

老师、同学们大家好：

接下来由我代表我们组参加进行展示。

我将从以下六个方面对我们的工作进行陈述。

小组由四人组成，我主要负责计算机视觉部分的设计与开发和项目管理，罗瑶同学负责 Web 界面的开发与需求分析，施卓余同学负责系统架构和接口的设计以及部分 Web 界面开发，冯焱同学主要负责数据库设计与实现以及开发和文档的辅助工作。

本系统系统分为 Web 端和计算机视觉 2 部分。通过摄像头实时拍摄到的画面，对老人的情感、摔倒、入侵、互动等事件进行检测与分析。一旦上述事件发生，该事件会被实时地更新在 Web 界面中，从而管理人员可以迅速做出反应。系统 Web 端仅供系统管理员使用。在对项目的问题、范围等作出初步分析后，我们制定了项目章程，以指导项目后续的开发、管理等工作。

我们通过获取原型的方法加速对系统的分析，利用 figma 对原型进行设计与简单构造，在此过程中明确系统需求。本系统的两个主要 Actor 为管理员和摄像头，管理员希望系统可以提供对老人、员工等的人员管理、对实时监控、微笑、摔倒等事件的查看以及其他的基本功能等。而摄像头则将采集到的数据和对检测到事件进行分析处理、并进行数据传输。

在项目章程中，我们制定了项目的时间管理计划和里程碑，以下为项目进行到现在的执行情况和部分交付物。项目采用 github 进行代码托管，目前累计提交 112 次，三天完成前后端的主体功能，具体表现为完成了注册/登录、Cookie 等基本功能、三种人员管理、事件统计等功能，老人分析待完成，完成了单样本人脸识别、摔倒检测等功能、待完成的有与义工互动检测和数据传输。

系统的开发采用较为成熟的 Vue 和 Django 框架，计算机视觉部分以 Python 作为主要的开发语言。

实时监控方面，采用本地推流的方式，在阿里云服务器上搭建 nginx-rtmp 流服务器，通过 Ffmpeg 将 opencv 捕获的视频流进行编码、推流，以 RTMP 协议进行传输。RTMP 协议基于 TCP 协议，具有实时性高、支持加密、易实现等优势，通常时延在 3s 以内，常用作直播协议。

卷积神经网络在图像分类任务上取得了较好的训练效果，经典的 CNN 有 LeNet-5, VGG, ResNet 和 DenseNet 等。在本系统的人脸识别中，我们希望通过输入一张人脸图片就能够识别对应的人的身份。而如果系统并没有成功地识别图片，是因为这个人脸图片并不在系统的数据库中。为了解决这个问题，我们不能仅使用单个卷积神经网络的原因是 CNN 在小数据集上是不起作用的；而如果每次向系统中加入一位新人的一幅图片时再去重新训练模型，也是极其不方便的。为此，我们在人脸识别中使用了孪生神经网络。

Siamese network 就像是“连体的神经网络”，神经网络的“连体”是通过共享权值来实现的，第一个子网络的输入是一幅图片，然后依次馈送到卷积层、池化层和全连接层，最后输出一个特征向量 ($f(x_1)$ 是对输入 x_1 的编码)。然后，向第二个子网络输入图片 x_2 做相同的处理，并且得到对 $x(2)$ 的编码 $f(x_2)$ 。为了比较图片 $x(1)$ 和 $x(2)$ ，我们计算了编码结果 $f(x_1)$ 和 $f(x_2)$ 之间的距离。如果它比某个阈值小，则意味着两张图片是同一个人，否则，两张图片中不是同一个人。在本项目中，通过第三方库提供的人脸检测器确定录入人脸的图片范围，提取人脸部的 128 个特征，在实时视频中用同样的模型提取 128 个特征，计算二者之间的距离，若小于阈值，则判断为同一个人。

在人脸检测的选择上，我们对比了 dLib 和 OpenCV 提供的不同检测器的有关指标，为达到实时、准确的目标，最终选择了 OpenCV 提供的 DNN 检测器。

在微笑检测上，我们尝试了逻辑回归等分类方法，由于表情模型和拍摄环境的复杂多变，并不适合这一任务。在阅读部分文献的基础上，我们借用一篇 2017 发表论文设计的模型，

通过 keras 提供的 Xception 模型进行表情分类, Xception 融合了残差模型和深度可分离卷积两大特性, 其设计的模型具有实时参数少等特点。

在摔倒检测方面, 我们设计了基于人体关键点和背景差分的检测算法, 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, 以下为推荐的编译环境。其鲁棒性、实时性较高, 但耗显存, 极特殊场景下检测效果差。

在摔倒检测的特征中, 我们主要参考宽度、高度、速度、加速度和角度五个指标。通过 OpenPose 提供的人体坐标计算画面中老人的速度, 加速度和长宽比等, 对符合摔倒特征的画面进行判断, 当较短时间连续发生则判断为摔倒。而正如之前提到的, 在光照, 人体姿势扭曲等极端条件下, OpenPose 不能很好地进行人体姿态估计, 这时通过背景差分法, 对运动物体的轮廓进行拟合, 检测其与水平夹角, 当多次大于阈值时, 判断为摔倒。

Web 界面已部署在公网服务器上, 以下为部分重要界面截图。

上面图片为自实现的人脸录入与笑脸检测功能, 下图为调用公用接口的人脸与笑脸检测的结果。

下面通过一段视频对摔倒检测进行演示, 可以看到当摔倒发生时, 运行窗口打印出了 fall。

针对现阶段的工作, 未来我们需要继续完成剩余功能开发, 并进行系统测试以及文档撰写。

以上就是我答辩的全部内容, 请老师批评指导。