

数字图像处理

|  |  |
| --- | --- |
| 题目： | **面向家庭办公的工作状态图片分析** |

|  |  |
| --- | --- |
| 组 员： | 陈明亮 徐伟元 赵俊祥 |
| 学 号： | 200121817 2001212882 2001212891 |
| 院 系： | 信息工程学院 |
| 专 业： | 计算机应用技术 |
| 选题方向： | 面向家庭办公的工作状态  视频分析 |

二〇二一年六月

目录

[1 选题背景 3](#_Toc74041496)

[2 问题分析 3](#_Toc74041497)

[2.1 居家办公工作状态 3](#_Toc74041498)

[2.2 数据集 4](#_Toc74041499)

[3 解决方案 4](#_Toc74041500)

[3.1 基线方案 4](#_Toc74041501)

[3.2 改进方案 5](#_Toc74041502)

[4 实验 5](#_Toc74041503)

[4.1 实验目的 5](#_Toc74041504)

[4.2 实验过程 5](#_Toc74041505)

[5 总结 5](#_Toc74041506)

[6 参考文献 5](#_Toc74041507)

# 1 选题背景

近年来，受新冠病毒疫情与各地防疫政策影响，许多企业的员工进入了长时间居家线上办公的新模式。线上办公的模式，使得员工独处在家，缺少与同事之间的互动与沟通。部分员工可能因为这种枯燥的办公环境，降低工作效率。为了解决居家办公时的沟通问题，线上会议软件应运而生。该类型软件的一大特点是，通过线上视频的方式，提高员工之间的沟通效率，模拟日常会话情景。除职场员工外，学生群体也在这次疫情期间体验了线上课堂。同样地，依托于线上会议软件，学生与老师可以模拟线下面对面的教学场景。

虽然线上会议软件通过在线视频的方式，模拟了线下面对面的场景，但是其中存在一个问题，即视频中的人是否按照预期在进行合理的活动。如职场员工居家办公时，是否处于工作状态；学生居家上课时，是否处于上课状态。也就是说，居家办公/学习时，视频对象是否处于合理状态，是一个值得研究的问题。

# 2 问题分析

线上会议软件给出的是持续视频流，分析持续视频流中对象的工作状态对模型软件需要较高的实时性。但是，针对离线视频序列中对象的工作状态分析的模型复杂度也较高。因此，本文选择分析某一时刻下，视频帧内对象的工作状态。即，本文关注面向家庭办公的工作状态图片分析。本章接下来将介绍本文对工作状态定义以及本文所使用数据集。

## 2.1 居家办公工作状态

判定图片对象是否处于工作状态是一个二分类问题，属于粗粒度任务。对于二分类问题，简单的机器学习模型，如支持向量机（Support Vector Machine，SVM）也可以表现出优秀的性能。为了贴近办公与学习两个场景，并充分发挥深度网络模型的性能优势，本文将图片对象工作状态判断任务划分为更细粒度的分类任务，即针对图片对象，定义多个工作状态标签。针对办公与学习的两个场景，其非工作状态具有一定共性，因此可以定义出相同的细粒度非工作状态标签。具体而言，本文使用的工作状态标注信息有5类，分别为：睡觉（Sleeping），发呆（Dazing），玩手机（Playing Phone），读书（Reading Book），在电脑上工作（Working On Computer）。其中，前三个类别属于图片对象不处于工作状态时，其行为的更细粒度划分，而后两个类别则对应办公与学习状态。

## 2.2 数据集

线上办公对象工作状态的分析属于课程提出的新课题，因此无开源数据集或现有数据集。本文使用数据集属于自主收集数据。数据集中出现人物为本文作者。为保证模型的真实测试效果，本文提供500张图片作为测试数据。相对的，为确保模型不会欠拟合，本文提供X张图片作为模型训练数据。本文使用“训练-验证-测试”范式进行研究。因此除训练图片与测试图片外，额外提供X张图片作为验证集。表1给出了本文数据集的统计数据。

表格 1 数据集统计数据

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **训练集** | **验证集** | **测试集** | **合计** |
| **数量** |  |  |  |  |
| **比例** |  |  |  | **1** |

从表1可以看到，本文总共需要收集X张图片作为数据集。如果通过拍照获取数据，成本较高。因此，本文采取拍摄视频，从视频中截取视频帧的方式获取图片数据。即，针对2.1节定义的不同工作状态标签，被拍摄对象在不同居家场景中，做出相应动作或状态，拍摄者则以被拍摄对象为中心环绕被拍摄者拍摄视频。视频序列中的每一帧，都对应相应的工作状态。通过增加视频数量与视频时长方式，可以获取大量的标注视频。通过代码脚本从视频中截取视频帧，可快速获得大量标注图片。

利用Python的OpenCV库，可以读取视频并获取视频帧。通过控制截取视频帧的频率，可以从单个标注视频中获得大量的标注图片。

# 3 解决方案

## 3.1 基线方案

线上办公对象工作状态图片分析，本质属于针对图片的分类问题。因此本文选取图像处理领域中的基准特征提取器，卷积神经网络（Convolution Neural Network，CNN），对图片进行特征提取。针对提取的图片特征，使用线性层（Linear Layer）进行分类。即基线方案为CNN+Linear。

## 3.2 改进方案

# 4 实验

## 4.1 实验目的

## 4.2 实验过程

# 5 总结

# 6 参考文献

[1] D. Katabi and J. Wroclawski. A framework for scalable global IPanycast (GIA). In ACM SIGCOMM, 2000.

[2] C. Metz. IP anycast point-to-(any) point communication. volume 6, pages 94–98. IEEE, 2002.