研究与开发

一种复杂场景下的视频流人脸隐私保护技术

张驰，陆晔，罗渝平，孙晓凯，祝涵珂
（中国电信股份有限公司上海分公司，上海 200122）

摘　要：查看视频监控的过程中，一些场景存在因为人脸面部信息暴露在监控视频中导致个人隐私信息泄露的风险，有必要对实时视频流中的行人隐私信息进行马赛克处理。目前市面上常见的基于人脸检测的打码方法在实时监控视频流上打码效果受行人姿态、光线影响较大，存在实时性差、漏检较多等问题。针对以上问题，提出了融合人脸检测算法、目标物体检测算法和前置帧关联检测方法的多检测模型，并与传统的人脸检测模型进行对比。实验结果表明，在人脸检测召回率上，所提模型相较于传统人脸检测算法提高了 532%。

关键词：人脸隐私保护；人脸打码；物体检测
中图分类号：TP311
文献标识码：A
doi: 10.11959/j.issn.1000−0801.2021015

A novel approach for face privacy protection based on surveillance video in complex scene

ZHANG Chi, LU Ye, LUO Yuping, SUN Xiaokai, ZHU Hanke
Shanghai Branch of China Telecom Co., Ltd., Shanghai 200122, China

Abstract: During the process of monitoring video surveillance, potential risks of leaking personal privacy information may occur due to the exposure of facial information to the surveillance video. It is necessary to mosaic the pedestrian privacy information in the real-time video stream. At present, the commonly used coding method based on face detection is heavily affected by pedestrian appearances and illumination, which often leads to weak real-time performance and undetected errors. To solve the problems above, a multi-detection model was proposed, combining face detection algorithm, target object detection algorithm and pre-frame association detection method, and compared with the traditional face detection model. Experimental results show that the recall rate of the proposed model is 532% higher than that of the traditional face detection algorithm.

Key words: face privacy protection, face mosaic, object detection

1　引言

当前，社会治安形势日趋复杂，传统的治安防控措施已经难以满足现实需求。中央已将公共安全视频监控系统建设纳入“十三五”规划和国家安全保障能力建设规划，开展“雪亮工程”建

收稿日期：2020−08−19；修回日期：2020−10−10

设。在视频监控广泛覆盖的背景之下，随着电信转型业务的发展，在智慧园区、楼宇等产品中，越来越多的产品嵌入园区、楼宇等内部视频流，助力用户第一时间发现问题，实时掌握现场情况。但许多监控视频中有大量人脸信息，存在生物识别 ID 乃至关联的身份、位置等个人信息暴露的可能，不仅存在一定的法律风险[1]，也存在被不法分子利用的可能性。因此，在非必要场合，应对实时监控视频中的人脸使用马赛克进行处理，规避相关的风险。

传统的视频流人脸隐私保护技术一般分为 3 个步骤：

（1）将视频流分解为图片序列；

（2）对图片序列中每一帧图片进行人脸检测并使用马赛克替换；

（3）将处理过的图片序列重新编码为指定格式的视频流。

通常使用例如 OpenCV[2]等开源工具箱从视频源读取一帧图片，然后通过人脸检测算法对图片进行检测，对检测到的人脸进行处理，输出分组视频流然后再读取下一帧图片，以此往复。这种做法对人脸检测和处理的性能提出比较高的要求，处理一帧图片的时间上限取决于视频流的每秒帧数，处理时间过长会导致视频严重卡顿。在许多场景下比如人脸较多时处理耗时增加不可避免，这便对人脸检测算法模型的性能提出了比较高的要求，限制了复杂而精确模型的发挥，导致技术进步困难，算法精度存在瓶颈。另外在网络条件相对不稳定出现轻微波动时，该方法会因为无法及时读取视频流而导致输出视频流的卡顿，对网络波动抵抗力较差。

传统的视频流人脸隐私保护技术[3]中人脸检测技术主要分为两种：手动检测和半自动检测。手动检测一般通过一定技术手段将视频流拆解为图片序列，然后对于每一张图片，通过人工对其中的人脸使用马赛克算法进行覆盖，然后再将图片序列整合为视频。该方式人脸检测召回率极高，打码位置一般也较为精确，但基本无法做到实时打码，人工成本高，不能满足对实时性存在要求的业务需求。半自动打码通常使用传统机器学习的方法定位人脸，通过马赛克算法、差分算法等对检测到的人脸进行处理[4-5]。传统的人脸检测算法通常基于模板匹配法或使用 Haar-like[6]等传统机器学习特征结合分类器 Ada-boost[7]或者 SVM[8]等分类器对视频中的人脸进行检测和定位。该类算法优点是检测速度较快、开销较小、检测效率较高，但对人脸的位置、角度要求极高，受亮度、光线影响较大，在例如监控场景等非配合式场景下效果不佳。然而在人脸隐私保护应用中，如果存在较多的漏打，会直接导致部分人脸直接暴露在公众视野之下，人脸隐私保护的效果便荡然无存。换而言之，人脸隐私保护应用对人脸检测算法的要求更偏重于召回率，要尽量减少漏打现象的出现，对于把非人脸识别为人脸的误打现象存在一定的包容性。所以传统人脸检测算法只能做到替代部分人工，仍需要对整个视频进行人工复核。这一点直接导致传统人脸检测算法不能胜任人脸隐私保护应用中人脸检测的工作。

目前主流的人脸检测算法主要研究推进方向多为提高算法的准确率，对于召回率的提高一般放在次要位置，本文尝试使用目前先进的基于深度学习框架的人脸检测算法——多任务卷积神经网络（multi-task convolutional neural network，MTCNN）[9]并进行了测试，虽然在正脸检测上取得了良好的效果，但在侧脸的检测上效果不佳，当行人脸部位置、姿态发生变化的时候检测不稳定，漏检率很高，在对人脸检测算法召回率要求较高的复杂场景下，隐私保护效果大打折扣。针对以上问题，本文通过将检测任务进行拆分，提出了一种融合人脸检测算法 MTCNN 和物体检测算法 YOLOv3[10]的融合检测模型来提高复杂场景

下的人脸检测召回率，其中 MTCNN 用于对人脸正脸进行检测，自训练的 YOLOv3 模型用于对人脸侧脸、背对镜头等姿态进行检测，极大地提高了人脸检测召回率。并针对视频流检测任务，利用视频流帧间关联特性，通过前置帧关联检测方法进一步降低人脸检测不稳定的现象，提高了算法的稳健性。

2 基于深度学习的人脸检测算法

多任务卷积神经网络是目前主流的基于深度学习的人脸检测算法，相对于传统人脸检测方法有着巨大的技术进步，效果也有着巨大的提升。它将人脸区域检测与人脸关键点检测放在一起，总体可分为 P-Net、R-Net 和 O-Net 三层网络结构。它是2016年中国科学院深圳光通技术研究院提出的用于人脸检测任务的多任务神经网络模型[9]，该模型主要采用了 3 个级联的网络，采用候选框架分类器的思想，进行快速高效的人脸检测。这 3 个级联的网络分别是快速生成候选窗口的 P-Net、进行高精度候选窗口过滤选择的 R-Net 和生成最终边界框与人脸关键点的 O-Net。与很多处理图像问题的卷积神经网络模型相同，该模型也用到了图像金字塔、边框回归、非最大值抑制等技术。

首先将图像进行不同尺度的变换，构建图像金字塔，以适应不同大小的人脸进行检测。

P-Net 全称为 Proposal Network，其基本的构造如图 1 所示。对上一步构建完成的图像金字塔，通过卷积神经网络进行初步特征提取与标定边框，并通过边框回归（bounding-box regression）调整候选框定位，通过非极大值抑制进行大部分候选框的过滤。P-Net 将图像输入 3 个卷积层之后，通过一个人脸分类器判断该区域是否是人脸，同时使用边框回归和一个面部关键点的定位器来得到人脸区域的初步定位，该部分最终将输出很多张可能存在人脸的候选框，并输入 R-Net 进行进一步处理。

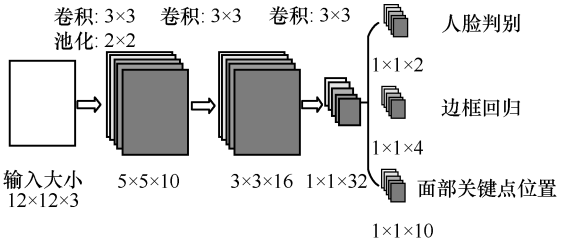

图 1 P-Net 的基本构造

R-Net 全称为 Refine Network，其基本的构造是一个卷积神经网络，如图 2 所示。与第一层的 P-Net 相比，增加了一个全连接层，因此对于输入数据的筛选会更加严格。R-Net 会滤除大量效果比较差的候选框，最后对选定的候选框通过边框回归和非极大值抑制进一步优化预测结果。

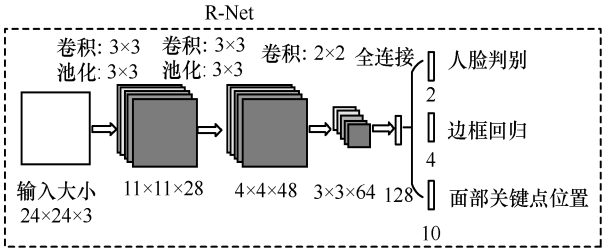

图 2 R-Net 架构

因为 P-Net 的输出只是具有一定可信度的可能的人脸区域，在这个网络中，将对输入进行细化选择，且舍去大部分的错误输入，并再次使用边框回归和面部关键点定位器进行人脸区域的边框回归和关键点定位，最后将输出较为可信的人脸区域，供 O-Net 使用。与 P-Net 使用全卷积输出的 1×1×32 的特征对比，R-Net 在最后一个卷积层之后使用了一个包含 128 个节点的全连接层，保留了更多的图像特征，准确率也优于 P-Net。

O-Net 全称为 Output Network，基本构造是一个较为复杂的卷积神经网络，相对于 R-Net 来说多了一个卷积层，O-Net 架构如图 3 所示。O-Net 的效果与 R-Net 的区别在于这一层结构会通过更多的监督来识别面部的区域，而且会对人的面部特征点进行回归，最终输出 5 个人脸面部特征点。

O-Net 是一个更复杂的卷积网络，该网络的输

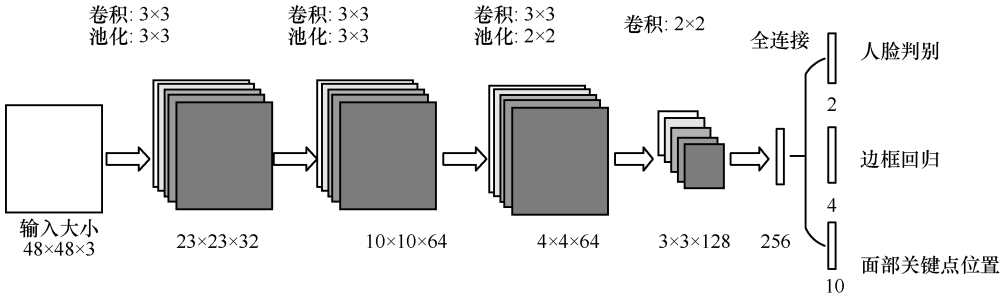

图 3　O-Net 的基本结构

入特征更多，在网络结构的最后是一个更大的具有 256 个节点的全连接层，从而保留了更多的图像特征。之后再进行人脸判别、边框回归和面部关键点定位，得到人脸区域的左上角坐标和右下角坐标与人脸区域的 5 个特征点，作为网络模型的最终输出。

3　基于目标检测的非正脸检测算法

在复杂场景下，行人的人脸姿态和角度存在着非常多的变化，人脸在监控视频中经常只有侧脸，或者可能只能拍到行人的头顶。目前主流的人脸检测算法 MTCNN 对于人脸的仰角俯角、左右旋转角度都有一定的限制要求，非正脸损失了太多的人脸信息导致人脸特征点无法计算，目前来看通过对人脸检测算法进行改进提升对非正脸的检测召回率相对比较困难，是业界目前尚未完全解决的难题，所以本文认为应将非正脸的检测作为与正脸检测截然不同的问题进行处理。考虑到行人在视频流中的位置、姿态、角度多种多样，为确保人脸隐私保护的召回率，侧脸、头顶、后脑等各个非正脸的角度都有检测的必要，即对人头的各个角度进行检测，非正脸检测问题便归结为了对人头的检测问题。人头作为一个形态特征比较明确的物体，可以通过目标检测算法解决人头检测问题。

基于以上考虑，本文将人脸检测模型拆分为正脸检测和非正脸检测两个部分。对于非正脸检测，使用兼顾性能和效率的目标检测模型 YOLOv3 训练人头模型用于检测视频流中的非正脸。虽然该人头检测模型通过训练除了非正脸以外也可以对正脸进行检测，但考虑人脸检测算法 MTCNN 在正脸检测上无论是准确率、召回率还是性能均全面优于自训练模型，最终采取算法实现方案为使用人脸检测算法 MTCNN 对正脸进行检测，使用自训练目标检测模型 YOLOv3 对非侧脸进行检测，二者检测结果的并集作为当前帧的人脸检测结果。

YOLOv3 是现在常用的一种目标物体检测算法。在基本的图像特征提取方面，YOLOv3 采用了称之为 Darknet-53 的网络结构（含有 53 个卷积层），它借鉴了残差网络的做法，在一些层之间设置了快捷链路。

YOLOv3 采用了 3 个不同尺度的特征图来进行对象检测，可以有效提升复杂场景下各种大小物体检测的准确率。卷积网络在 79 层后，会通过多个不同尺度的卷积层得到多种检测结果，卷积层的大小为原始输入图像的 1/1 024，即如果原始输入大小是 416×416，这里的特征图就是 13×13。由于下采样倍数高，这里特征图的感受野（receptive field）比较大，因此适合检测图像中尺寸比较大的对象。

为了实现细粒度的检测，第 79 层的特征图又开始作上采样（从 79 层往右开始上采样卷积），然后与第 61 层特征图融合，这样得到第 91 层较细粒度的特征图，同样经过几个卷积层后得到相对输入图像 16 倍下采样的特征图。它具有中等尺度的感受野，适合检测中等尺度的对象。

最后，第 91 层特征图再次上采样，并与第 36 层特征图融合，最后得到相对输入图像 8 倍下采样的特征图。它的感受野最小，适合检测小尺寸的对象。

随着输出的特征图的数量和尺度的变化，先验框的尺寸也需要相应的调整。YOLOv3 采用 *K*-means 聚类得到先验框的尺寸，为每种下采样尺度设定 3 种先验框，总共聚类出 9 种尺寸的先验框。在 COCO 数据集[11]实验这 9 个先验框分别是：（10×13）、（16×30）、（33×23）、（30×61）、（62×45）、（59×119）、（116×90）、（156×198）、（373×326）。

在最小的 13×13 特征图上（有最大的感受野）应用较大的先验框（116×90）、（156×198）、（373×326），适合检测较大的对象。中等的 26×26 特征图上（中等感受野）应用中等的先验框（30×61）、（62×45）、（59×119），适合检测中等大小的对象。较大的 52×52 特征图上（较小的感受野）应用较小的先验框（10×13）、（16×30）、（33×23），适合检测较小的对象。特征图、感受野与先验框对应情况见表 1。

YOLOv3 借鉴了残差网络结构，形成更深的网络层次以及多尺度检测，提升了 mAP 及小物体检测效果。如果采用 COCO 数据集上 50%重叠度下平均准确率（COCO mAP50）做评估指标，在精确度相当的情况下，YOLOv3 的速度是其他模型的 3~4 倍。

4 前置帧关联检测方法

虽然在融合了人脸检测和人头检测算法以后，本文的方法在图片数据集上取得了良好的检测效果。但在实际复杂场景的视频流中，行人的移动、姿态变化、光线变化等各种不可控因素，导致算法模型人脸检测的稳健性依然不高，在对连续帧的检测中依然会存在某些帧中的某些人脸的漏检，反映在输出视频流上是人脸上的马赛克时断时续，隐私保护效果依然不够理想。

针对这个问题，利用视频流时序信息，提出了一种通过前置帧进行关联检测的改进方法。一般而言，本文可以假设行人不会从视频流中凭空出现或者消失，而且行人的移动速度并不会太快。所以对于一个行人而言，他在视频流中的移动速度将和他的人脸隐私保护马赛克一同连续运动，在相邻几帧这种很短的时间内，位置并不应发生较大位移或者凭空消失。如果某一帧中某个人脸马赛克并非在屏幕边缘但突然消失了，大概率是这一帧发生了漏检，这时可以使用几帧以内上一次人脸马赛克出现的位置作为本帧应打马赛克位置的预测。在本文的测试中，设置阈值为 10 帧以内。融合人脸检测模型、人头检测模型和前置帧关联检测方法的多检测模型流程如图 4 所示。

整个人脸检测流程将分两个线程进行。

- 线程一负责对图片帧进行人脸检测定位并进行马赛克处理，然后将处理后的图片帧输出到检测结果资源池中进行缓存；
- 线程二负责根据时序和预设的帧率等参数从资源池中抽取图片帧进行输出。

双线程架构形成了内部缓冲架构，当线程一因为间歇性输入视频流卡顿或者部分图片帧由于人脸较多等原因检测较慢的时候，可以维持线程二图片帧输出的稳定，保障视频流的平稳输出。

在线程一中，本文首先使用自训练物体检测

表 1 特征图、感受野与先验框对应的情况

对比项	特征图		
	13×13	26×26	52×52
感受野	大	中	小
先验框	（116×90）（156×198）（373×326）	（30×61）（62×45）（59×119）	（10×13）（16×30）（33×23）

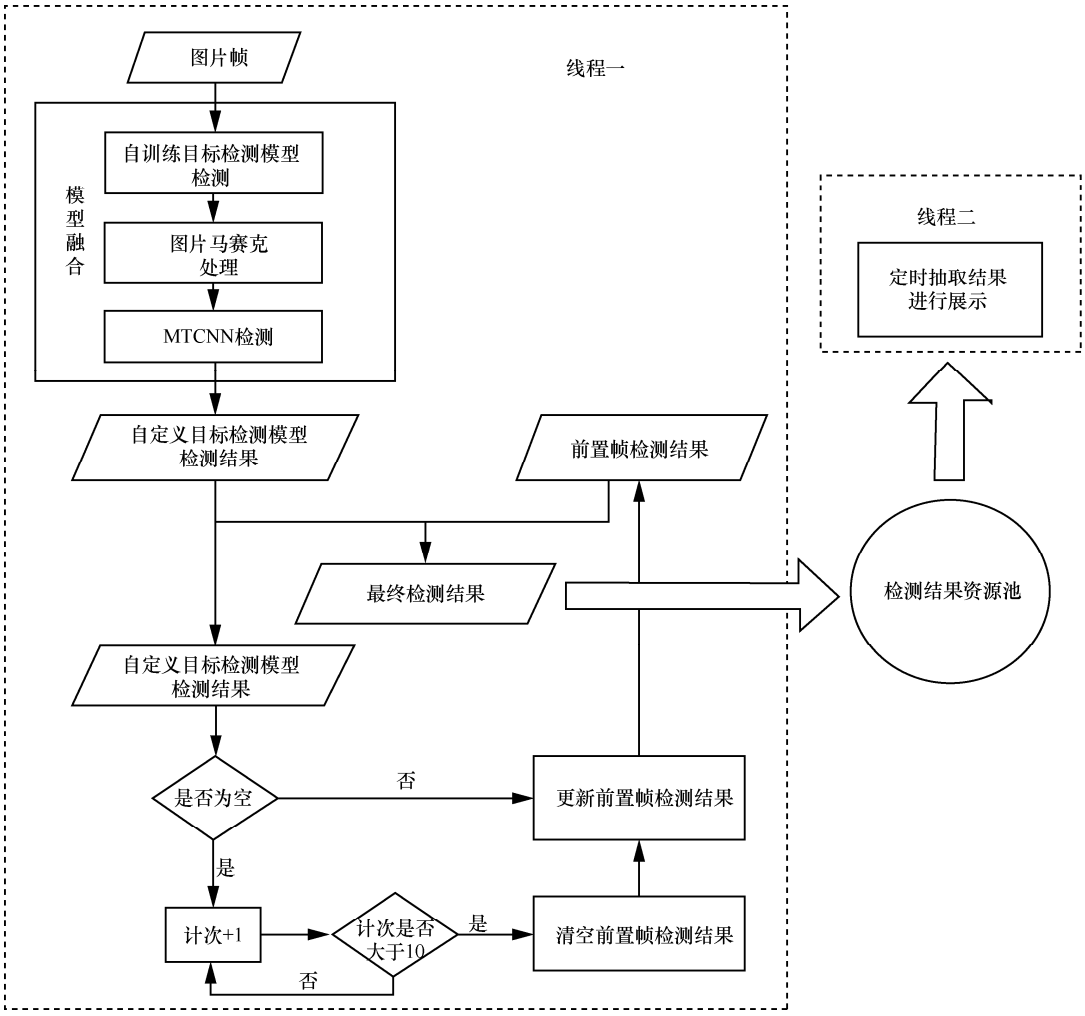

图 4　多检测模型流程

模型来对非正脸人头进行检测并进行马赛克处理，然后使用人脸检测模型对图片帧进行人脸检测并标记位置，然后与前置帧的标记区域进行合并，得到本文图片帧的所有标记区域并进行马赛克处理，处理后的图片帧将输入检测结果资源池中。然后自训练人头检测模型和人脸检测模型的合并结果，如果不为空值，将作为新的前置帧检测结果应用于后续帧检测。如果连续 10 帧均为空值，本文认为画面中确实没有需要马赛克的人脸，将清空前置帧检测结果。

5　实验与分析

在实验中，本文对所提算法框架进行了测试，并与其他两个传统的人脸检测模型——基于 Haar-like 特征和 Ada-boost 的传统人脸检测模型、基于深度学习的人脸检测模型 MTCNN 进行了对比。其中传统 Haar-like 特征结合 Ada-boost 分类器的模型已集成在 OpenCV 工具箱中[12]，MTCNN 使用默认预训练权重，本文算法中物体检测模型 YOLOv3 使用自建人头数据集进行训练。自建人头数据集分为两个部分：一部分为通过开源数据集获取具有人头的图片数据，场景多为教室等室内场景；另一部分通过在部分实际监控场景中进行采集，多为室外场景。人脸打码算法采用均值滤波算法，效果如图 5 所示。

图 5　均值滤波人脸打码效果

模型的优劣通常通过准确率和召回率两个指标来进行评价，准确率和召回率按式（1）和式（2）进行计算：

$$准确率=\frac{预测正确的人脸数量}{模型输出人脸的总数} \quad (1)$$

$$召回率=\frac{预测正确的人脸数量}{测试集中真实人脸总数} \quad (2)$$

本文准备了一段 1 min 时长的视频流并分别用上述 3 个模型进行人脸隐私保护处理，得到 3 段处理过的视频流。然后从 1 min 内随机抽取了 100 个时间点并分别从 3 段视频流中抽取对应时间点图片作为测试集。每段视频流抽取 100 张图片，其中共有 337 个真实人脸。在计算准确率和召回率的时候，模型计算的人脸定位坐标和实际人脸定位坐标的重叠度阈值设定为 50%，即当模型检测人脸定位框与实际人脸定位框重叠度高于 50%时认定模型检测正确，小于 50%认定为不正确。3 种模型在实验测试集上测试结果见表 2。

可以看出 3 种方法在准确率上取得了比较接近的效果，几乎全部的预测框都是正确的。但是在召回率上，传统 Haar-like 特征的表现较差，基本只能检测出正对摄像头且比较完整的人脸，基本不具备实用价值。基于深度学习的 MTCNN 算法相较于传统方法在人脸检测召回率上有了极大的提升，但总体而言检测效果也不尽如人意，虽然对人脸的各种姿态变化具有一定的稳健性，但对于测试集中在实际监控场景下经常出现的侧脸或者低头并不能很好地检测，在人脸转正或者抬起头的时候也无法在人脸即将可以被人眼辨识的时候及时把马赛克打上，依然存在人脸隐私泄露的风险。而本文多检测模型在设计上将人脸隐私保护任务进行了细化，划分为正脸检测的人脸检测任务和非正脸的物体检测任务，并利用视频流图片帧相互关联的特性对检测结果进行优化，在复杂场景视频流中检测效果取得了极大的提升，召回率相较于只使用人脸检测算法提高 532%。

6　结束语

在实验中，本文分别对基于传统方法和深度学习的人脸检测算法在基于监控视频场景的人脸隐私保护应用中的准确率和召回率进行了测试。实验表明，以上两种算法在准确率上表现均比较优秀，但由于监控场景中人脸姿态、角度变化相对复杂，两种算法在召回率指标上均表现不佳，不能很好地满足人脸隐私保护应用对于高召回率的要求，无法满足实际应用要求。针对以上问题，本文提出的结合人脸检测和目标检测的多检测模型并结合前置帧关联检测方法有效解决了人脸姿态、角度变化大的检测问题，大幅度提高了检测召回率，极大降低了马赛克漏打现象的出现，为人脸隐私保护应用落地商用往前迈进了一大步。

表 2　3 种算法模型准确率与召回率比较

	模型检测人脸数总数	检测正确人脸数	真实人脸数	准确率	召回率
Haar-like	16	16	337	100.00%	4.75%
MTCNN	52	52	337	100.00%	15.43%
多检测模型	334	329	337	98.50%	97.63%

在实验中，本文发现部分行人因为佩戴帽子或者口罩，系统在进行检测的时候存在一定困难，对在视频的边缘部分出现的行人或者由于前后遮挡导致的不完整人脸检测效果也不是很稳定。分析原因本文认为是在组织训练数据训练非正脸模型的时候对以上类别的数据收集过少，在以上情景下模型训练不够充分。在未来考虑将进一步通过丰富训练数据集，并通过随机擦除增强技术[13]进一步增加数据种类，提升模型对于不完整人脸的检测效果，降低模型在各种情景下的漏检现象。

参考文献：

[1] 毕玉谦, 洪霄. 民事诉讼生成权利规制探析——以“人脸识别第一案”为切入点[J]. 法学杂志, 2020, 41(3): 53-62.
BI Y Q, HONG X. Research on regulations of generating rRights in civil litigation: based on the first case of facial recognition[J]. Law Science Magazine, 2020, 41(3): 53-62.

[2] BRADSKI G, KAEHLER A. Learning openCV: computer vision with the openCV library[J]. IEEE Robotics & Automation Magazine, 2008(9): 100.

[3] 陈玲. 人脸检测技术综述[J]. 科学与信息化, 2018(16): 1.
CHEN L. Overview of face detection technology[J]. Technology and Information, 2018(16): 1.

[4] 侯小毛, 徐仁伯. 云环境中考虑隐私保护的人脸图像识别[J]. 沈阳工业大学学报, 2018(2): 203-207.
HOU X M, XU R B. Face image recognition based on privacy protection in cloud environment[J]. Journal of Shenyang University of Technology, 2018(2): 203-207.

[5] 张啸剑, 付聪聪, 孟小峰. 面向人脸图像发布的差分隐私保护[J]. 中国图像图形学报, 2018, 23(9): 1305-1315.
ZHANG X J, FU C C, MENG X F. Facial image publication with differential privacy[J]. Journal of Image and Graphics, 2018, 23(9): 1305-1315.

[6] PAPAGEORGIOU C P, OREN M, POGGIO T. A general framework for object detection[C]//Proceedings of the International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 1998: 555-562.

[7] 崔鹏, 燕天天. 融合YCbCr肤色模型与改进的Adaboost算法的人脸检测[J]. 哈尔滨理工大学学报, 2018(2): 91-96.
CUI P, YAN T T. Face detection combining the YCbCr skin color model with improved adaboost algorithm[J]. Journal of Harbin University of Science and Technology, 2018(2): 91-96.

[8] CHEN P H, LIN C J, SCHLKOPF B. A tutorial on ν-support vector machines[J]. Applied Stochastic Models in Business and Industry, 2005, 21(2): 111-136.

[9] ZHANG K, ZHANG Z, LI Z, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-1503.

[10] 陈正斌, 叶东毅, 朱彩霞, 等. 基于改进 YOLOv3 的目标识别方法[J]. 计算机系统应用, 2020, 29(1): 49-58.
CHEN Z B,YE D Y, ZHU C X, et al. Object recognition method based on improved YOLOv3[J]. Computer Systems Applications, 2020, 29(1): 49-58.

[11] LIN T Y, MAIRE M S, BELONGIE J, et al. Microsoft COCO: common objects in context[C]//Proceedings of European Conference on Computer Vision. [S.l:s.n.], 2014: 740-755.

[12] 刘卫华, 吴丹. 基于 OpenCV 的人脸检测[J]. 科技创新与应用, 2019(31): 18-19.
LIU W H, WU D. Face detection based on OpenCV[J]. Technology Innovation and Application, 2019(31): 18-19.

[13] 金翠, 王洪元, 陈首兵. 基于随机擦除行人对齐网络的行人重识别方法[J]. 山东大学学报: 工学版, 2018, 48(6): 71-77.
JIN C, WANG H Y, CHEN S B. Person re-identification based on random erasing pedestrian alignment network method[J]. Journal of Shandong University (Engineering Science), 2018, 48(6): 71-77.

[作者简介]

张驰（1990– ），男，中国电信股份有限公司上海分公司工程师、产品业务架构师，主要从事大视频和人工智能等技术工作。

陆晔（1975– ），女，中国电信股份有限公司上海分公司工程师、信息通信技术运营部总经理，主要研究方向为基于物联网、智能视频、大数据和人工智能的转型业务和产品的架构设计和全面推进。

罗渝平（1973– ），男，中国电信股份有限公司上海分公司工程师、信息通信技术运营部副总经理，主要研究方向为基于物联网、智能视频、大数据和人工智能的转型业务和产品的需求设计。

孙晓凯（1986– ），男，中国电信股份有限公司上海分公司工程师、智能视频应用中心副主任，主要研究方向为基于大视频的人工智能算法应用研发。

祝涵珂（1977– ），男，中国电信股份有限公司上海分公司工程师、DICT 运营中心副主任，主要研究方向为基于视频的算法算力平台运营。