

# 基于 WebSocket 的动态人脸识别系统构建

张海荣, 王正玉, 齐梅, 刘则芬

(安徽开放大学, 合肥 230022)

**摘要:**传统人脸识别系统的客户端与服务端交互基本上采用 HTTP 轮询或者定时任务触发方式,该方式不能满足动态人脸识别的实时性需求。为进一步提高识别精度和时效性,采用 MTCNN 多任务学习算法,构建了基于 WebSocket 的动态人脸识别系统。测试表明:该系统提高了人像识别率,支持浏览器与 Web 服务器的双向通信,能将识别出的人像信息及预警结果经服务端主动推送至客户端实时显示和告警,有效降低数据传输量,提高数据传输的稳定性和及时性。

**关键词:**WebSocket;动态人脸识别;视频数据处理

**中图分类号:**TP391.41

**文献标识码:**A

**文章编号:**2097-0625(2022)03-0087-05

## 一、引言

近几年,随着各地平安城市、智慧城市等项目的推进,大量的监控摄像机投入使用。网络监控作为视频监控系统的的重要组成部分,得到了广泛的应用,是社会治安防控的重要手段,对平安城市、智慧城市的建设有着巨大的促进作用<sup>[1-4]</sup>。动态人脸识别技术结合了人脸识别和动态视频监控技术,能够满足特定领域视频监控需要,成为当前研究热点之一。近年来,研究出现很多新的改进算法<sup>[5-9]</sup>,如文献[7]以 Resnet34 为基础网络,结合联合损失和恒等映射提出的动态人脸识别算法有效提高了识别效率。但传统人脸识别系统客户端与服务端的交互基本上采用 HTTP 轮询或者定时任务触发,需要客户端不断地与服务器建立 HTTP 连接,然后等待处理。这种不断反复请求解析的模式非常消耗资源,也导致了消息同步的延迟,甚至告警频率过高时,会造成数据丢失。数据延迟和丢失会带来极大的弊端,可能导致安保人员错失先机或发现不了问题,不利于处理突发状况。因此,有效解决人脸识别系统中告警数据推送是十分必要的。

随着 HTML5 技术广泛普及,主流浏览器均对

其具有较好的兼容性,客户端和服务端之间的实时数据交换变得更加容易。WebSocket 是 HTML5 新标准中的通信机制<sup>[10-11]</sup>,能够实现稳定全双工实时通信。本文基于此设计了基于 WebSocket 的动态人脸识别系统,能够使服务端有信息及时主动推送给客户端,避免了反复解析协议和数据头信息重复发送,提高了信息交换效率和服务器使用率,有效解决了信息延迟<sup>[12-14]</sup>。

## 二、动态人脸识别技术概述

动态人脸识别技术是由人脸识别技术融合了图像处理技术和生物统计学理论形成的,是当前比较新的一项生物识别技术。该技术主要基于人的脸部特征,对输入的人脸图像或者视频流进行检测判断,进而对其识别的。通常动态人脸识别包括以下四个流程:

- (1)人脸检测:指经检测算法从动态的场景中判断是否存在人像并分离出图像的过程。
- (2)人脸追踪:对检测到的目标图像进行追踪。
- (3)人脸特征提取:指对人脸进行特征建模的过程,是用来衡量图片相似度的,又称人脸表征。
- (4)人脸识别:将要查询的人脸特征模板与过程

**收稿日期:**2022-02-23

**基金项目:**安徽省高等学校自然科学研究项目“基于计算机视觉的在线学业情绪分析及学习倦怠预测研究”(项目编号:KJ2021A1255);安徽开放大学校级科研项目“数字鸿沟视角下老年远程教育智慧助老策略研究”(项目编号:QN202108)

**作者简介:**张海荣(1987—),女,安徽阜阳人,助教,硕士。研究方向:图像处理、远程教育。

(3)中所提取的待识别的人脸特征进行比较,设定阈值,当比较所得相似度超过阈值,则认为匹配成功。

### 三、动态人脸识别系统设计

#### (一)系统总体设计方案

动态人脸识别系统设计基于 WebSocket 协议实现。系统包括视频识别终端、人脸识别算法服务、人脸识别系统、消息服务等。图 1 为系统组成。

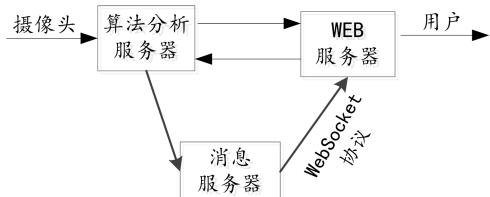


图 1 系统组成

系统首先在 web 服务器上发布布控任务,随后通过摄像头捕获视频数据,经算法分析服务器中人脸识别算法进行人脸检测、追踪、特征提取及识别。最后将识别出的人像及预警结果经消息服务器主动推送至 Web 服务器进行实时显示和告警。图 2 所示为系统数据处理流程,其中框选部分为图像抓拍结果输出数据流程。

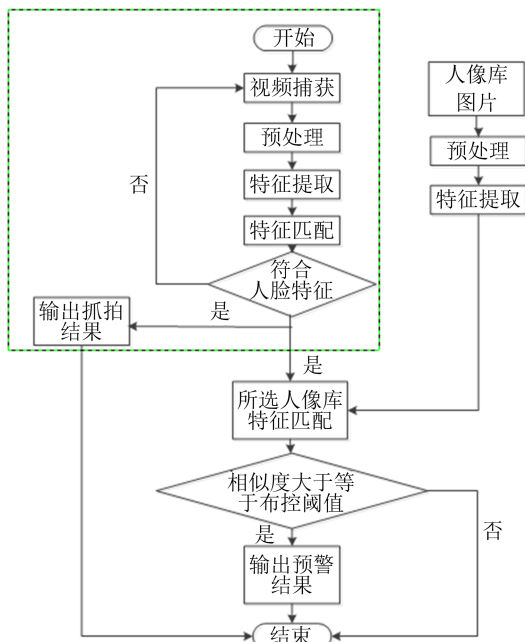


图 2 数据处理流程

#### (二)系统实现

##### 1. 服务部署

在算法分析服务器上部署人脸识别算法服务;在 Web 服务器上部署人脸动态识别系统 Web 服务;在

消息服务器上部署 java 消息服务,该消息服务采用支持 WebSocket、AMQP 和 STOMP 协议的 Apache Apollo 软件。

##### 2. 准备样本图像

将至少一个人像样本库上传至 Web 服务器,所述人像样本库中至少包含一个需要进行对比的样本人像。

##### 3. 建立服务连接,发布识别任务

Web 服务器向算法分析服务器下发布控任务;Web 服务器向消息服务器发送订阅消息请求并与消息服务器建立连接;Web 服务器向算法分析服务器下发布控任务,指定一个人像样本库、指定摄像头、设定相似度阈值和发送识别任务指令。

如图 3 所示,(a)和(b)分别为 HTTP 协议和 WebSocket 协议的消息传送原理图,可以看到 HTTP 交互方式,要不断地建立和关闭 HTTP 协议,且 HTTP 是无状态性的,每次都要重新传输 identity info(鉴别信息);而 WebSocket 只需要建立一次连接,直到你关闭请求,同时由客户端主动询问,转换为服务端有信息时主动推送。这就避免了反复解析 HTTP 协议和查看 identity info 信息,提高了宽带利用率和信息交换效率,有效地实现了消息的及时性。

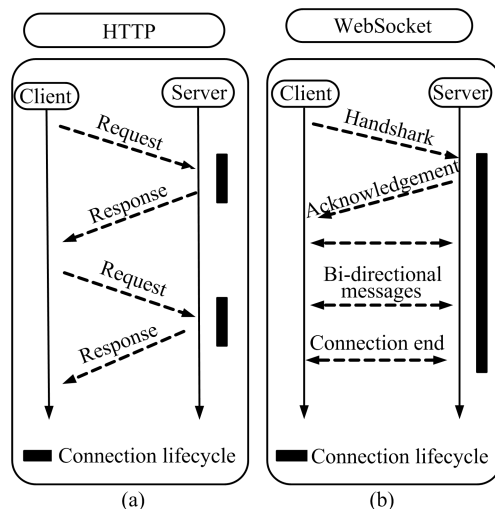


图 3 消息传送原理

如图 4 所示,Web 服务器向消息服务器发送订阅消息请求并与消息服务器建立连接,具体步骤如下:

(1)Web 服务器调用 WebSocket 的初始化接口 URL。

(2) Web 服务器使用 STOMP 客户端库 stomp.js 中 over 方法, 获取 STOMP 子协议的客户端对象, 即消息服务器。

(3) Web 服务器向消息服务器发起 WebSocket 连接请求并发送 CONNECE 帧。

(4) 消息服务器接收 Web 服务器发送来的连接请求并与 Web 服务器成功连接。消息服务器接收 Web 服务器发送来的连接请求, 并判断数据包数据格式和两端协议版本号是否匹配, 匹配则接受本次连接请求, 并与 Web 服务器成功连接; 不匹配消息服务器则将错误通知反馈给 Web 服务器, Web 服务器端找出并解决对应的错误, 直至消息服务器与 Web 服务器成功连接为止。

(5) 经过三次步骤 4 的连接请求并成功连接后, 触发 connectCallback 事件, Web 服务器与消息服务器之间形成 TCP 通道连接。

#### 4. 执行识别任务

算法分析服务器对目标人像与人像样本库中样本人像进行对比, 并将对比结果发送给服务器。具体步骤如下:

(1) 算法分析服务器接收摄像头抓拍的目标人像和 Web 服务器中的人像样本库中的样本人像, 对目标人像和样本人像进行预处理操作, 利用 MTCNN 多任务学习算法<sup>[15]</sup>提取图像人脸特征。人脸面部关键点坐标包括左眼、右眼、鼻子、左嘴角和右嘴角等坐标。

(2) 通过卷积神经网络技术将待识别人像与对比样本人像进行特征匹配, 匹配后得到的所有对比相似度值与 Web 服务器设定的相似度阈值进行比较, 并将目标人像和对比相似度值大于或等于相似度阈值的人像样本库中的相关样本人像发送给消息服务器。若对比相似度值均小于相似度阈值, 则仅将目标人像发送给消息服务器。

#### 5. 结果显示

消息服务器接收目标人像和对比结果, 并主动发送给 Web 服务器, Web 服务器显示目标人像和对比结果。随着摄像头不停地抓拍, 通过上述步骤, Web 服务器可以不停地显示抓拍的目标人像和样本库中的相关样本人像对比相似度值大于或等于相似度阈值的人像, 从而实现动态人脸识别。

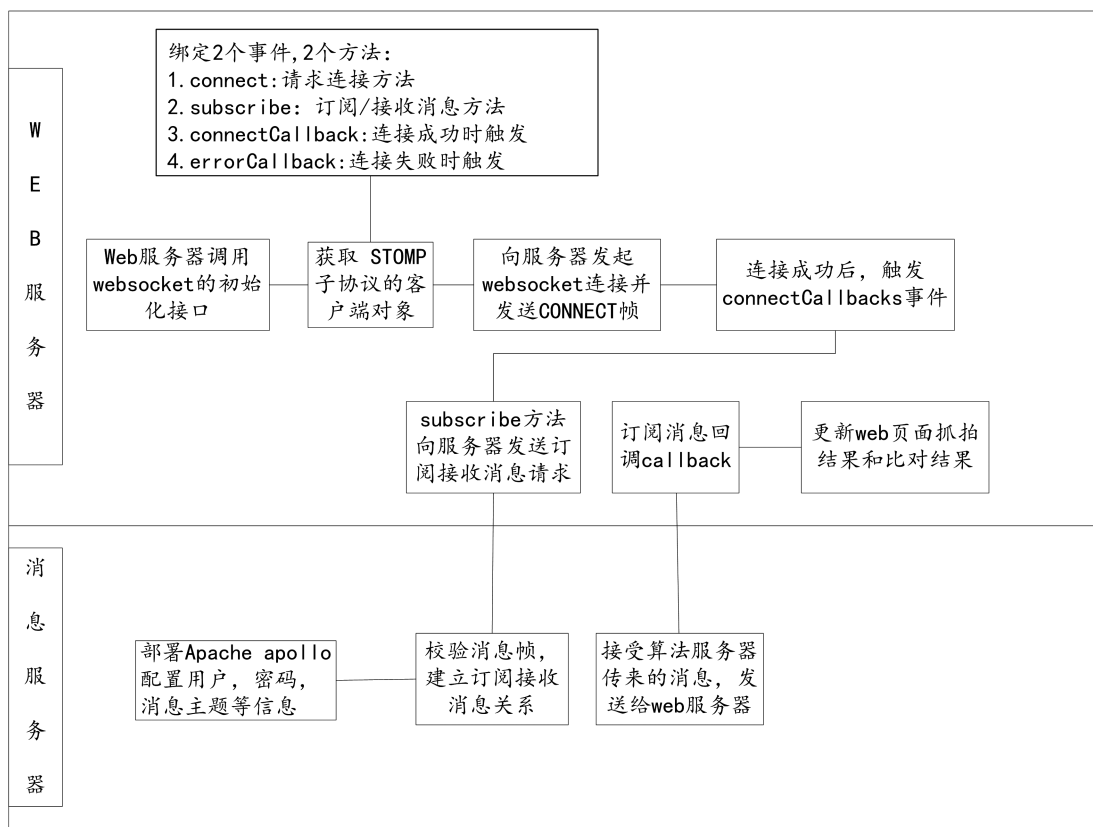


图 4 Web 服务器与消息服务器之间的通信

### (三) 系统测试

系统 Web 前端基于 vue 框架开发, 后端采用 Java 语言。人脸检测及识别算法采用 MTCNN 多任务学习算法<sup>[15]</sup>, 具有更高的精度和实时性能。系统在设计实现之后, 即部署在某产业园区进行测试。测试环境如下: 服务器硬件配置为 16 核, 8 GB 内存。服务器操作系统为 Linux, 开发语言是 Java, 编译环境为 JDK1.8, 数据库使用 Oracle 11.0.2.3 数据库。

表 1 系统测试结果

| 人像库容量(人) | 误识别数量(人) | 正确识别数量(人) | 识别准确率(%) | 平均检测速度(ms) | 图片传输丢失率(%) |
|----------|----------|-----------|----------|------------|------------|
| 50       | 3        | 47        | 94       | 513        | 0          |
| 100      | 5        | 94        | 95       | 512        | 0          |

从表 1 可以看出, 在不同大小人像库, 系统均实现了较好的识别准确率; 检测速度快, 图片前后端传输零丢失, 验证了 WebSocket 数据传输的稳定性和安全性。针对人像未识别情况进行分析, 发现是测试过程中员工经过摄像头抓拍区一直处于低头状态, 导致无法采集到人脸。此外, 系统对浏览器兼容性进行了测试, 结果表明该系统具有较好的兼容性, 支持新版 Chrome、IE11、火狐等主流浏览器。综上所述, 该系统运行稳定, 识别率高且识别出的人像信息实现了

客户端环境为 Windows 10 操作系统, 浏览器为 Chrome 99.0.4844.84 及其自带开发工具。

为保证比对人像库图片的高质量, 人像库图片直接采集部门员工证件照, 测试时间选择员工中午就餐时间, 阈值设定 0.8, 其中识别准确率 =  $TP / (TP + FP) * 100\%$ 。TP 指算法正确判定的识别数量, FP 指算法错误判定的识别数量。具体测试结果见表 1。

实时显示和告警; 数据传输稳定安全, 实现图片传输零丢失; 具有较好的兼容性, 满足了实际生产需求。

### (四) 运行界面

图 5 为系统最终运行界面, 界面下方展示为实时抓拍区, 右侧展示为抓拍识别区, 中间实时播放视频监控影像。当摄像头抓拍识别到目标人像时, 消息服务器将实时推送数据至抓拍区。当人像对比相似度高于设置阈值时, 服务器同时推送至抓拍识别区并产生预警信息。



图 5 系统运行界面

### 四、结语

本文研究设计实现了基于 WebSocket 的动态人脸识别系统, 其工作流程如下: Web 服务器向算法分析服务器下发识别任务; Web 服务器向消息服务器发送订阅消息请求并与消息服务器建立连接; 算法分析服务器对抓拍的人像与人像样本库中的样本人像

进行对比, 结合设定阈值, 将对对比结果发送给消息服务器; 消息服务器接收对比结果, 并主动发送给 Web 服务器, Web 服务器显示结果。Web 服务器通过 WebSocket 协议与消息服务器建立全双工 (full-duplex) 通信, 避免了反复解析协议和数据头信息重复发送的问题, 提高了信息交换效率和服务器 CPU 使

用率, 实现了人脸识别消息的及时性, 有效避免了因数据滞后等问题, 满足了实际应用需求。下一步, 将

研究根据人像识别位置绘制人像轨迹, 更加便捷地识别出人员运动轨迹。

#### 参考文献:

- [1] 王成刚. 互联网时代下平安城市网络监控研究[J]. 数字技术与应用, 2019, 37(10): 51-51.
- [2] 杜庆灵. 平安城市视频监控系统网络安全研究[J]. 科技与创新, 2019(22): 59-60.
- [3] 单晓辉, 单洪伟, 刘祥. 公安视频数据安全面临的机遇与挑战[J]. 中国公共安全, 2019(5): 117-121.
- [4] 周正. 城市视频监控应用系统中监控中心的设计与分析[J]. 电子世界, 2020(5): 150-151.
- [5] 杜艳明, 刘齐跃. 基于多姿态脸和小波变换的动态人脸识别[J]. 计算机应用, 2007(S2): 83-84.
- [6] 李宏菲, 李庆, 周莉. 基于多视觉描述子及音频特征的动态序列人脸表情识别[J]. 电子学报, 2019, 47(8): 1643-1653.
- [7] 刘成攀, 吴斌, 杨壮. 基于联合损失和恒等映射的动态人脸识别[J]. 传感器与微系统, 2021, 40(9): 153-156.
- [8] 温荷, 罗频捷. 基于改进脉冲耦合神经网络的动态人脸识别[J]. 计算机科学, 2021, 48(S1): 85-88.
- [9] 何元烈, 刘峰, 孙盛. 基于深度学习的动态人脸识别方法[J]. 计算机工程与设计, 2019, 40(11): 3299-3305.
- [10] 李代立, 陈榕. WebSocket 在 Web 实时通信领域的研究[J]. 电脑知识与技术, 2010, 6(28): 7923-7925.
- [11] 薛陇彬, 刘钊远. 基于 WebSocket 的网络实时通信[J]. 计算机与数字工程, 2014, 42(3): 478-481.
- [12] LIU Zhiqiang, ZHU Wenbo, ZHANG Hongzhou, et al. Reliability Evaluation of Dynamic Face Recognition Systems Based on Improved Fuzzy Dynamic Bayesian Network[J]. International Journal of Distributed Sensor Networks, 2020, 16(3): 1-12.
- [13] 赵晗. 基于 WebSocket 协议的 Web 应用性能测试[D]. 北京: 中国科学院大学, 2014: 51-60.
- [14] 郑强, 徐国盛. WebSocket 在服务器推送中的研究[C]//中国通信学会. 第九届中国通信学会学术年会论文集. 北京: 中国通信学会, 2012: 376-381.
- [15] 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.

## Construction of Dynamic Face Recognition System Based on WebSocket

ZHANG Hairong, WANG Zhengyu, QI Mei, LIU Zefen

(Anhui Open University, Hefei 230022, China)

**Abstract:** The interaction between the client and the server of the traditional face recognition system basically adopts HTTP polling or timed task triggering, which can not meet the real-time requirements of dynamic face recognition. In order to further improve the recognition accuracy and timeliness, the dynamic face recognition system based on WebSocket is constructed by using MTCNN multi-task learning algorithm. The test shows that the system improves the portrait recognition rate, supports the two-way communication between the browser and the web server, and can actively push the identified portrait information and early warning results to the client through the server for real-time display and alarm, effectively reduce the amount of data transmission, and improve the stability and timeliness of data transmission.

**Keywords:** WebSocket; dynamic face recognition; video data processing

[责任编辑 李潜生]