的 LBP^u 被转换为一个向量 $\mathbf{x}^{(i)}$, n ($n = 28 \times 28 = 784$) 维样本集 $D = (\mathbf{x}^{(1)} \mathbf{x}^{(2)} \cdots \mathbf{x}^{(m)})$ 降维后为 n' ($n' = 149$) 维样本集 $D' = (\mathbf{z}^{(1)} \mathbf{z}^{(2)} \cdots \mathbf{z}^{(m)})$; 样本的协方差矩阵为:

$$C = \sum_{i=1}^n (\mathbf{x}^{(i)} - \bar{\mathbf{x}}) (\mathbf{x}^{(i)} - \bar{\mathbf{x}})^T \quad (4)$$

其中, $\bar{\mathbf{x}}$ 为训练样本的平均值 ($\bar{\mathbf{x}} = \frac{1}{m} \sum_{j=1}^m \mathbf{x}^{(j)}$)。

而协方差矩阵的特征向量 (\mathbf{P}_n) 为:

$$C \mathbf{P}_n = \lambda_n \mathbf{P}_n \quad (5)$$

其中, λ_n 为 C 的特征值。

取出最大的 n' 个特征值对应的特征向量 ($\omega_1, \omega_2, \cdots, \omega_{n'}$) 将所有的特征向量标准化后, 组成特征向量矩阵 W ; 对样本集中的每一个样本 $\mathbf{x}^{(i)}$, 转化为新的样本 $\mathbf{z}^{(i)} = W^T \mathbf{x}^{(i)}$; 最后得到输出样本集 D' 。于是 n' 个特征值对应的特征向量将原始样本转换到 n' 个特征向量构建的新空间中。在测试阶段, 将转换后的 LBP^u 投影到新的特征向量空间中, 可以使图像的维度降得很低。

(3) 反向传播神经网络: 卷积神经网络广泛用于人脸识别, 并且检测率非常高。但在该模型中, 样本图像经 PCA 降维后输出为 149×1 的灰度值, 因此, 使用普通的反向传播神经网络即可完成分类。如图 3 所示, 该神经网络由输入层 149 个神经元, 隐藏层 10 个神经元和 1 个输出神经元所构成。微笑样本图像输出标签用 1 表示, 而非微笑用 -1 表示, 其他三种分类模型的标签表示方法与此一致。

1.3 分类模型 F_2 (GPS)

分类模型 F_2 由 Gabor 滤波, PCA 和 SVM 组成。

(1) Gabor 小波变换: 对预处理后的图像进行 Gabor 小波变换, 由于 Gabor 函数形成的二维 Gabor 小波变换具有在空间域和频率域同时取得最优局部化的特性, 能充分对物体纹理特征进行提取; 其属于加窗傅里叶变换, 是 Gabor 理论和小波理论结合的产物。受其

启发, Gabor 小波变换在此模型中以 4 种尺度 (7, 12, 17 和 21) 同时从六个方向 (角度分别为 $0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ$ 和 150°) 对灰度图像进行变换。二维的 Gabor 小波变换^[17]的计算公式为:

$$g(x, y; \lambda, \theta, \varphi, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \sin\left(\frac{2\pi x'}{\lambda} + \varphi\right) \quad (6)$$

$$x' = x \cos \theta + y \sin \theta \quad (7)$$

其中, λ, θ, φ 和 σ 分别为波长 (比例), 方向, 相位偏移和高斯标准差。

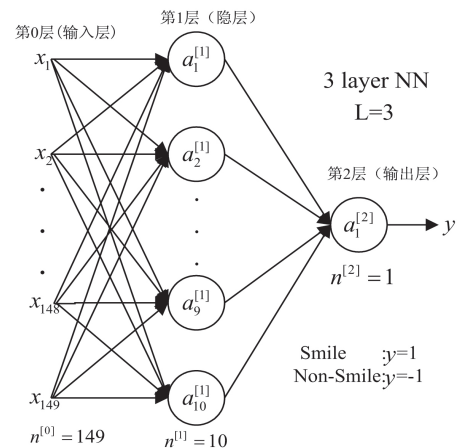


图3 反向传播神经网络

在该分类模型中, Gabor 小波变换具有非常重要的作用。图 4 为对一个样本面部图像应用 24 个小波核卷积的结果; 人脸的某些部分在特定的尺度和方向上呈现剧烈的变化趋势, 通过 Gabor 小波变换之后, 这些变化可以很直观地表现出来。

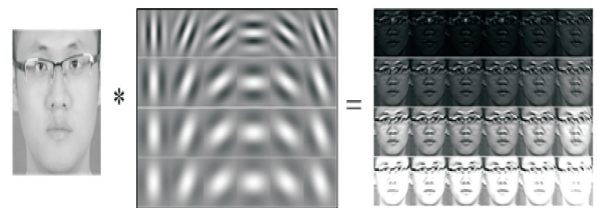


图4 图像卷积结果

(2) 主成分分析: 在进行 Gabor 滤波时使用了 24 个小波核, 同时每个核的尺寸为 30×30 , 因此, 将滤波器结果连接起来并将其整合成一个矢量之后, 该矢量维数将达到 21 600 ($30 \times 30 \times 24$); 为了避免出现维度危机, 与第一个模型中所述的一样, 再次使用 PCA 进行降维。

(3) 支持向量机 (SVM): SVM 基本模型为特征空间上间隔最大的线性分类器, 即学习策略为间隔最大化, 最终通过转化为一个凸二次优化问题的求解, 或者是在高维空间中寻找一个合理的超平面将数据点分隔开来, 其中涉及到非线性数据到高维的映射以达到数据线性可分的目的。在此模型中, 经 PCA 降维后, 其维数仍有 $146 \times m$, 属于高维空间。非线性 SVM 的决

策函数^[18]可以表示为:

$$y^{(i)} = \text{sign} \left(\sum_{i=1}^N \alpha^{(i)} y^{(i)} K(x^{(i)}, x^{(i)}) + b \right) \quad (8)$$

其中, x_i 和 x_i 分别表示训练和测试样本; b 为偏差; α_i 为拉格朗日乘数; N 为所有训练样本的总量; $y_i \in \{+1, -1\}$ 为训练样本的标签。

支持向量机通过映射函数 $\varphi(x)$, 将输入空间中的样本特征映射到高维特征空间, 从而降低计算量; 式 8 中 $K(x^{(i)}, x^{(i)})$ 为 SVM 的核函数, 与映射函数的关系为 $K(x^{(i)}, x^{(i)}) = \varphi(x^{(i)}) \varphi(x^{(i)})$ 。在此模型中, 选取径向基函数作为核函数:

$$K(x^{(i)}, x^{(i)}) = e^{-\|x^{(i)} - x^{(i)}\|^2} \quad (9)$$

1.4 分类模型 F_3 (AdaBoost)

分类模型 F_3 主要包括提取面部的下半部分, 对裁剪后的图像进行重采样, 然后使用 AdaBoost 进行分类。

(1) 提取脸部的下半部分: 脸部的下半部分^[19], 即从鼻尖到下巴, 包含影响检测结果的关键信息。换言之, 从纹理图案的平滑度或粗糙度的差异出发, 仅脸部的下部就可用于微笑检测。该模型提取下半脸、鼻尖、下巴和两个侧面轮廓特征信息。

(2) 重采样: 为了减少数据样本维数和保持维数均衡, 将提取的部分使用 30×30 均匀网格进行重采样。然后, 将重采样后的数据重新调整为 900×1 维的矢量。

(3) AdaBoost: AdaBoost 是一种迭代(集成学习)方法^[20], 既可以用于分类, 也可以用于回归, 其核心思想是针对不同的训练集训练同一个弱分类器, 然后把在不同训练集上得到的弱分类器集合起来, 构成一个最终的强分类器。该分类模型将用 AdaBoost 级联结构分类器(建立在弱分类器单层决策树之上)并对 $900 \times m$ 维的样本进行二元分类, 其流程为:

输入: 样本集 $D = (x^{(1)}, x^{(2)}, \dots, x^{(m)})$ 输出: 标签 $y_i \in \{-1, +1\}$, 弱分类器迭代次数 K

过程:

1. 初始化样本权重为: $W(1) = (w_{11}, w_{12}, \dots, w_{1m}); w_{1i} = \frac{1}{m}; i = 1, 2, \dots, m$

2. for $k = 1, 2, \dots, K$ do:

3. 使用具有权重 $W(k)$ 的样本集来训练数据, 得到弱分类器 $G_k(x)$

4. 计算 $G_k(x)$ 的分类误差率: $e_k = P(G_k(x_i) \neq y_i) = \sum_{i=1}^m w_{ii} I(G_k(x_i) \neq y_i)$

5: if $e_k > 0.5$ then break

6: 计算弱分类器的系数: $\alpha_k = \frac{1}{2} \log \frac{1 - e_k}{e_k}$

7: 更新样本集的权重分布: $w_{k+1,i} = \frac{w_{ki} \exp(-\alpha_k y_i G_k(x_i))}{Z_k}$

$i = 1, 2, \dots, m$

其中 Z_k 为规范化因子: $Z_k = \sum_{i=1}^m w_{ki} \exp(-\alpha_k y_i G_k(x_i))$

8: end for

输出: 强分类器: $f(x) = \text{sign}(\sum_{k=1}^K \alpha_k G_k(x))$

1.5 分类模型 F_4 (LDA)

分类模型 F_4 直接使用线性判别分析(LDA)对面部的侧面轮廓特征进行分类。

(1) 侧面轮廓特征点: 在预处理阶段已经对面部侧面轮廓特征进行了检测。而人类生物学^[21]显示, 人在微笑和非微笑时, 面部的几何特征有着明显的差异, 表现在脸部、眼睛、眉毛、下巴、眶上脊、颧骨、下颌骨、腭部、眼睛和鼻尖、眼睛和嘴唇中心之间的垂直距离, 或者鼻子宽度、嘴唇宽度、脸部高度和脸部阴影的侧面轮廓。但是这些特征可能会受到外界因素的影响, 如人的肥胖程度, 同时面部侧面轮廓的上半部分可能会被头发覆盖, 因此, 在此模型中仅考虑从左耳到右耳的下侧轮廓。然后, 将所有特征标记点的坐标整合成一个列向量送入分类器进行分类。

(2) 线性判别分析: LDA(linear discriminant analysis)不但是一种监督学习的降维技术, 同时也是一个强大的分类器^[22], 也就是说它的数据集的每个样本是有类别输出的。其思想可以用一句话概括, 就是“投影后类内方差最小, 类间方差最大”。就是要将样本数据在低维度上进行投影, 投影后希望每一种类别数据的投影点尽可能接近, 而不同类别的数据的类别中心之间的距离尽可能大。其二分类原理和流程为:

数据集 $D = (x^{(1)}, x^{(2)}, \dots, x^{(m)})$, 其中任意样本 $x^{(i)}$ 为 n (本模型中 $n = 68$) 维向量, $y^{(i)} \in \{-1, 1\}$; 定义 $N_j(j = -1, 1)$ 为第 j 类样本的个数, $X_j(j = -1, 1)$ 为第 j 类样本的集合, 而 $\mu_j(j = -1, 1)$ 为第 j 类样本的均值向量, 定义 $\sum_j(j = -1, 1)$ 为第 j 类样本的协方差矩阵(严格说是缺少分母部分的协方差矩阵)。

μ_j 的表达式为:

$$\mu_j = \frac{1}{N_j} \sum_{x \in X_j} x \quad (j = -1, 1) \quad (10)$$

\sum_j 的表达式为:

$$\sum_j = \sum_{x \in X_j} (x - \mu_j)(x - \mu_j)^T \quad (j = -1, 1) \quad (11)$$

由于是二分类数据样本, 因此只需要将数据投影到一条直线上即可。假设投影直线是向量 w , 则对任意一个样本 $x^{(i)}$, 它在直线 w 的投影为 $w^T x^{(i)}$, 对于两个类别的中心点 μ_{-1}, μ_1 , 在直线 w 的投影为 $w^T \mu_{-1}$ 和 $w^T \mu_1$ 。由于 LDA 需要让不同类别数据的类别中心之间的距离尽可能大, 也就是说要最大化 $\|w^T \mu_{-1} - w^T \mu_1\|_2^2$, 同时希望同一种类别数据的投影点尽可能接

近,也就是要同类样本投影点的协方差 $w^T \sum_{-1} w$ 和 $w^T \sum_1 w$ 尽可能小,即最小化 $w^T \sum_{-1} w + w^T \sum_1 w$ 。

综上所述,最终的优化目标为:

$$\arg \max J(w) = \frac{\|w^T \mu_{-1} - w^T \mu_1\|_2^2}{w^T \sum_{-1} w + w^T \sum_1 w} = \frac{w^T (\mu_{-1} - \mu_1) (\mu_{-1} - \mu_1)^T w}{w^T (\sum_{-1} + \sum_1) w} \quad (12)$$

一般定义类内散度矩阵 S_w 为:

$$S_w = \sum_{-1} + \sum_1 = \sum_{x \in X_{-1}} (x - \mu_{-1}) (x - \mu_{-1})^T + \sum_{x \in X_1} (x - \mu_1) (x - \mu_1)^T \quad (13)$$

同时定义类间散度矩阵为 S_b :

$$S_b = (\mu_{-1} - \mu_1) (\mu_{-1} - \mu_1)^T \quad (14)$$

然后优化目标重写为:

$$\arg \max J(w) = \frac{w^T S_b w}{w^T S_w w} \quad (15)$$

根据广义瑞利商的性质, $J(w)$ 最大值为矩阵 $S_w^{-1} S_b$ 的最大特征值,而对应的 w 为 $S_w^{-1} S_b$ 的最大特征值对应的特征向量。由于在二分类时, $S_b w$ 的方向恒为 $\mu_{-1} - \mu_1$, 所以不妨设 $S_b w = \lambda (\mu_{-1} - \mu_1)$, 将其带入: $(S_w^{-1} S_b) w = \lambda w$, 可以得到 $w = S_b^{-1} (\mu_{-1} - \mu_1)$, 也就是说只要求出原始二类样本的均值和方差就可以确定最佳的投影方向 w ; 然后对样本集中的每一个样本特征 $x^{(i)}$ 转化为新的样本 $z^{(i)} = w^T x^{(i)}$ ($z^{(i)}$ 表示的是 $x^{(i)}$ 投影到直线上的点到原点的距离)。

1.6 加权投票

加权投票是集成学习的一种方法,是一种很直观的方法,其给分类性能高的分类器赋予一个高的权值,投票结果往往能利用单分类模型间的互补功能来减少单个分类器的误差,提高预测性能和分类精度。其集成四个分类模型的加权投票流程^[23]如下:

(1) 计算每个分类模型的正确率:

$$P_i^s = \frac{m_s}{m_s} \quad (m_s \text{ 为 } R_i^s = y_i \text{ 的累积次数}) \quad (16)$$

其中, $i = 1, 2, 3, 4$; $s = \text{Smile}, \text{Non-Smile}$ 。

(2) 计算每个模型的权值 w_i 为:

$$w_i = \frac{t_i}{\sum_{i=1}^4 t_i} \quad i = 1, 2, 3, 4 \quad (17)$$

$$\text{其中, } t_i = \frac{P_i^{\text{Smile}} + P_i^{\text{Non-Smile}}}{2}。$$

(3) 将各模型得到的检测结果分别与所对应的权值 w_i 相乘,得到 sum_i 。

(4) 将各结果的 sum_i 相加,得到 sum ,然后将 sum 与阈值 0 相比,即根据式 18 得到最终的检测结果。

$$\text{Results} = \begin{cases} \text{Smile} & \text{if } \text{sum} \geq 0 \\ \text{Non-Smile} & \text{otherwise} \end{cases} \quad (18)$$

每个分类模型的权重可以认为是在一个训练子集上的检测率。换句话说,权重可以展现其相应模型的检测的准确度。加权投票步骤如图 1 所示。

2 实验

2.1 实验数据

实验在 GENKI-4K^[24] 人脸数据集上进行,该数据集图片是在不同的成像条件下拍摄的,无论室内还是室外,姿势(大部分的头部的偏转,俯仰和滚动参数图像距离正面位置在 20°以内),年龄,性别,或是否戴帽子或眼镜;而且所有图像都是手动标记。GENKI-4K 分别包含的微笑和非微笑面部图像分别为 2 162 和 1 838 个。文中在数据集中随机选用了 200 张图片进行实验,100 张微笑的,100 张非微笑的。75% 的图片用于训练,其余的用于后期测试验证。在训练和测试阶段,微笑与非微笑照片的比例一样。

2.2 结果分析

从表 1 可得四个分类模型对微笑,非微笑和整体的检测率,以及最终经过加权投票后的检测率。可以看出,表现最好的和最差的分别为模型 F_3 和模型 F_1 。此外,一些模型对微笑的检测率更好,一些对非微笑的检测率更为准确。最后表明不同模型的加权投票可以产生一个完全适用于微笑和非微笑检测的整体模型(见表 1 的最后一列,其中对微笑和非微笑的检测都非常好且非常接近)。对于 GENKI-4K 人脸数据集,基于加权投票的方法的整体识别率为 95.8%。表 2 为文中所用方法与 GENKI-4K 数据集中的其他微笑检测方法进行比较的结果。

表 1 微笑检测率

	%			
	F_1	F_2	F_3	F_4
Smile	84.1	80.6	100	81.2
Non-Smile	73.9	91.4	82	90.8
总体	79	86	91	86
权重	0.231	0.251	0.267	0.251

表 2 比较结果

	Smile	Non-Smile	总体
文献[10]	?	?	94.9%
文献[25]	95.3%	96.3%	95.8%
文中方法	97.2%	94.4%	95.8%

注: ? 表示没有相关的检测率。

3 结束语

应用了一种基于多分类模型加权投票的方法对面

部微笑进行检测识别,这些分类模型分别受现今最流行的人脸识别或性别检测方法的启发;首先对面部图像进行预处理,使之规范化并增强其对比度,然后利用各分类模型优势对图像进行检测分类,最后对各模型的结果进行加权投票。

显而易见,该方法包含了不同的分类检测方法,即神经网络、SVM、AdaBoost 和 LDA,使总体的分类能力明显提高。此外,运用不同的特征提取方法,如 LBP、Gabor、PCA、面部下半部和特征标记检测,把基本能用于微笑检测所需的重要特征都使用了。文中纹理和几何信息均被用于面部微笑检测。第一和第二种模型使用纹理信息,第三和第四种模型使用几何信息。利用这两种类型的信息以及不同的分类模型使得运用的方法具有很强的分类检测能力。

参考文献:

- [1] EKMAN P. Self-deception and detection of misinformation [J]. *Self-deception: An Adaptive Mechanism*, 1988, 16(1): 229–257.
- [2] PANTIC M, PENTLAND A, NIJHOLT A, et al. Human computing and machine understanding of human behavior: a survey [M]// *Artificial intelligence for human computing*. Berlin: Springer, 2007: 47–71.
- [3] POTAPOVA E, EGOROVA M, SAFONOV I. Automatic photo selection for media and entertainment applications [C]// *Proceedings of 19th international conference on computer graphics*. [s.l.]: [s.n.], 2009: 117–124.
- [4] YADAPPANAVAR H, SHYLAJA S S. Machine learning approach for smile detection in real time images [J]. *International Journal of Image Processing and Vision Sciences*, 2012, 1(1): 32–37.
- [5] FREIRE D, SANTANA M C, DÉNIZ-SUÁREZ O. Smile detection using local binary patterns and support vector machines [C]// *Proceedings of the fourth international conference on computer vision theory and applications*. Lisboa, Portugal: [s.n.], 2010: 398–401.
- [6] DÉNIZ O, CASTRILLON M, LORENZO J, et al. Smile detection for user interfaces [C]// *Proceedings of the 4th international symposium on advances in visual computing*. Las Vegas, NV: Springer, 2008: 602–611.
- [7] SHAN C. Smile detection by boosting pixel differences [J]. *IEEE Transactions on Image Processing*, 2012, 21(1): 431–436.
- [8] BITOUK D, KUMAR N, DHILLON S, et al. Face swapping: automatically replacing faces in photographs [J]. *ACM Transactions on Graphics*, 2008, 27(3): 39.
- [9] LECUN Y, BENGIO Y. Convolutional networks for images, speech and time series [M]// *The handbook of brain theory and neural networks*. Cambridge: MIT Press, 1995: 255–258.
- [10] ZHANG Kaihao, HUANG Yongzhen, WU Hong, et al. Facial smile detection based on deep learning features [C]// *3rd IAPR Asian conference on pattern recognition*. Kuala Lumpur, Malaysia: IEEE, 2015: 534–538.
- [11] GONZALEZ R C, WOODS R E. *Digital image processing* [J]. Prentice Hall International, 2005, 28(4): 484–486.
- [12] COOTES T F, TAYLOR C J, COOPER D H, et al. Active shape models – their training and application [J]. *Computer Vision and Image Understanding*, 1995, 61(1): 38–59.
- [13] BALTRUSAITIS T, ROBINSON P, MORENCY L P. Constrained local neural fields for robust facial landmark detection in the wild [C]// *IEEE international conference on computer vision workshops*. Sydney, NSW, Australia: IEEE, 2013: 354–361.
- [14] VIOLA P, JONES M J. Robust real-time face detection [J]. *International Journal of Computer Vision*, 2004, 57(2): 137–154.
- [15] 沈伟听, 王海波, 师小宇. 一种人脸微笑检测方法 [P]. 中国: CN106548160A, 2017-03-29.
- [16] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 971–987.
- [17] PRINCE S J D. *Computer vision: models, learning and inference* [M]. Cambridge: Cambridge University Press, 2012.
- [18] BURGESS C J C. A tutorial on support vector machines for pattern recognition [J]. *Data Mining and Knowledge Discovery*, 1998, 2(2): 121–167.
- [19] BISSOON T, VIRIRI S. Gender classification using face recognition [C]// *International conference on adaptive science and technology*. Pretoria, South Africa: IEEE, 2013: 1–4.
- [20] 于微波, 赵琳, 佟冬. 基于 Adaboost 人脸检测算法 [J]. *吉林大学学报: 信息科学版*, 2014, 32(5): 539–544.
- [21] ISCAN M Y, STEYN M. *The human skeleton in forensic medicine* [M]. [s.l.]: Charles C Thomas Publisher, 2013.
- [22] FRIEDMAN J, HASTIE T, TIBSHIRANI R. *The elements of statistical learning* [M]. New York: Springer Series in Statistics, 2001.
- [23] 周志华. *机器学习* [M]. 北京: 清华大学出版社, 2016: 171–184.
- [24] WHITEHILL J, LITTLEWORT G, FASEL I, et al. Toward practical smile detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(11): 2106–2111.
- [25] GUO Xin, POLANÍA L F, BARNER K E. Smile detection in the wild based on transfer learning [C]// *13th IEEE international conference on automatic face & gesture recognition*. [s.l.]: IEEE, 2018: 3–11.