人工智能与数据挖掘课程设计

报告题目

**组员信息：**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 编号 | 姓名 | 学号 | 成员贡献 | 实践成绩 |
| 1 | 陈昶源 | 2021211902 |  |  |
| 2 | 刘腾钰 | 2021211916 |  |  |
| 3 | 吴炫 | 2021211908 |  |  |
| 4 |  |  |  |  |

# **引言**

表情包是一种通过图像和文字来表达情感和态度的网络文化现象，它在社交媒体和聊天软件中广泛流行，具有传播性、趣味性和创造性。表情包的制作通常需要借助一些图像处理软件，如Photoshop、AI等，这些软件的操作较为复杂，需要一定的专业技能和美术基础。而Python是一种简单易学、功能强大的编程语言，它提供了一些图像处理库，如PIL、turtle等，可以实现一些基本的图像处理功能，如裁剪、缩放、旋转、合成、添加文字等。利用Python来制作表情包，不仅可以提高编程能力，还可以培养创造力和审美能力，同时也可以增加编程的乐趣和趣味性。此外，通过python还可以实现将人像提取出素描像，再以此为素材来实现表情包的制作，在本课程设计中，我们组以此为理论基础，将从百度图片网站中以不同表情的关键词来进行爬取素材，再通过对这些素材进行筛选、裁剪，得到大小一致的人像图。而后我们将对这些人像提取素描像，而后再实现图片的融合以及文字在图片中的定位等功能，来实现特定表情包的制作。

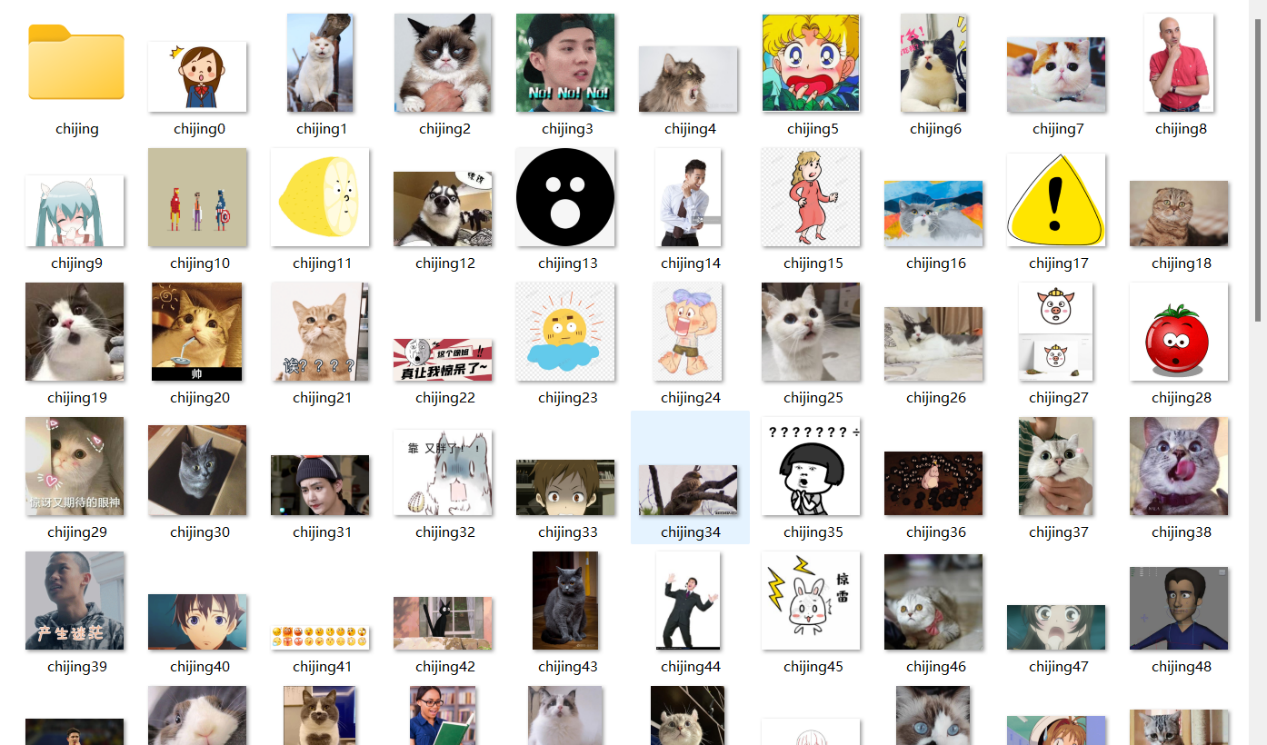
# **实验数据**

## 数据获取或预处理

## 首先没有任何数据的情况下，该如何开始数据的获取将是一个难题，我们组考虑到了两种方法，一种是自己采集，例如用摄像头抓帧，另一种是爬虫爬取数据。而在该实验中我们组将采用爬虫爬取百度图片中的数据。然后自己设计自己需要的表情，我自己设置了10类。这里以“吃惊”表情为例，在百度图片搜索中可以看到：

## 

而我们所需要的就是将这些图片提取并下载到本地文件夹，我们组针对此问题，使用了python，借此获取了初步数据：



从百度中获得的图片大小不一，因此需要进一步进行处理

## 2.1预处理

# 爬取完人脸表情之后，我们需要裁剪处图像中的人脸，这里设置裁剪大小为128\*128。裁剪过程需要用到opencv的人脸识别工具，关于该文件，可从opencv库的根目录中查找，经过以上几步，基本可以得到人脸数据集了，但是有的表情的数据量很少，进行模型训练的时候必然会因为数据不足带来影响，因此还需要进行数据增广处理。

# 一般地，数据增广处理的方法包括：旋转，镜像，随机裁剪，噪声，变形，颜色变化，对比度拉伸等方法。对于人脸来说，这里所选择的处理方法仅有：镜像，即左右反转；随机裁剪，将128\*128影响随机裁剪为120\*120，再将其resize成128\*128；噪声，添加少量的随机噪声。考虑到人脸数据集的特殊性，其他方法暂时没有选择。

# **方法/模型**

功能描述与实现方法介绍。

# **结果评测/展示**

方法评测或功能展示。

# **总结**

工作总结、局限性分析。