基于Neural Networks的京剧脸谱特征识别

1. 背景引入

京剧，是中国五大戏曲剧种之一，作为中国的国粹，位列中国戏曲三鼎甲的榜首。作为展现京剧特色的京剧脸谱，也随着京剧的发展一同变化，并展现出自己独特的魅力。脸谱最初的作用，只是夸大剧中角色的五官部位及面部纹理，用夸张的手法表现人物的性格、心理和生理上的特征，以此来为整个戏剧的情节服务，然而后来，脸谱由简到繁、由粗到细、由浅到深、由表及里，本身就逐渐成为一种具有民族特色的、以人的面部为表现手段的图案艺术了。

因此，在人脸特征识别发展如此迅速和完善的当今，考虑将特征识别应用于京剧脸谱就非常有价值和意义，不但能将较为成熟的技术予以新的应用层面，还能对传统文化加以深入了解和传扬。

首先，脸谱对色彩的应用就有讲究，一般是根据性格和性情为特殊人物所设计的。

红色有脸谱表示忠勇士义烈，如关羽、姜维、常遇春；



黑色的脸谱表示刚烈、正直、勇猛甚至鲁莽，如包拯、张飞、李逵等；



黄色的脸谱表示凶狠残暴，如宇文成都、典韦;



蓝色或绿色的脸谱表示一些粗豪暴躁的人物，如窦尔敦、马武等;



白色的脸谱一般代表坏人、奸臣，如曹操、赵高等；



除此之外，京剧还分为生旦净丑四个行当；京剧角色的分行十分细致入微，原本十个不同的行当，后来才归并为现在的四个大类。生是男子扮相，又可以细分为老生、小生、娃娃生及武生；京剧形成初期，老生承担着极为重要的角色责任。旦则指女子，可以细分为青衣、花旦、武旦、刀马旦、老旦等，上世纪40年代涌现出以梅艳芳为首的四大名旦使得旦角的地位发生了巨大的改变，并对京剧发展也造成了很大的影响；净用来代表性格刚烈或粗暴的男性；丑指滑稽人物，鼻梁上要抹白粉，称小丑或花脸。

1. 预想结果规划

最后的结果希望以微信小程序的形式呈现。程序的界面设计如下：





用户可以通过上传或者拍照提供一张清晰的包含京剧角色的照片

程序铺捉到图像中的人脸，根据人物的面部妆容特征或者借助人物服饰来判断人物的身份

程序对图片进行分类，增加判断标签以便用户更加清楚京剧的分类以及判断标准

初步成果未涉及前端的小程序页面设计。

基础功能希望可以根据上传的简单的面部脸谱数据，通过脸谱的主要色调如红色、黑色、黄色或白色来判断剧情人物的人物性格。

原本预期希望进一步的程序功能可以通过图片中人物的脸部妆效，脸谱的特征以及服装风格判断人物所处的戏曲情节以及人物的身份，但在做过调查之后发现，对于同一个人物有多段不同的戏曲故事，并且即使是相同的戏曲片段，也可能由于剧场或者演员的原因有不同的服装，面部特征等变化，因此目前无法对大量数据集做出相应处理以满足判断出此种细节的程度。

1. 主要技术分析

想要实现预期中的程序功能，需要用到深度学习以及神经网络的相关知识。其中，大量数据的训练集是必须的。通过利用高效高级的算法，让机器对训练集中已经做好标准的大量京剧图片进行学习和分析，从而可以对新导入的图片进行分析分类。因此，该程序所涉及到的主要就是对图片进行分类并加注标签的问题。

当前，图片分类作为计算机视觉的一个开放性问题被众多学者广泛研究，同时也是机器学习领域中极其热门以及有研究价值的方向之一。现如今对该问题的研究方法绝大多数都是受到了2013年由Geffory Hinton提出的Back Propagation的启发，并基于此研究并发展出的基于神经网络的深度学习框架。

其中，由Fei Fei Li于2006年提出的image net competition的比赛是一个提供数据测试集给世界各国学者测试自己算法模型优劣的公平公正的平台，吸引了众多机器学习(Machine Learning) & 计算机视觉(Computer Vision)学者的参与。

在这一过程中，算法由lenet, alexnet, resnet 到attention resnet的发展和变迁标志着这个问题在da大范围标注（large scale labels）和shallow resolutions的情况下已经得到基本的解决。

我们想做的京剧脸谱识别，虽然作为一个只有生旦净丑4个标签分类的高分辨率图像分类问题, 也应当可以从经典方法中学习汲取经验，采取相应的处理手段。我们尝试了多种算法，包括从traditional support vector machine, random forest到由神经网络组成的lenet, alexnet, transformers, ()，发现其中由state-of-the-art model () 所训练出来的分类器在测试集上达到了最好的p1分数的最佳的识别精确度。

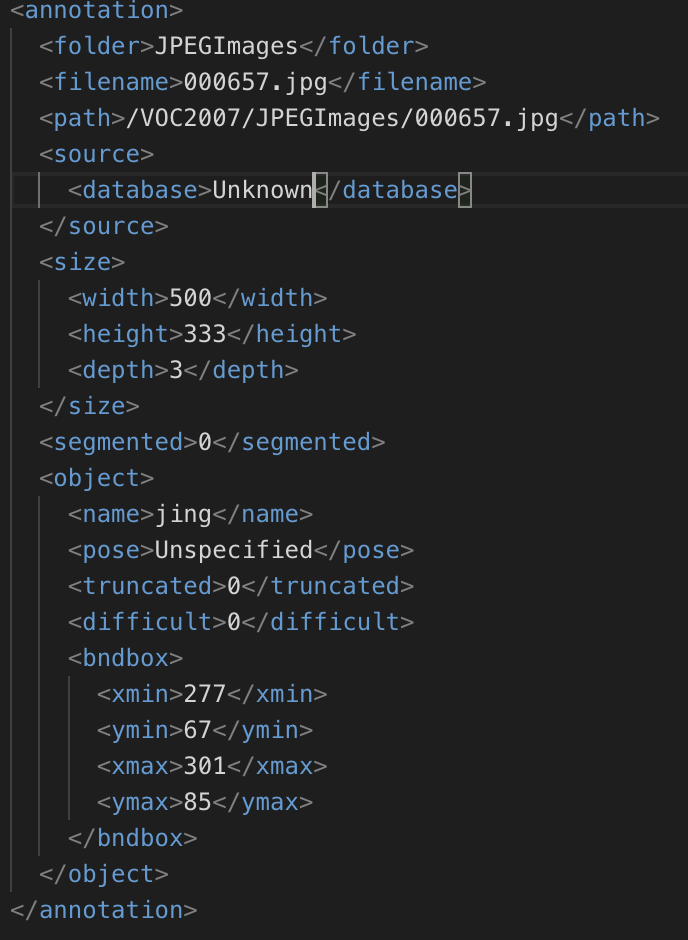
1. 目前的工作内容和进度

在开始着手做该项工作时，由于对利用机器学习实现的图片分类相关算法以及技术没有指向性和统一性地了解，因此考虑从获取系统的知识入手。

我用了两周时间来阅读和学习相关的文献，包括Collaborative Learning of Semi-Supervised Segmentation and Classification for Medical Images、Residual Attention Network for Image Classification和Generating Classification Weights with GNN Denoising Autoencoders for Few-Shot Learning，并最终决定使用state-of-the-art模型。

目前拟使用python来完成后端代码的编写和实现。

首先将获取到的大量图片一一对应进行了pickle标注，包括图片的长宽高等属性、以及图片内容的姿势、名称等属性，样例如下：

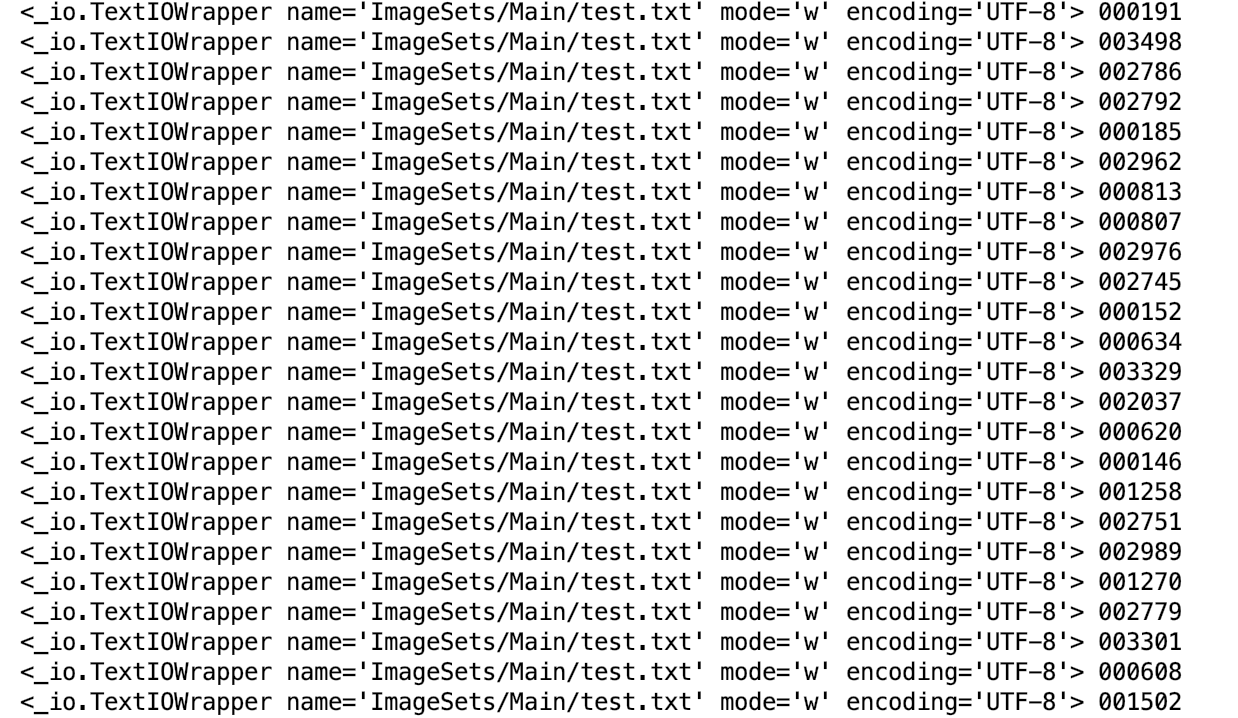


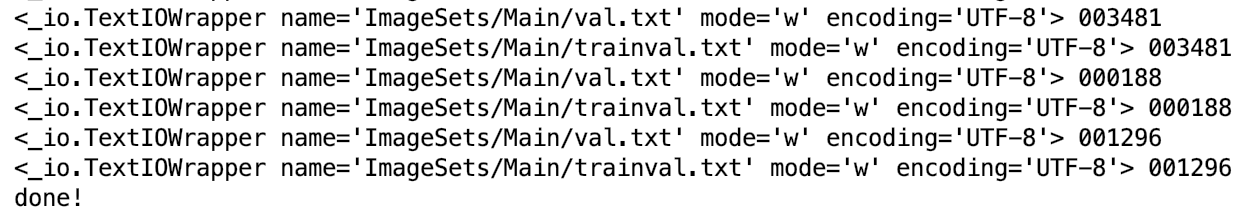
并生成相应的测试pkl文件

随后将数据分为训练集，测试集，测试验证集和验证集，分别将文件名放入对应的txt文件中，对数据集进行初步测评。



数据结果：





1. 未来的任务规划及安排

在未来的半个月左右，设想在保证准确度的情况下，首先满足后端代码的正确执行，即训练机器可以正确接收图片，并对图片进行最基本的分析分类。从颜色、服饰、装饰等多个方面判断，得出最后结论。

当保证后端代码正确性达到80%以上之后，在考虑将前后端连接，以小程序为基本载体，呈现多元化的中国传统文化色彩。