基于人脸特征点的疲劳度检测系统

第二小组

陈博楷 陈帅 陈劲宇 陈鹏宇

目录

[一、 研究背景简介 2](#_Toc105834185)

[二、 研究问题定义 2](#_Toc105834186)

[三、 数据集介绍 2](#_Toc105834187)

[四、 算法模块介绍 3](#_Toc105834188)

[（一） MTCNN网络 3](#_Toc105834189)

[1. 构建图像金字塔 3](#_Toc105834190)

[2. P-net网络 3](#_Toc105834191)

[3. R-net网络 4](#_Toc105834192)

[4. O-net网络 4](#_Toc105834193)

[5. 非极大值抑制（NMS） 4](#_Toc105834194)

[6. MTCNN网络训练过程 4](#_Toc105834195)

[（二） CNN网络 5](#_Toc105834196)

[（三） 网络连接方式 5](#_Toc105834197)

[五、 系统运行结果 6](#_Toc105834198)

[六、 总结体会 7](#_Toc105834199)

[（一） 组员分工情况 7](#_Toc105834200)

[（二） 学习体会 7](#_Toc105834201)

1. 研究背景简介

随着人们生活水平的逐渐提高以及各大城市道路交通系统的不断完善，我国的汽车保有总量也在不断增加，这在为人们的出行带来便捷的同时也导致了交通事故的频频发生，对驾驶员和行人的生命财产安全构成了巨大的威胁。

当驾驶员长时间内持续驾驶而没有得到足够休息时，其生理机能和心理机能将会逐渐失调，对车辆的控制能力、注意力以及面对突发事件时的反应能力都会明显下降，甚至可能坠入短暂睡眠，导致车辆彻底失控，这种现象叫做疲劳驾驶。

当前疲劳度检测的类型总共有三大类，分别为：基于人体生理信号检测、基于车辆行驶路径变化检测、基于人体可视特征检测。基于人体生理信号检测的方法优点在于检测精度高，但是缺点也很明显，这种检测方法用到的仪器复杂，并且接触式测量会影响正常驾驶。基于车辆行驶路径变化检测是一种间接检测疲劳度的方式，其优点是信息容易采集，不会影响驾驶，缺点在于这种方式容易受到路况、驾驶习惯等因素的影响，导致检测精度不足。基于人体可视特征检测是目前应用最广泛的检测方式，它在保证检测精度的同时不会影响驾驶员正常驾驶。基于人体可视特征的方式不需要复杂的仪器，但是这种方式要求输入检测系统的图像要具有较高的质量，并且要求系统拥有较强的实时性，。尽管如此，综合各项因素分析，基于人体可视特征检测的方式已然成为现在的主流方式。

1. 研究问题定义

本系统属于基于人体特征检测的系统，系统主体功能是判断驾驶员的眨眼情况，根据眨眼情况确定驾驶员是否处于疲劳状态。系统整体功能可以拆分为以下几个步骤：

1. 收集图像并进行预处理。
2. 图像人脸位置及关键点判定。
3. 检测眨眼情况。
4. 输出结果。

上述步骤中，对图像中人脸位置和关键点的判定以及眨眼检测是系统中的重要功能。同时也要保证检测正确率的同时兼顾检测速度，保证实时性。

1. 数据集介绍

CEW数据集

项目中用到的数据集为CEW（Closed eyes in the wild）闭眼数据集。该数据集由南京航空航天大学提供，包含了2423个测试者睁眼与闭眼状态的照片，照片的差异化体现于测试者个体的差异以及环境变化，如光照、模糊度、等因素。

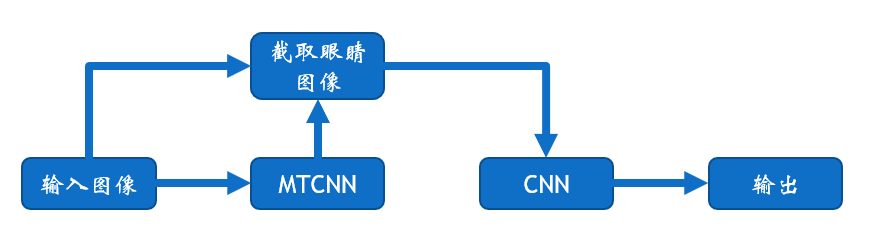


原始数据集包含1192张左眼或右眼的闭眼图像，1231张左眼或右眼的睁眼图像，总共4846张灰度眼睛图像。为保证检测数据的准确度，通过数据增强手段将4846张眼睛图像扩充为14538张眼睛图像。

该数据集用于训练CNN网络，由于MTCNN网络训练成本过高，电脑运行时出现内存溢出情况，因此使用了预训练好的模型。

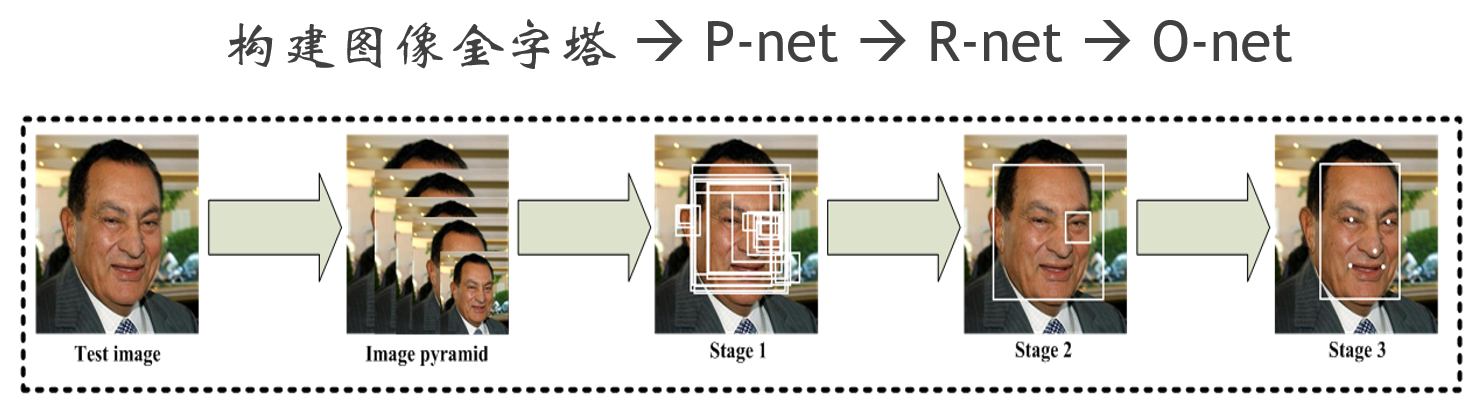
1. 算法模块介绍

本系统使用的算法模块有两部分，分别是MTCNN网络和CNN网络。MTCNN网络用于判断并框选人脸位置，并在人脸框内使用五个点标定人脸关键点。CNN网络用于判断眨眼情况，输入为一张眼部图片，输出为眼睛睁闭情况。整体结构如下：



1. MTCNN网络

MTCNN网络主要采用三个级联的网络，利用候选框加分类器的思想，进行快速高效的人脸检测。三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的R-Net和生成最终边界框与人脸关键点的O-Net。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。其主要流程如下：

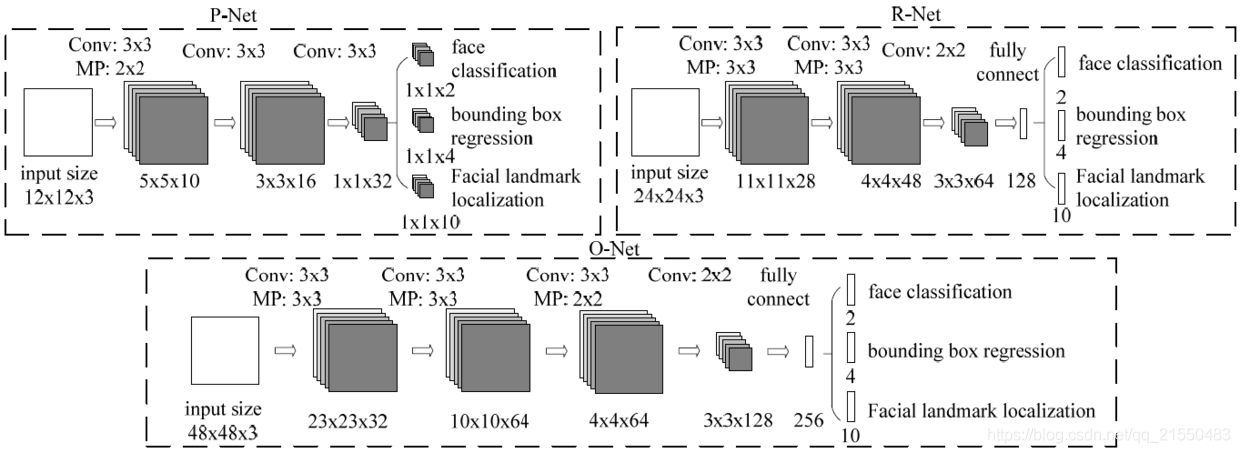


1. 构建图像金字塔

为了网络能够对不同大小的图像进行标定，训练时首先将图片进行不同尺度变换，以适应不同大小的人脸检测。构建方式为通过不同缩放系数factor对图片进行缩放，每次缩小为原来的factor大小，但是缩放最小尺寸不可以小于12。MTCNN网络规定最小的图像大小为12。

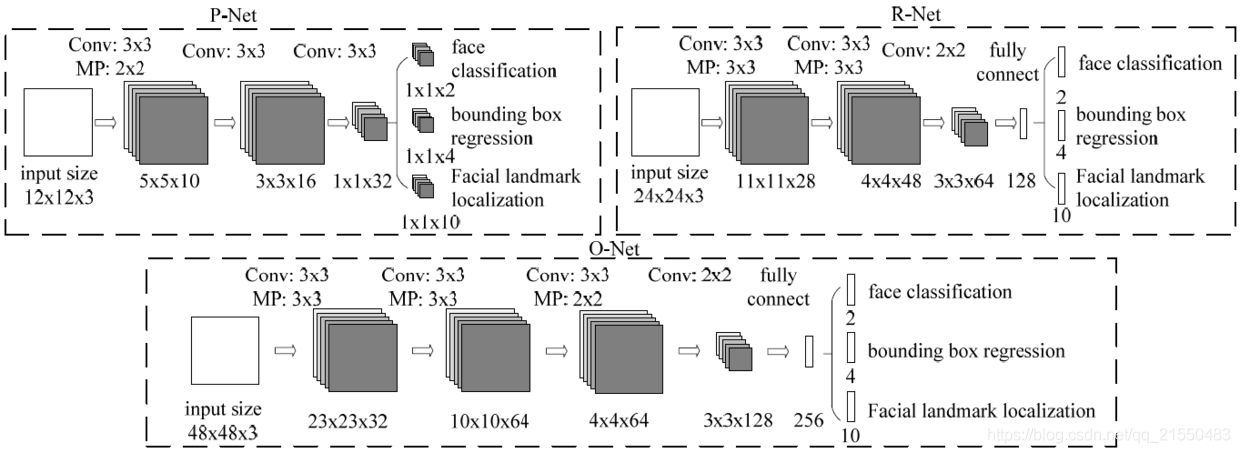
1. P-net网络

P-net网络的基本构造是全卷积网络。对构建完成的图像金字塔，通过一个FCN进行初步特征提取与标定边框。输出为类别、标定框位置和关键点坐标。这部分网络的结构如下：



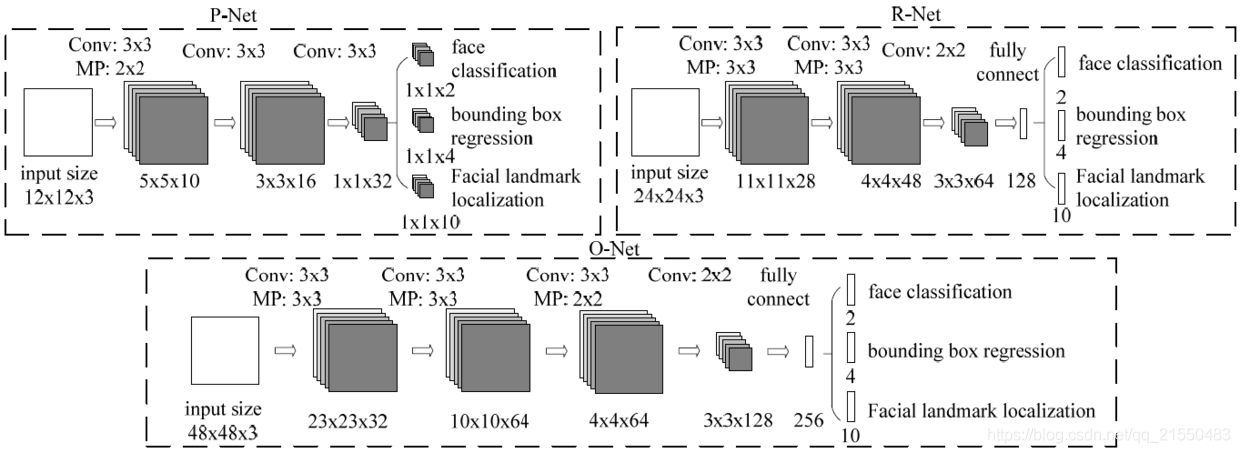
1. R-net网络

R-net的基本构造是一个卷积神经网络，相对于P-Net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。图片经过P-Net后会留下许多预测窗口，将所有的预测窗口送入R-net，R-net会滤除大量效果比较差的候选框。输出同P-net网络。



1. O-net网络

O-net的基本结构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，最终输出类别，标定框和五个人脸面部特征点。



1. 非极大值抑制（NMS）

非极大抑制的主要作用在于将一定区域内属于同一种得分最高的框筛选出来并将与该框重合度较高且得分相对低的框剔除掉。操作方式为：把所有候选框按置信度大小排列，将置信度最大的候选框与其余候选框依次计算交并比（IoU），当交并比大于某一阈值时将这个候选框删除。然后在剩余未删除的框中重复操作。

1. MTCNN网络训练过程

使用Wider\_face和CelebA数据集进行训练，其中Wider\_face用于检测任务的训练，CelebA用于关键点的训练。训练集分为四种:负样本，正样本，部分样本，关键点样本。三个样本的比例为3: 1: 1: 2。

训练主要包括三个任务：

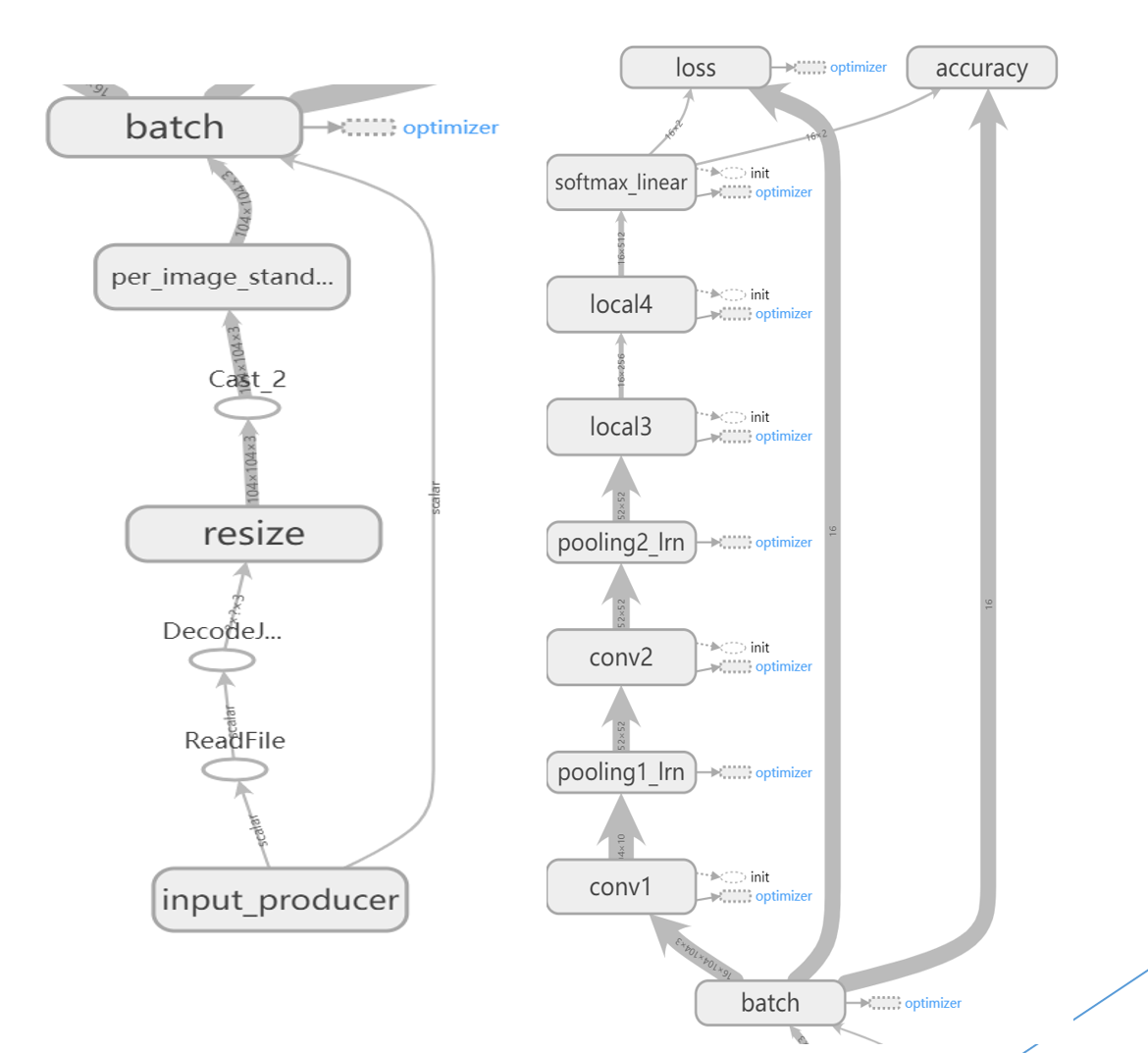
人脸分类任务：利用正样本和负样本进行训练。

人脸边框回归任务：利用正样本和部分样本进行训练。

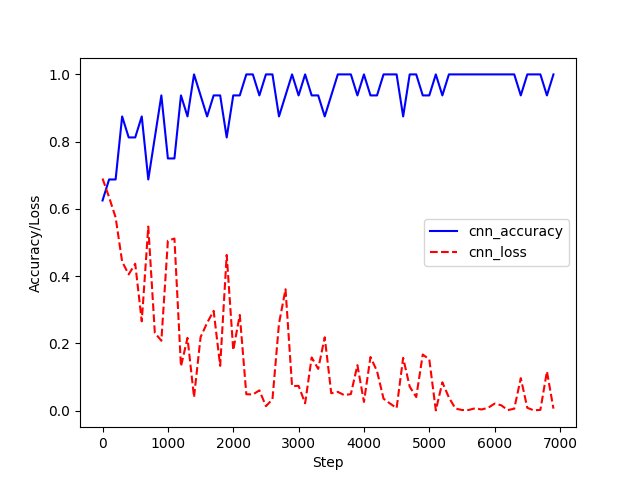
关键点检测任务：利用关键点样本进行训练。

1. CNN网络

CNN网络主要用于眨眼检测，眼睛睁闭是一个二分类问题，本系统使用CNN网络进行眨眼判断。网络结构由TensorFlow插件tensorboard生成，如下图：

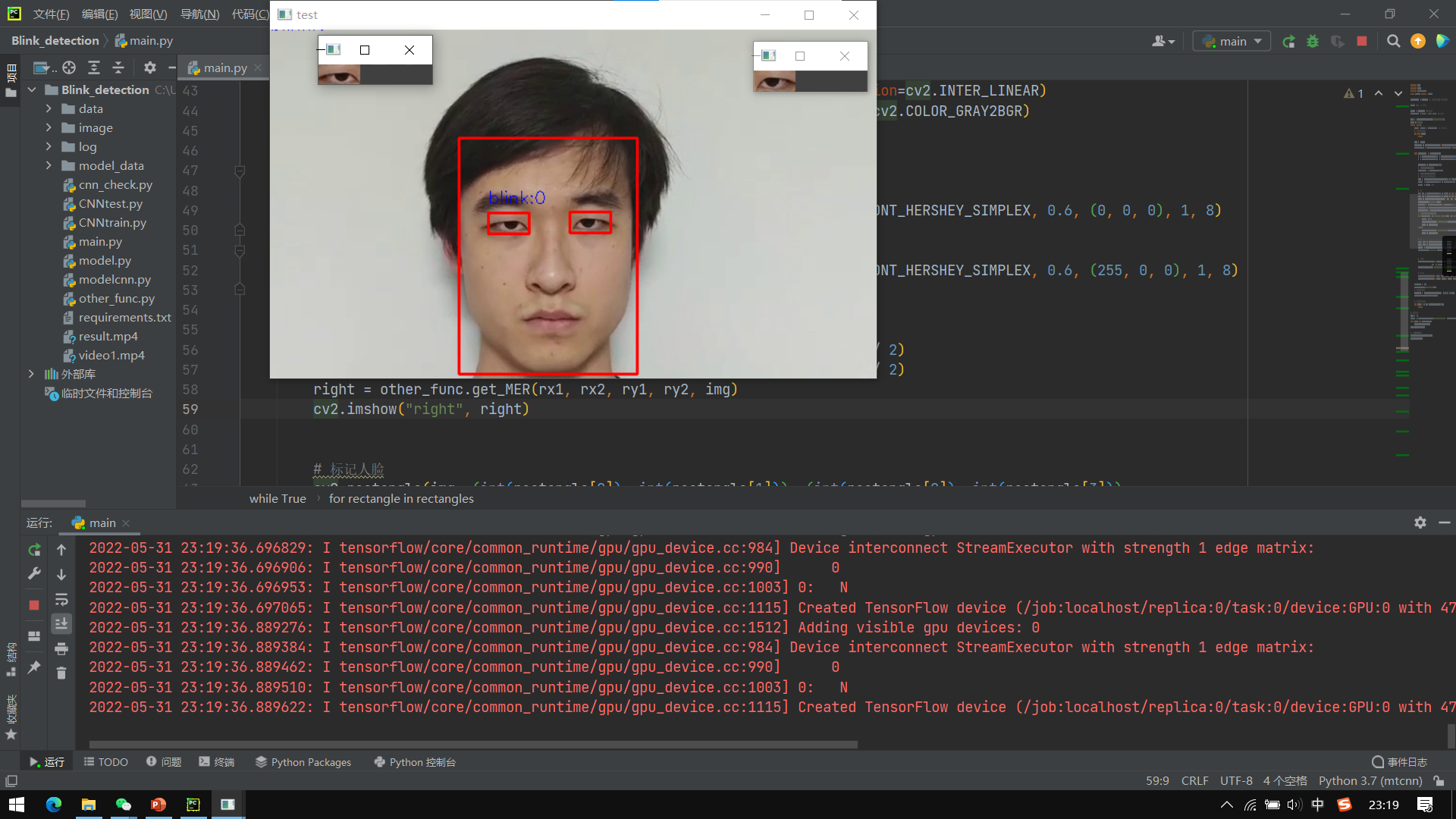
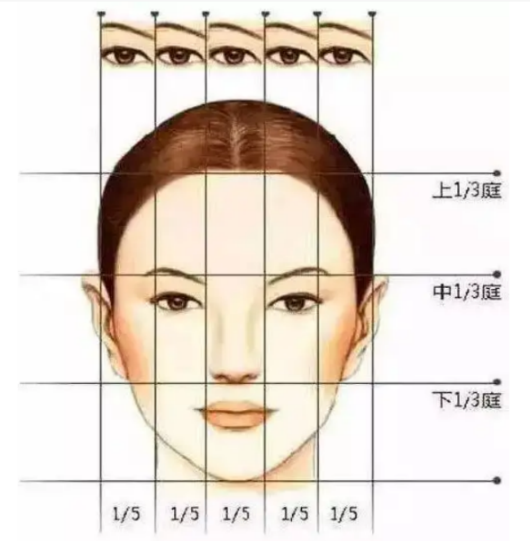


使用CEW训练集对CNN网络进行训练，训练后在4360张测试集图像中预测错误数量为303张，正确率为93.05%，训练损失函数图像如下：



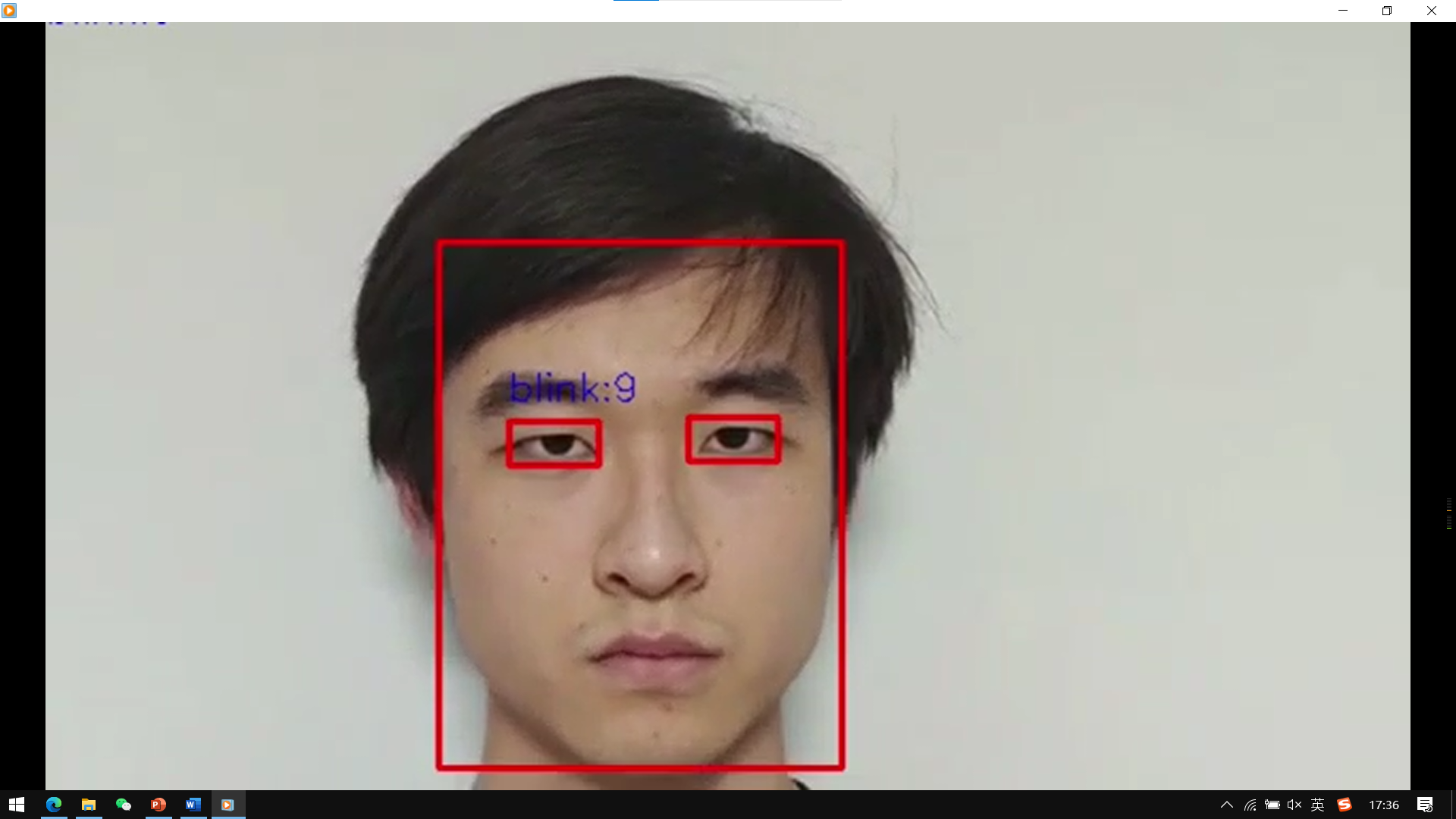
1. 网络连接方式

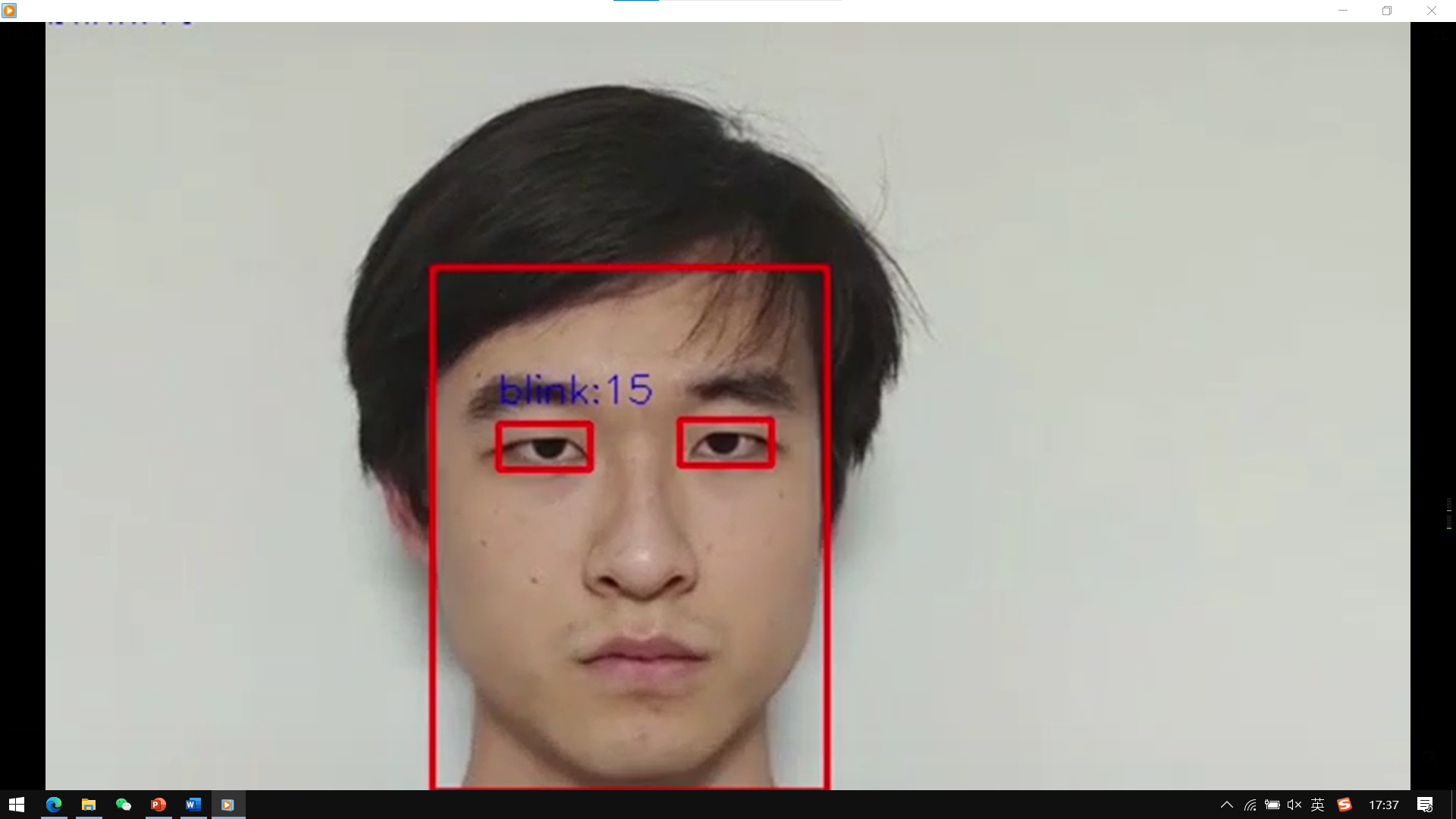
MTCNN网络输出结果为人脸位置框坐标以及人脸五关键点坐标，根据五点坐标中双眼坐标可以确定眼睛框选大小，按照眼睛选定框对眼睛进行截取，将眼睛图像传入CNN网络进行眨眼判定。由坐标点建立眼睛位置框的理论基础为“三庭五眼”理论，MTCNN网络检测到人脸的五个关键点，定义左、右眼中心点连线与水平方向的夹角为 θ，眼部区域宽度为W， 高度为 H=w/2，即可得到眼睛框的具体坐标。其理论图和效果图如下：



1. 系统运行结果

系统运行结果图如下：





获取驾驶员眨眼情况后，即可根据眨眼检测标准设定阈值，进行驾驶员疲劳度判定，根据判定结果输出提示语音。

1. 总结体会
2. 组员分工情况

本小组成员一共有四人，项目主要工作分为文献收集、数据集图像收集与整理、MTCNN网络构建、CNN网络构建，网络整合与测试。具体分工情况如下图所示：

|  |  |  |  |
| --- | --- | --- | --- |
| 姓名 | 工作比例 | 主要工作内容 | 备注 |
| 陈博楷 | 25% | MTCNN网络构建、系统测试 |  |
| 陈帅 | 25% | MTCNN网络构建 |  |
| 陈劲宇 | 25% | CNN网络构建 | 尝试使用Linux环境 |
| 陈鹏宇 | 25% | 文献收集、数据集收集、图像预处理 |  |

1. 学习体会

这次课程设计中，我们小组设计了一个驾驶员疲劳度检测系统的雏形，并且从零开始搭建了整个检测网络。在这个过程中我们学到了许多在课堂上没学到的知识和技术，总结一下有以下体会：

1、互联网很强大，在学习上将是一个非常高效的助手。我们所用的很多资料都能够在网上找到。从linux虚拟机的安装，linux的各种命令操作，到搭建Linux网络运行环境的详细解析，最后到MTCNN网络的基本原理，都能在网上找到。也因为这样，整个课程设计下来，我们才能解决大部分碰到的棘手问题。当然网上的东西很乱很杂，自己要能够学会筛选。对于有些问题，我们无法确定是否可行，这时就要去逐一尝试，而这又是非常要花时间的，因此多找参考，相互对比，才能提高我们的效率，事半功倍。

2、第二点是要注重小组间的讨论，这是很重要的。能够完成本次的课程设计，我们小组组员之间的讨论是不能忽略的。大家在研究问题时往往会出现不同想法，我们在有限地时间里能够尝试很多想法，完成系统的初步制作，离不开大家的共同努力，经过一番讨论之后，我们能够把思路理清楚，相互帮助，可以大大提高效率。

3、第三点就是编写系统的过程中要养成记录和备份的习惯。记录有助于对于不同版本的复盘，备份有助于在改进之后有对比，同时如果出现了运行或其他问题还能改回去，这是非常有用的。对系统的备份和记录不只应用在这次课程设计中，在未来的其他工作中也要时刻保持这一习惯。

当然，本次的课程设计还有一些可以改进的地方，但由于时间有限，我们组只进行了思路的谈论，希望以后可以改进。系统主要问题有两点，第一是增加CNN网络辨认的准确度，第二是增强系统的实时性。对于增加CNN网络的准确度，我们认为可以通过调整网络参数和修改网络结构实现，比如增加网络层数，换损失函数等方式；对于系统的实时性，则需要优化代码结构，或是更换部分网络结构，例如将网络的识别部分换成dilb识别，将其68关键点检测修改为只检测双眼，以此来降低dilb检测的时间。由于时间和设备问题，对MTCNN网络的训练过程我们小组只从训练方式上进行了解，实际操作过程中由于训练集数量过大，电脑出现内存溢出的情况，就没有完成MTCNN网络的训练。虽然这次课程设计已经接近结束，我们会在之后的时间里继续对它进行更多修改和优化，争取能够达到应用的水准。

总体来说，我们认为这是一次比较成功的课程设计，虽然不能说非常完美，但是我们已经尽了最大努力的去研究和学习。我们通过一步步的尝试，知道了神经网络的运行过程，大致了解了目前流行的视觉方面的优秀网络架构，我们相信这次动手实践一定会使我们今后的学习受益匪浅。