

系统的视频流传输结构如图所示，通过 Nginx 和 FFMPeg 搭建流媒体服务器。这个流媒体服务器可以支持 RTMP 协议。RTMP 协议基于 TCP 协议的，能够为数据的传输提供可靠保障，因此数据在网络上传输不会出现丢包的情况。通常延迟在 1-3s。

在人脸验证方面，考虑到系统可能只有一个或几个给定人脸的样本，且必须从新照片中正确识别出该人，即通过单样本学习进行人脸识别。

上面的图为传统卷积神经网络下的图像分类任务，而下图则是用孪生神经网络进行图像相似度计算。孪生神经网络是单样本学习中最具代表性的模型。其目标是寻找两个对象的相似程度。这个网络有两个相同的子网络，两个子网络有相同的参数和权重。

FaceNet 是谷歌提出的基于孪生神经网络的通用系统，可以用于人脸验证，识别和聚类。FaceNet 直接使用基于 triplets 的 LMNN (最大边界近邻分类) 的 loss 函数训练神经网络，网络直接输出为映射到欧式空间的 128 维的向量。

triplet 是一个三元组，从训练数据集中随机选一个样本，该样本称为 Anchor，然后再随机选取一个和 Anchor 属于同一类的样本和不同类的样本，这两个样本对应的称为 Positive 和 Negative，由此构成一个 (Anchor, Positive, Negative) 三元组。

人脸验证的思路是通过人脸检测器将人脸从图像中提取出来，通过 facenet 进行验证，我们对比了不同的人脸检测器，发现 OpenCV 中提供的 DNN 运行速度快，对遮挡等情况适用。此外，从用户角度出发，在建立人脸数据库时，通过 OpenCV 和 win32com 从文字和语音方面进行操作提示。

在微笑检测上，我们使用了基于 fer2013 数据集的 Mini-Xception 模型，该模型移除了全连接层，结合了残差模块和深度可分离卷积减少参数，达到实时检测的目的。

以下是对 face++ 提供的人脸和微笑检测接口与 Mini-Xception 的对比。

综上，我们首先建立人脸数据库，通过人脸检测器提取出图片中的人脸记性处理，利用 facenet 计算新图片中人脸与数据库中人脸的欧式距离，若最小值大于选定的阈值，则发出陌生人警告，反之，若新图片中的人脸为老人，则通过 Mini-Xception 选择置信度最大的情绪识别结果。

摔倒检测方面，从单幅图像上来看，我们可以提取出人的高度、宽度和夹角信息，从时间序列上来看，除了上述特征，我们还可以提取出速度和加速度，摔倒行为的特征表现如图所示。背景差分法由于易受到摄像头摆动等干扰测量准确度，故我们使用 OpenPose 提取人体关键点。OpenPose 是由 C++ 开发的开源库，可用于实时人体姿态估计等，为了在 python 中使用，需要对其进行编译。但 OpenPose 在人体姿势极端的情况下，识别准确度下降，为识别摔倒的老人通过挥手臂等方式求救，我们实现了基于人体关键点和背景差分的摔打检测算法，利用人力关键点提取出速度、高度、宽度等描述特征，当发生异常通过背景差分提取出角度特征，当特征连续多次满足摔倒检测阈值时，则发出摔倒报警。

在互动检测上，我们假定老人与义工的距离小于 50cm 即发生互动。因此需要建立像素坐标系与世界坐标系的映射关系。以 $h(x)=2x^6+2x^5+x^4+3x^3+2x^2+2x+1$ 作为本原多项式生成伪随机序列，通过伽罗华域下的四则运算，生成 65×63 的伪随机矩阵，窗口大小为 2×3 。编码图案以菱形作为基本图案，菱形的角点作为特征点，以红、蓝、绿、黑四种颜色作为菱形的颜色。分别对应矩阵中的不同数值，通过颜色及位置对特征点进行编码，从而可以唯一确定图案中某一角点的位置，进而确定不同角点在世界坐标系下的相对位置。通过设计的解码算法，可以将角点位置精确到亚像素，在 Lab 演色空间下进行颜色分类，完成解码。如图分别为拍摄原图，角点位置和颜色分类结果。

项目管理上，我们采用 github 进行代码托管，三端分别提交 101 + 63 + 18 次，每天保证进行每日站会和进度汇报，形成 28 次会议纪要。生成项目章程，需求规格说明书、用户手册等五份文档。通过甘特图对项目进度进行管理，项目进行过程中，执行情况良好。