

基于惩罚因子的PNMS算法的人脸检测和对齐

李振东^{1,2}, 钟 勇^{1,2}, 陈 蔓^{1,2}, 陶 攀^{1,2}

(1.中国科学院 成都计算机应用研究所, 四川 成都 610041; 2.中国科学院大学, 北京 100049)

摘 要:针对人脸相互遮挡、人脸朝向等不确定性因素给人脸检测和对齐任务带来的困难问题, 提出了基于惩罚因子的PNMS算法用以改进人脸检测和对齐的准确性。该算法首先根据人脸候选窗口相互之间的重叠度和候选窗口相应的检测得分, 提出非连续的线性函数和基于高斯分布的连续函数, 作为非极大值抑制算法的两种惩罚因子, 用以改进并替代传统非极大值抑制算法对候选窗口的检测得分进行重分配。在此基础上, 综合前两种惩罚因子的优缺点以及窗口之间的重叠度值, 进一步提出连续非线性函数作为非极大值抑制算法的惩罚因子, 使得窗口之间重叠度值越大则相应的惩罚权重越严重, 且函数在整个重叠取值区间连续。将提出的算法在FDDB和WIDER FACE这2个人脸检测数据集上进行详尽的人脸检测实验验证, 以及在AFLW人脸对齐数据集上进行人脸对齐实验验证。结果表明, 提出的基于惩罚因子的PNMS算法相比于其他算法, 在保持一定实时性的同时不仅有效地提高了人脸检测和对齐的准确率和可靠性, 并且解决了一定程度的人脸相互遮挡被漏检的问题, 降低了被遮挡人脸的漏检率。

关键词:人脸检测; 人脸对齐; 卷积神经网络; 非极大值抑制

中图分类号:TP391.1

文献标志码:A

文章编号:2096-3246(2018)06-0225-07

PNMS Algorithm Based on Penalty Factors for Face Detection and Alignment

LI Zhendong^{1,2}, ZHONG Yong^{1,2}, CHEN Man^{1,2}, TAO Pan^{1,2}

(1.Chengdu Inst. of Computer Application, Chinese Academy of Sciences, Chengdu 610041, China;

2.Univ. of Chinese Academy of Sciences, Beijing 100084, China)

Abstract: In order to solve the difficult problems of caused by mutual occlusion of face and face orientation, a PNMS algorithm based on penalty factors was proposed to improve the accuracy of face detection and alignment. Firstly, according to the overlap degree between face candidate windows and the corresponding detection scores of candidate windows, a non-continuous linear function and a continuous function based on Gaussian distribution were proposed and used as penalty factors for non-maximum suppression. Then the traditional non-maximum suppression algorithm was improved and replaced, and the detection score of the candidate window was redistributed. On this basis, combining the characteristics of the first two kinds of penalty factors, the continuous nonlinear function was further proposed as the penalty factor of the non-maximum suppression algorithm. Consequently, the greater the overlap value between windows, the more severe the penalty is, and the function is continuous throughout the overlapping value range. The proposed algorithm performed detailed face detection experiment verification on two face detection data sets of FDDB and WIDER FACE. The face alignment experiments were verified on the AFLW data set. The results showed that the proposed PNMS algorithms compared with other algorithms not only effectively improves the accuracy and reliability of face detection and alignment, but also solves a certain degree of face occlusion, and reduces the rate of detection failure of occluded faces.

Key words: face detection; face alignment; convolutional neural network; non-maximum suppression

人脸检测和人脸对齐任务是计算机视觉领域的一个基本研究问题。但人脸面部多样性的视觉变化, 如人脸遮挡、夸张表情和光照等因素使得人脸检测

和对齐在现实场景中的应用具有很大的挑战。

Viola等^[1]提出AdaBoost级联分类器并使用简单的人脸Haar-like特征对级联分类器进行训练, 实现了

收稿日期:2017-12-22

基金项目:四川省科技支撑计划资助项目(2014GZ0013)

作者简介:李振东(1990—), 男, 博士生。研究方向:图像处理。E-mail: lizhendong13@mailsucas.ac.cn

网络出版时间:2018-10-23 09:25:00

网络出版地址: <http://kns.cnki.net/kcms/detail/51.1773.TB.20181022.0909.005.html>

<http://jsuese.ijournals.cn>

<http://jsuese.scu.edu.cn>

人脸检测高效实时的性能。然而, Yang等^[2]研究表明, 这种人脸检测器的性能在现实应用中也会随着人脸视觉多样性变化而显著降低。除了级联结构外, Yan等^[3]引入了可变形零件模型(deformable part models, DPM)进行人脸检测, 并实现了卓越的性能。然而, 上述方法计算复杂并且通常在训练阶段需要昂贵的注释。最近, 卷积神经网络(convolutional neural network, CNN)在人脸检测^[4]、人脸对齐^[5]和人脸识别^[6]等计算机视觉任务中取得了显著的进展。Zhang等^[5]提出使用深度卷积神经网络进行人脸对齐任务; Yang等^[7]通过深度卷积神经网络进行人脸属性识别。但是, 由于文献^[5, 7]中深层、复杂的CNN结构使计算复杂度高, 实时性低。Li等^[8]提出使用级联CNN进行人脸检测, 但该算法需要对边界框校准, 增加了额外的计算开销。Zhang等^[9]使用多任务CNN提高多视角人脸检测的准确性, 但检测召回受到弱人脸检测器产生的初始人脸检测窗口的限制。Zhan等^[10]提出使用多任务级联卷积神经网络(multi-task convolutional neural network, MTCNN)学习同时进行人脸检测和对齐任务, 并取得了显著的结果和时效性。

以上算法都将研究注意力集中在提出不同的人脸检测和人脸对齐的算法及不同的模型设计。但是, 如果其中一个简单的模块可以提高人脸检测和对齐的整体性能, 而不再需要再用额外的时间对现有模型进行再训练, 那么它将被广泛地采用。为此, 作者提出基于惩罚因子的惩罚非极大值抑制算法(penalize non-maximum suppression, PNMS), 替代当前人脸检测和对齐任务中传统的非极大值抑制算法^[11](non-

maximum suppression, NMS)。

基于卷积神经网络的人脸检测和对齐算法在网络生成大量的人脸候选窗口和相应的窗口得分后, 使用NMS算法根据窗口之间的交并比重叠度(intersection over union, IoU)去除冗余窗口进行后处理操作, 获得最终的人脸检测窗口。但是, NMS算法的一个主要问题是它将邻近检测窗口的分数根据重叠阈值强制设置为零。因此, 在复杂多目标的环境中当人脸之间存在部分相互遮挡的情况下, 进行人脸检测和对齐任务时相互遮挡人脸之间的候选窗口相互重叠的阈值刚好出现在预设的重叠阈值中, 则被遮挡的人脸检测窗口将被错误的移除, 这将导致被遮挡的人脸漏检, 从而使得人脸检测算法的准确度下降。针对以上问题作者提出基于惩罚因子的PNMS算法, 将检测人脸候选窗口之间的重叠度作为窗口得分函数, 对邻近重叠度大于预设阈值的窗口的分数进行重分配, 得到一个较低的窗口得分, 使得重叠度较高的窗口不会被去除仍保留在候选窗口列表中用于下一步的精确人脸检测和对齐, 以解决被遮挡人脸漏检的问题。

1 人脸检测和对齐算法

1.1 人脸检测和对齐总体框架

多任务人脸检测和对齐算法的总体框架采用现有的MTCNN中提出的级联CNN模型, 如图1所示。模型结构由3个级联的卷积神经网络组成, 分别为人脸候选区域提取网络P-Net、人脸候选窗口细调网络R-Net以及人脸边界框和关键点输出网络O-Net。

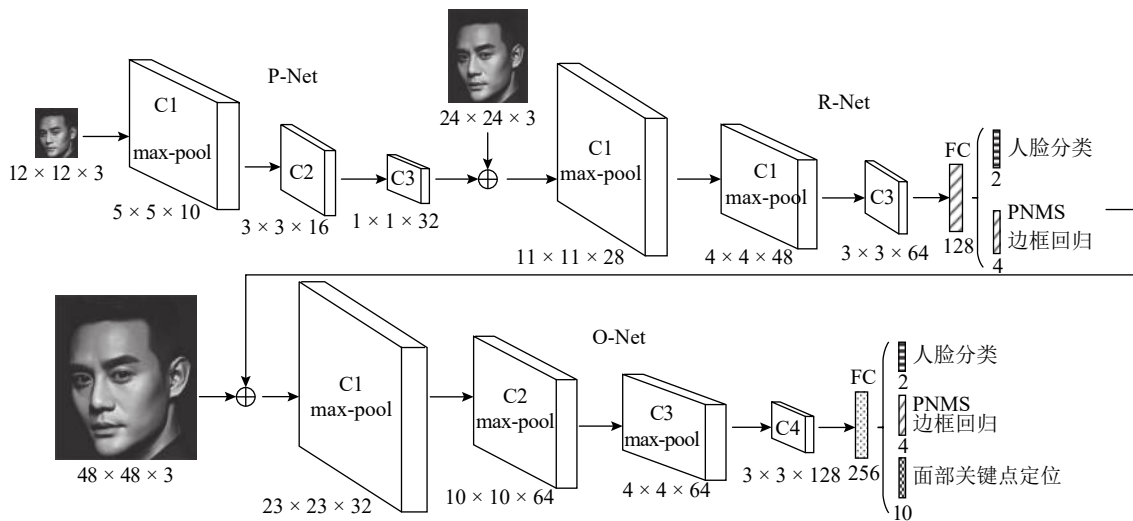


图 1 多任务人脸检测和对齐框架

Fig. 1 Multitask face detection and alignment framework

给定大小为 $h \times w$ 的图像 I_0 , 调整其初始大小建立一个图像金字塔, 作为P-Net的输入; 首先, 通过P-

Net网络输出大量的人脸检测候选窗口 $B = \{B_1, B_2, \dots, B_i, \dots, B_N\}$ 及窗口回归向量 $t = \{t_1, t_2, \dots, t_i, \dots, t_N\}$, 其

中: $\mathbf{B}_i$ 是由人脸候选窗口的4个参数化坐标组成的向量,即 x_1, y_1, x_2, y_2 分别表示人脸候选窗口左上角坐标 (x_1, y_1) 和右下角坐标 (x_2, y_2) ; $\mathbf{t}_i$ 为候选窗口回归向量的4个参数化坐标组成的向量,与 $\mathbf{B}_i$ 类似。然后,基于估计的候选窗口回归向量对人脸候选窗口进行校准并利用NMS算法去除高度重叠的人脸候选窗口;之后,所有人脸候选窗口结合图像 I_0 作为R-Net的输入,通过R-Net输出精确的人脸候选窗口 $\mathbf{B}$, 并利用NMS算法去除高度重叠的人脸候选窗口;最后, R-Net输出的精确人脸候选窗口和图像 I_0 作为O-Net的输入,利用NMS算法得到最终的人脸候选窗口并输出5个人脸面部标志位置,分别为左眼、右眼、鼻子、左嘴角和右嘴角。

1.2 人脸特征提取

图1网络结构中,每个卷积层都利用前一层的输出作为本层的输入,定义为:

$$\mathbf{h}_{w,b}^{m(l)}(x) = \text{ReLU}\left(b^{m(l)} + \sum_n \mathbf{W}^{mn(l)} * \mathbf{h}_{w,b}^{n(l-1)}\right) \quad (1)$$

式中: $\mathbf{h}^{m(l)}$ 和 $\mathbf{h}^{n(l-1)}$ 分别为 l 层的第 m 个输出通道和 $l-1$ 层的第 n 个输入通道; $\mathbf{W}^{mn(l)}$ 为相应的卷积核滤波器; $b^{m(l)}$ 为相应的网络偏置项; $*$ 表示卷积运算符。为了保证所有前景元素的非负性,选择校正线性函数 $\text{ReLU}(x)$ ^[12] 为激活函数:

$$\text{ReLU}(x) = \begin{cases} x, & x > 0; \\ 0, & x \leq 0 \end{cases} \quad (2)$$

由于从卷积层直接提取的图像特征存在特征数据量大、数据维度高问题导致结果容易出现过拟合及计算量大,因此,在P-Net的C1层、R-Net的C1、C2层以及O-Net的C1、C2和C3层采用最大特征值池化方法(max-pooling)对卷积层的不同位置进行特征聚合统计:

$$\mathbf{h}_{(i,j)}^{m(l)} = \max_{(p,q) \in \Omega_{(i,j)}} \left\{ \mathbf{h}_{(p,q)}^{m(l)} \right\} \quad (3)$$

式中, $\Omega_{(i,j)}$ 表示索引为 (i, j) 特定区域, (p, q) 是在 Ω 上具体位置的索引。

为了保证前景(人脸)和背景(非人脸)的有效分类,在卷积层和完全连接层之后(输出层除外)采用 PReLU^[13] 作为激活函数:

$$\text{PReLU}(x) = \begin{cases} x, & x > 0; \\ ax, & x \leq 0 \end{cases} \quad (4)$$

式中, a 初始化为0.25。

1.3 传统NMS算法

在每个网络输出人脸候选窗口集合 $\mathbf{B} = \{\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_i, \dots, \mathbf{B}_N\}$ 和每个窗口对应的人脸得分列表 $\mathbf{s} = \{s_1, s_2, \dots, s_i, \dots, s_N\}$ 后,采用传统NMS算法根据阈值 N_t 对冗余的候选窗口进行筛选删除,算法实现为:

算 法 NMS($\mathbf{B}, \mathbf{s}, N_t$)

输入: 候选窗口 $\mathbf{B}$, 得分 $\mathbf{s}$, 阈值 N_t ;

输出: $\mathbf{B}', \mathbf{s}$;

1. $\mathbf{B}' = \{\}$
2. While $\mathbf{B} \neq \emptyset$ do
3. $s_k \leftarrow \arg \max s$
4. $\mathbf{B}' \leftarrow \mathbf{B}' \cup \mathbf{B}_k; \mathbf{B} \leftarrow \mathbf{B} - \mathbf{B}_k$
5. For $\mathbf{B}_i$ in $\mathbf{B}$ do
6. If $I(\mathbf{B}_k, \mathbf{B}_i) \geq N_t$ then
7. $\mathbf{B} \leftarrow \mathbf{B} - \mathbf{B}_i; s \leftarrow s - s_i$
8. End
9. End
10. End
11. Return $\mathbf{B}', \mathbf{s}$

从NMS算法实现过程可知, NMS算法设定一个硬阈值 N_t 以决定是否保留或者移除与当前得分最高的候选窗口 $\mathbf{B}_k$ 邻近的所有窗口 $\mathbf{B}_i$, 即根据 $\mathbf{B}_k$ 与 $\mathbf{B}_i$ 之间的IoU值 $I(\mathbf{B}_k, \mathbf{B}_i)$ 对满足阈值条件的邻近窗口 $\mathbf{B}_i$ 及相应的得分 s_i 保留不变, 而移除不满足条件的邻近窗口 $\mathbf{B}_i$ 并将相应的得分 s_i 置为0, 该过程实质是根据设定的阈值 N_t 对邻近候选窗口进行重置得分的过程, 表示如下:

$$s_i = \begin{cases} s_i, & I(\mathbf{B}_k, \mathbf{B}_i) < N_t; \\ 0, & I(\mathbf{B}_k, \mathbf{B}_i) \geq N_t \end{cases} \quad (5)$$

式中, $\mathbf{B}_k$ 与 $\mathbf{B}_i$ 之间的IoU值 $I(\mathbf{B}_k, \mathbf{B}_i)$ 的定义如下:

$$I(\mathbf{B}_k, \mathbf{B}_i) = \frac{\mathbf{B}_k \cap \mathbf{B}_i}{\mathbf{B}_k \cup \mathbf{B}_i} \quad (6)$$

这样, 经过NMS算法重置得分为0的冗余候选窗口就会被移除, 只保留了少量得分较高且相互之间重叠度较低的人脸候选窗口。但是, NMS算法的一个主要问题是它将相互遮挡的邻近检测窗口的分数根据重叠阈值强制设置为零从而使被遮挡人脸出现漏检问题, 因此作者提出PNMS算法以解决被遮挡人脸的漏检问题。

2 PNMS算法

当候选窗口 $\mathbf{B}_i$ 与最大得分窗口 $\mathbf{B}_k$ 有非常高的重叠时将重新分配一个很低的窗口得分; 反之, 当重叠度较小时则保持 $\mathbf{B}_i$ 原来的检测分数 s_i 不变。因此, PNMS将检测窗口相互之间的重叠度作为降低窗口分数的函数, 以替代NMS算法中将窗口分数直接置为0而被移除, 具体定义如下

$$s_i = \begin{cases} s_i, & I(\mathbf{B}_k, \mathbf{B}_i) < N_t; \\ s_i f(I(\mathbf{B}_k, \mathbf{B}_i)), & I(\mathbf{B}_k, \mathbf{B}_i) \geq N_t \end{cases} \quad (7)$$

式中, $I(\mathbf{B}_k, \mathbf{B}_i)$ 定义如式(6)所示, 这样被遮挡对象的检测窗口的分数经惩罚函数 $f(I(\mathbf{B}_k, \mathbf{B}_i))$ 后会降低, 但仍保留在候选窗口列表中。

2.1 线性惩罚函数

根据以上分析可知,惩罚与当前最大得分窗口 B_k 重叠的窗口 B_i 的得分 s_i , 使得 s_i 降低, 这是一个有效改进NMS算法的途径。而且对于重叠度比较高的窗口 B_i 的得分 s_i 应该受到严重的惩罚, 因为这样的窗口有更大可能性是假阳性样本, 因此作者提出线性惩罚函数作为PNMS算法的惩罚因子(PNMS-linear):

$$s_i = \begin{cases} s_i, I(B_k, B_i) < N_t; \\ s_i(1 - I(B_k, B_i)), I(B_k, B_i) \geq N_t \end{cases} \quad (8)$$

上述函数将与窗口 B_k 重叠度在检测阈值 N_t 以上的窗口检测分数进行衰减的线性函数。因此, 远离 B_k 或重叠度较小的检测窗口不会受到影响, 而那些非常接近的检测窗口将被分配更大的惩罚。

2.2 基于高斯分布的惩罚函数

然而, 在窗口重叠方面线性惩罚函数是不连续的, 并且当窗口之间的重叠度达到PNMS阈值 N_t 时突然惩罚, 这样会导致窗口列表中的突变。理想的惩罚函数应该是连续的, 随着窗口之间重叠度的增加而相应的惩罚逐渐增加。因此, 当重叠较低时, B_k 对其重叠的窗口的分数影响很低, 然而, 当一个窗口 B_i 与 B_k 的重叠度 $I(B_k, B_i)$ 接近1时, B_i 的得分 s_i 应该受到显著的惩罚。考虑到这一点, 作者结合高斯函数提出连续的高斯惩罚函数作为PNMS算法的惩罚因子(PNMS-Gaussian):

$$s_i = s_i e^{-\frac{I^2(B_k, B_i)}{\sigma}}, \forall B_i \notin B' \quad (9)$$

式中, 根据高斯函数的分布, σ 取值为0.5。每次迭代时应用高斯连续惩罚函数更新规则, 更新所有剩余检测框的分数。

2.3 非线性连续惩罚函数

因为窗口 B_k 与附近任意窗口 B_i 的重叠度的取值范围为[0, 1], 因此PNMS-linear和PNMS-Gaussian在区间[0, 1]上的函数图像如图2中带“*”和“o”标记的曲线所示。图2中, 横坐标表示窗口之间的重叠度, 纵坐标表示对应的惩罚权重比例。

由图2所示, PNMS-linear中当窗口之间的重叠度大于NMS阈值 N_t 时突然对得分进行惩罚, 并且惩罚力度一样; 而PNMS-Gaussian对任意的窗口 B_i 的惩罚权重会随着窗口之间的重叠度增加而增大, 并且比PNMS-linear更平滑。但是, 由图2中PNMS-Gaussian图像曲线的斜率可知, 当重叠度取值范围较大(例如在[0.7, 1]中)时其惩罚力度会相对减弱, 并未达到最理想的状态。

针对PNMS-linear和PNMS-Gaussian的不足, 作者提出理想的连续非线性惩罚函数作为PNMS算法的惩罚因子(PNMS-nonlinear):

$$s_i = s_i (1 - I^3(B_k, B_i)) \quad (10)$$

PNMS-nonlinear在区间[0, 1]上的函数图像如图2中带“+”标记的曲线所示。随着窗口之间重叠度的增加, 相应的惩罚权重会越严重, 且函数在整个重叠取值区间连续, 惩罚权重更加平滑合理。

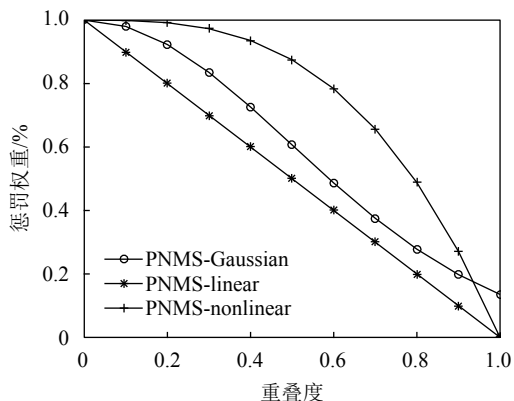


图2 3种不同惩罚函数的函数图像

Fig. 2 Function image of three different penalty functions

3 实验结果与分析

3.1 实验数据

将提出的基于不同惩罚因子的PNMS算法(PNMS-linear、PNMS-Gaussian、PNMS-nonlinear)与其他现有的人脸检测和对齐算法在人脸检测数据集FDDB^[14]、WIDER FACE^[15]和人脸对齐数据集AFLW^[16]上进行评估比较。FDDB数据集包含2 845幅图像, 共标注5 171个人脸; WIDER FACE数据集由32 203幅图像中的393 703个标记的人脸边界框组成; AFLW包含24 386个人脸的面部标志注释, 作者使用与文献[5]相同的测试子集。

3.2 PNMS算法检测时效对比

图3为在FDDB人脸检测数据集上基于3种惩罚因子的PNMS算法以及MTCNN进行人脸检测时所用时间的对比。

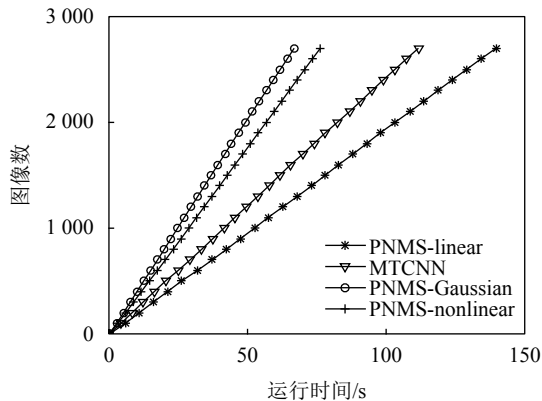


图3 不同算法在FDDB上检测时间对比

Fig. 3 Comparison of detection time between different algorithms on FDDB

从图3中可以看出: PNMS-Gaussian算法所用的时间最短, 共用时70 s, 平均每幅图像处理时间为0.024 s, 达到实时处理速度, 具有明显的优势; PNMS-Nonlinear算法用时次之但优于MTCNN, 平均处理每幅图像用时0.028 s; 而PNMS-linear算法检测每幅图像所用的时间最长, 平均处理每幅图像用时0.051 s。

3.3 评估人脸检测

为了评估提出的人脸检测和对齐算法的性能, 将提出的与其他主流算法在FDDB数据集上进行人脸检测性能比较; 同时, 在WIDER FACE数据集上进行人脸检测性能比较。

图4为提出的PNMS-linear、PNMS-Gaussian、PNMS-nonlinear算法与ACF^[2]、Faceness^[7]、Cascade CNN^[8]、MTCNN^[10]在FDDB人脸检测数据集上进行人脸检测的评估结果对比。从图4中可以看出, 在FDDB人脸检测数据集上PNMS-linear算法和PNMS-nonlinear算法表现出最优的检测性能, 人脸检测平均准确率分别达到93.1%和92.9%。

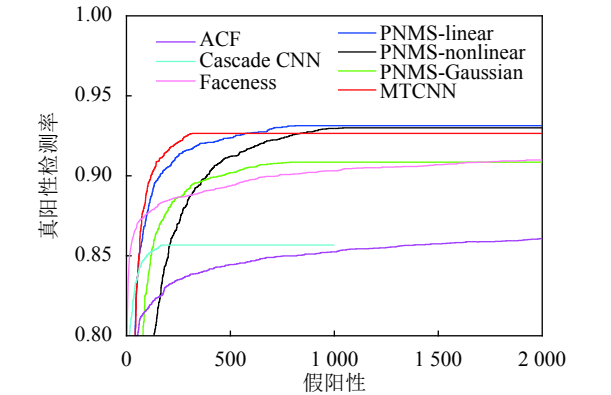


图 4 不同算法在FDDB上的人脸检测评估对比
Fig. 4 Comparison of face detection evaluation of different algorithms on FDDB

表1为提出的PNMS-linear、PNMS-Gaussian、PNMS-nonlinear算法与ACF^[2]、Faceness^[7]、Cascade CNN^[8]、MTCNN^[10]、Two-stage CNN^[15]在WIDER FACE人脸检测数据集的3个子集上的验证评估结果对比。表1中, Easy set、Medium set和Hard set是根据人脸图像复杂度划分的3个WIDER FACE部分人脸图像子集。从表1可以看出: 在WIDER FACE数据集的3个子集上PNMS-linear和PNMS-nonlinear算法表现出最优的人脸检测性能, 这两种算法在Easy set子集上准确率分别达到74.6%和74.4%; 而PNMS-Gaussian在3个子集上的检测性能仅次于MTCNN, 与在FDDB人脸检测数据集上的评估结果一致。

3.4 评估人脸对齐

图5为提出的PNMS-linear、PNMS-Gaussian、PNMS-nonlinear算法与TDCN^[5]、MTCNN^[10]、RCPR^[17]、

表 1 不同算法在WIDER FACE上的人脸检测评估对比
Tab. 1 Comparison of face detection evaluation of different algorithms on WIDER FACE

算法	检测平均准确率/%		
	Easy set	Medium set	Hard set
ACF ^[2]	65.9	54.1	27.3
Faceness ^[7]	71.3	63.4	34.5
Cascade CNN ^[8]	69.1	66.4	42.4
MTCNN ^[10]	71.6	68.9	45.7
Two-stage CNN ^[15]	68.1	61.8	32.3
PNMS-Gaussian	70.7	68.1	45.2
PNMS-nonlinear	74.4	70.8	46.5
PNMS-linear	74.6	71.5	47.3

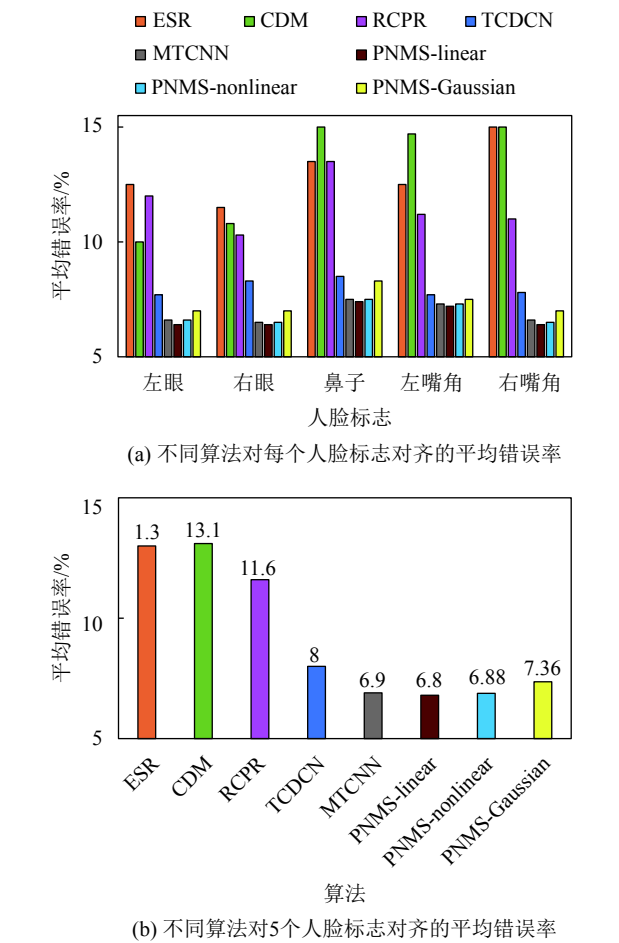


图 5 在AFLW数据集上人脸关键点对齐评估
Fig. 5 Evaluation face key point alignment on AFLW dataset

ESR^[18]和CDM^[19]等人脸对齐算法在AFLW人脸对齐数据集上进行性能评估比较。

图5中, 平均误差是通过估计各个人脸面部标志(左眼、右眼、鼻子、左嘴角和右嘴角)与人脸标注的真实值之间的距离进行测量的。图5(a)为各个算法对每个人脸面部标志对齐的平均错误率, 图5(b)为

各个算法对5个人脸面部标志对齐的平均错误率。从图5中结果可以看出：PNMS-linear算法在人脸对齐上的平均错误率最低，性能最优；PNMS-nonlinear算法性能次之，也具有一定的优势；而PNMS-Gaussian与前两种算法相比平均错误率较高，这与在FDDDB和WIDER FACE人脸检测数据集上的检测性能评估结果一致。

3.5 算法运行时间对比

为验证提出的基于惩罚因子的PNMS级联人脸检测和人脸对齐算法能否达到实时的时效性。将提出的PNMS-linear、PNMS-Gaussian、PNMS-nonlinear算法与人脸检测和人脸对齐算法Faceness^[7]、Cascade CNN^[8]、MTCNN^[10]的运行时效进行比较，结果如表2所示。

从表2中可以看出：PNMS-nonlinear和PNMS-Gaussian时效性最优，可以达到实时性；而PNMS-linear在保持较高的人脸检测和对齐准确率的同时速

度依然可以达到40 fps，速度略低于MTCNN，但仍然可以满足实时检测速度。

表 2 不同算法人脸检测平均运行时效比较

Tab. 2 Comparison of average running time of face detection by different algorithms

算法	GPU	速度/fps
PNMS-linear	GeForce GTX TITAN X	40
PNMS-nonlinear	GeForce GTX TITAN X	71
PNMS-Gaussian	GeForce GTX TITAN X	83
MTCNN ^[10]	GeForce GTX TITAN X	50
Cascade CNN ^[8]	Nvidia Titan Black	100
Faceness ^[7]	Nvidia Titan Black	20

3.6 部分遮挡人脸的检测结果对比

图6为提出的基于3种惩罚函数的PNMS算法以及MTCNN在部分遮挡人脸图像上的检测结果对比。



图 6 不同算法对部分遮挡人脸图像的检测结果对比

Fig. 6 Comparison of detection results of partial occlusion face images by different algorithms

图6中,绿色边界框表示检测到的人脸边界框,绿色的点表示左眼、右眼、鼻子、左嘴角和右嘴角人脸5个关键面部标志位置。从图6的检测结果可知,提出的PNMS-linear、PNMS-Gaussian以及PNMS-nonlinear算法相对于MTCNN可以更加有效地解决一定程度上的人脸相互遮挡检测问题,避免了一定程度上被遮挡人脸的漏检。

4 结 论

提出了基于3种惩罚函数的PNMS算法用于人脸检测和对齐的优化和改进。根据人脸候选窗口相互之间的重叠度和检测得分提出3种关联函数作为PNMS算法的惩罚因子,用以改进优化并替代传统的NMS算法。实验结果表明,提出的算法能够有效地解决一定程度的人脸之间相互遮挡导致被遮挡的人脸漏检的问题,有助于提高人脸检测性能,在多个人脸检测和人脸对齐验证基准测试集(包括人脸检测FDDB数据集和WIDER FACE数据集以及人脸对齐AFLW数据集)上与现有的其他算法相比具有一定的优势,在提高人脸检测准确率性能的同时保证了检测的实时性能。需要指出的是所提算法是基于图像的人脸检测和对齐,因此,下一步的研究方向是进行人脸活体检测和人脸识别,以提高人脸认证等应用的安全性。

参考文献:

- [1] Viola P, Jones M J. Robust real-time face detection[J]. *International Journal of Computer Vision*, 2004, 57(2): 137–154.
- [2] Yang Bin, Yan Junjie, Lei Zhen, et al. Aggregate channel features for multi-view face detection[C]//Proceedings of the IEEE International Joint Conference on Biometrics. *Clearwater: IEEE*, 2014: 1–8.
- [3] Yan Junjie, Lei Zhen, Wen Longyin, et al. The fastest deformable part model for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. *Columbus: IEEE*, 2014: 2497–2504.
- [4] Wu Qingjia. Rotation invariant face detection system based on neural network ensemble[J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2013, 43(Supp1): 433–438. [吴清佳. 基于神经网络集成的旋转人脸快速检测系统[J]. *吉林大学学报(工学版)*, 2013, 43(增刊1): 433–438.]
- [5] Zhang Zhanpeng, Luo Ping, Loy C C, et al. Facial landmark detection by deep multi-task learning[M]//Computer Vision-ECCV 2014. *Cham: Springer*, 2014: 94–108.
- [6] Sun Yi, Chen Yuheng, Wang Xiaogang, et al. Deep learning face representation by joint identification-verification[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. *Cambridge: MIT Press*,

2014: 1988–1996.

- [7] Yang Shuo, Luo Ping, Loy C C, et al. From facial parts responses to face detection: A deep learning approach[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. *Santiago: IEEE*, 2015: 3676–3684.
- [8] Li Haoxiang, Lin Zhe, Shen Xiaohui, et al. A convolutional neural network cascade for face detection[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. *Boston: IEEE*, 2015: 5325–5334.
- [9] Zhang Cha, Zhang Zhengyou. Improving multiview face detection with multi-task deep convolutional neural networks[C]//Proceedings of the 2014 IEEE Winter Conference on Applications of Computer Vision (WACV). *Steamboat: IEEE*, 2014: 1036–1041.
- [10] Zhang Kaipeng, Zhang Zhanpeng, Li Zhifeng, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. *IEEE Signal Processing Letters*, 2016, 23(10): 1499–1503.
- [11] Neubeck A, Gool L V. Efficient non-maximum suppression[C]//Proceedings of the 18th International Conference on Pattern Recognition. *Hong Kong: IEEE*, 2006: 850–855.
- [12] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems. *New York: ACM*, 2012: 1097–1105.
- [13] He Kaiming, Zhang Xiangyu, Ren Shaoqing, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. *Santiago: IEEE*, 2015: 1026–1034.
- [14] Jain V, Learned-Miller E. Fddb: A benchmark for face detection in unconstrained settings[R]. Amherst: University of Massachusetts Amherst, 2010.
- [15] Yang Shuo, Luo Ping, Loy C C, et al. Wider face: A face detection benchmark[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. *Las Vegas: IEEE*, 2016: 5525–5533.
- [16] Köstinger M, Wohlhart P, Roth P M, et al. Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization[C]//Proceedings of the 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops). *Barcelona: IEEE*, 2011: 2144–2151.
- [17] Burgos-Artizzu X P, Perona P, Dollár P. Robust face landmark estimation under occlusion[C]//Proceedings of the 2013 IEEE International Conference on Computer Vision. *Sydney: IEEE*, 2013: 1513–1520.
- [18] Cao Xudong, Wei Yichen, Wen Fang, et al. Face alignment by explicit shape regression[J]. *International Journal of Computer Vision*, 2014, 107(2): 177–190.
- [19] Yu Xiang, Huang Junzhou, Zhang Shaoting, et al. Pose-free facial landmark fitting via optimized part mixtures and cascaded deformable shape model[C]//Proceedings of the 2013 IEEE International Conference on Computer Vision. *Sydney: IEEE*, 2013: 1944–1951.

(编辑 赵 婧)

引用格式: Li Zhendong, Zhong Yong, Chen Man, et al. PNMS algorithm based on penalty factors for face detection and alignment[J]. *Advanced Engineering Sciences*, 2018, 50(6): 225–231. [李振东, 钟勇, 陈蔓, 等. 基于惩罚因子的PNMS算法的人脸检测和对齐[J]. *工程科学与技术*, 2018, 50(6): 225–231.]