

集美大学诚毅学院大学生创新创业训练计划项目 结题报告

经费编码： CJ20030

项目名称： 基于卷积神经网络和迁移学习的面部表情识
别研究

项目类型：☒ 创新训练项目 项目等级：☐ 国家级
☐ 创业训练项目 ☒ 省级
☐ 创业实践项目 ☐ 院级

负责人姓名： 陈青芬

所在系、专业、班级： 信息工程系 计算 1892

项目起止时间： 2020.6-2021.6

指导教师： 武斌

指导老师联系电话： 15960290093

填表日期： 2020.6.13

集美大学诚毅学院

2013 年制

填写说明及注意事项

一、结题报告逐项认真填写，填写内容必须实事求是，表达明确严谨。

二、表格中的字体为小四号仿宋体，1.5 倍行距；需签字部分由相关人员以黑色钢笔或水笔签名。均用 A4 纸双面打印，于左侧装订成册。

三、报告书一式一份报送教务部，同时提交电子文档。

一、基本情况

项目名称		基于卷积神经网络的面部表情识别研究				项目级别	省级
立项时间		2020.6		完成时间	2021.4		
负责人	姓名	学号	专业班级	所在系	联系电话	E-mail	
	陈青芬	201841051044	计算1892	信息工程系	18959262480	2571788766@qq.com	
项目主要研究人员	杜晓岚	201841054103	软件1894	中软国际学院	13960399887	1370484311@qq.com	

二、研究成果简介

1、研究目的

该项目主要研究如何准确的识别面部表情所表达的情绪。实验先下载相关表情数据库，其中数据库中包含各种不同情绪的表情，后搭建神经网络，导入数据、优化权重进行训练，最终实现识别不同面部表情所表达的情绪。

2、研究内容

在过去的 20 年中，面部表情的自动识别已经取得了显著的进展。面部表情是人类传递情感状态和意图的最有力、最自然、最普遍的信号之一。由于自动面部表情分析在社交机器人、医疗、驾驶员疲劳监测和许多其他人机交互系统中的实际重要性，人们对它进行了大量的研究。在计算机视觉和机器学习领域，各种面部表情识别系统已经被探索用来从面部表情编码信息。面部表情识别中的表情被分为 7 类，包括快乐、惊讶、悲伤、愤怒、恐惧、厌恶和中性表情。每种情感都有自己的特征，都会通过一种面部表情反映出来。如果计算机等智能设备能够对这些特征进行学习，可以准确判断和理解人类的情感，再通过人类的情感去理解人类的内心想法，那么无障碍的人机交互的实现指日可待。

本项目研究了一种基于迁移学习用于表情识别的卷积神经网络，方法包括：（1）获取两个面部表情数据库，分为 A、B 两组，数据库中划分好训练集和测试集；（2）建立适用于迁移学习的卷积神经网络；（3）将 A、B 两个面部表情数据库分别在该卷积神经网络上进行训练、测试；（4）将训练所得的测试结果取准确率最高的权重参数进行保存；（5）将 A 数据库训练中得到的准确率最高的结果的权重参数迁移至 B 数据库的训练中，并按照步骤（3）进行新的一次训练，得到新的训练结果。同理，将 B 数据库训练中得到的准确率最高的结果的权重参数迁移至 A 数据库的训练中，并按照步骤（3）进行新的一次训练，得到新的训练结果；（6）与没有进行迁移学习的训练结果进行比较。

本项目实验使用 CK+ 数据库和 RAF-DB 数据库进行实验，选用 Resnet18 网络进行实验。实验研究发现将在 Resnet18 网络上测试训练后的、不同数据库的权重参数迁移至另外的数据库并再次进行训练测试后，实验结果的准确率有提升。

数据库	原实验结果	本研究实验的结果	前后结果对比
RAF-DB	76.09%	79.02%	+2.93%
CK+	98.76%	98.92%	+0.16%

4、成果的创新特色

该项目是基于机器学习与深度学习的基础上进行研究，如果实现后可以用于高速公路中所拍摄到的驾驶员的面部表情进而检测是否疲劳驾驶、通过课堂中学生的面部表情进而检测学生是否走神等情况。但是现在的面部情感识别还不够成熟存在缺陷，比如在自然场景下的表情识别率低，情感识别的准确度也不高。针对面部表情数据集图片数量少、依靠人工选取特征的传统面部表情识别精度不高的问题，我们提出了一种在卷积神经网络上融合迁移学习的办法，将人脸库上的预训练的权重用在面部情感识别上，可以有效提高面部表情情感的识别准确率，有利于实际运用。

迁移学习：在目标领域分类模型训练时就可以借助从源 领域数据和特征中已经提取的知识，实现已学习知识在相似或 相关领域间的复用和迁移，使传统的从零开始学习变成可积累 学习，不仅降低了模型训练的成本，而且可以显著提高机器学习的效果。

三、项目研究总结报告

一、项目研究和实践情况

在前期准备工作中，查阅了大量相关文献资料，学习了机器学习基础和深度学习基础内容，了解 VGG、Resnet、

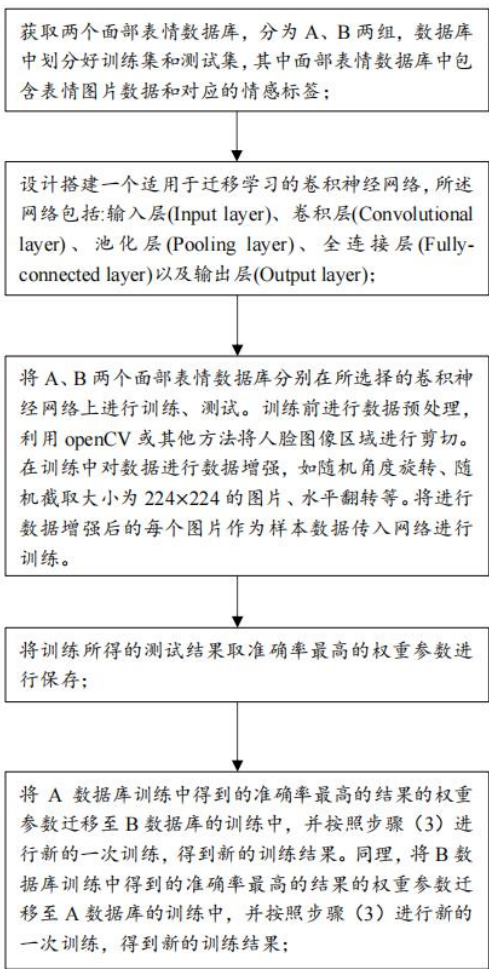
AlexNet、Googlenet 等网络及迁移学习相关知识；并安装所需的软件及搭建相关环境。随后开始尝试搭建神经网络，先进行简单的实验后应用到本项目中，逐步调节参数使神经网络的学习能力和测试后的准确率能够有所提高。

在此次项目进行过程中，通过近半年的学习与不断的实验测试，我们提高了分析问题、解决问题的能力，创新能力得到了锻炼。虽然完成了项目大部分功能，但仍有许多不足之处。

二、主要成绩和收获

1 实验过程

本项目实验过程如图



本项目研究了一种基于迁移学习用于表情识别的卷积神经网络，方法包括：（1）获取两个面部表情数据库，分为 A、B 两组，数据库中划分好训练集和测试集；（2）建立适用于迁移学习的卷积神经网络；（3）将 A、B 两个面部表情数据库分别在该卷积神经网络上进行训练、测试；（4）将训练所得的测试结果取准

准确率最高的权重参数进行保存；（5）将 A 数据库训练中得到的准确率最高的结果的权重参数迁移至 B 数据库的训练中，并按照步骤（3）进行新的一次训练，得到新的训练结果。同理，将 B 数据库训练中得到的准确率最高的结果的权重参数迁移至 A 数据库的训练中，并按照步骤（3）进行新的一次训练，得到新的训练结果；（6）与没有进行迁移学习的训练结果进行比较。

2 相关工作

2.1 卷积神经网络

CNN 的基本结构由输入层、卷积层 (convolutional layer)、池化层 (pooling layer, 也称为取样层)、全连接层及输出层构成。卷积层和池化层一般会取若干个, 采用卷积层和池化层交替设置, 即一个卷积层连接一个池化层, 池化层后再连接一个卷积层, 依此类推。由于卷积层中输出特征面的每个神经元与其输入进行局部连接, 并通过对应的连接权值与局部输入进行加权求和再加上偏置值, 得到该神经元输入值, 该过程等同于卷积过程, CNN 也由此而得名。Y. LeCun, L. Bottou, Y. Bengio, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.

2.2 网络模型

Resnet 是何凯明团队在 2015 年提出的一种深度卷积神经网络模型, 其基本模块如图。在 Resnet 提出之前, 训练非常深的网络模型存在一定困难, 因为随着网络深度的增加, 会存在造成梯度爆炸、消失, 导致系统不能收敛等问题。Resnet 与一般网络不同, Resnet 在 CNN 中引入跳跃结构进行跨层连接, 原始输入 x 通过跳跃连接与输出直接连接, 即不经过卷积神经网络直接送入到输出, 保留了原始数据, 在一定程度上避免了传统神经网络在信息传递时的丢失与损耗, 从而使得网络深度可以达到更深的层度。

2.3 迁移学习

迁移学习的定义是: “运用已存有的知识对不同但相关领域问题进行求解的多种机器学习方法”, 其目标是使知识在相关领域之间进行迁移。迁移学习在卷积神经网络中的作用就是将数据集在网络是训练所得的权重、参数等运用到新的领域中。

3 数据处理

3.1 RAF-DB

RAF-DB 是一个大型的面部表情数据库，从互联网上下载了大约 3 万张非常多样化的面部图像。基于众包标注，每张图片都有大约 40 个标注者独立标注。数据库中的图像在受试者的年龄、性别和种族、头部姿势、光照条件、遮挡(如眼镜、面部毛发或自遮挡)、后处理操作(如各种过滤器和特殊效果)等方面具有很大的可变性。raf-db 具有丰富的多样性、数量多、注释丰富等特点，包括:29672 张真实世界的图片，每个图像的 7 维表达式分布向量，两个不同的子集:单标签子集，包括 7 类基本情绪;双标签子集，包括 12 类复合情绪，5 个准确的地标位置，37 个自动地标位置，边界框，种族，年龄范围和性别属性每个图片注释，基线分类器输出基本情绪和复合情绪。本文聚焦研究的是单标签人脸表情，所以选用单标签子集。为了能够客观地度量关注者条目的性能，数据库被分为训练集和测试集，其中训练集的大小是测试集的五倍，两个集合中的表达式具有几乎相同的分布。

3.2 CK+

Cohn-Kanade 扩展数据集 (the extended Cohn-Kanade, CK+) [6]是目前比较通用的人脸表情数据集，适合于人脸表情识别的研究。 [1] Lucey P, Cohn J F, Kanade T, et al. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression[C]// Computer Vision & Pattern Recognition Workshops. IEEE, 2010.

CK+数据集为实验室环境下收集的数据集，该数据集包含八种人脸表情，分别为自然 (Neutral)、高兴 (Happy)、悲伤 (Sadness)、生气 (Anger)、恐惧 (Fear)、厌恶 (Disgust)、惊讶 (Surprise)、蔑视 (Contempt)。含有来自 123 个受试者的 593 个视频序列，仅从 106 个受试者中选择了 309 个序列，然后提取每个序列中表情强度达到峰值的最后 3 帧，得到 927 幅图，其中中性的表情数量为 0。为了使所使用的两个数据集具有相同标注标签，本文提取了每个序列表情强度最低的前 3 帧作为中性表情。最后，将 CK+分为 10 个子集，进行十折交叉验证。



CK+数据集部分图片

在 CK+数据集中已将表情按照 0-中性, 1-愤怒, 2-蔑视, 3-厌恶, 4-恐惧, 5-高兴, 6-悲伤, 7-惊讶进行标注。而 RAF-DB 数据集中不存在蔑视 (Contempt) 表情, 所以本文中的 CK+数据集将蔑视表情删除, 并按照 RAF-DB 数据集中表情标注的 label 对应数字进行修改, 使两个数据集的标注标签相同, 便于迁移学习。

4 实验方法

4.1 图像预处理

为了让模型更加关注人脸区域, 而不是去学习与面部变化无关的背景部分, 本文利用 face++接口, 返回 4 个人脸框坐标, 分别为人脸框左上角纵坐标 (top), 左上角横坐标 (left), 人脸框宽度 (width), 人脸框高度 (height)。再利用 openCV 进行剪裁。剪裁前后图片如图



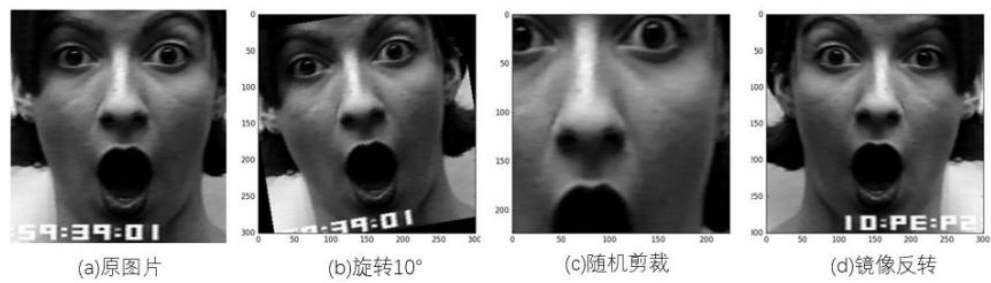
剪切前后对比图片

4.2 训练模型

本实验在 Resnet18 网络的基础上加上迁移学习，在训练心始前先加载 Resnet18 模型，再将先前训练好的准确率高的模型参数文件迁移到即将要训练的模型上，盼后再往始进行训练。训练过程中，学习率的值设置为 0.001，batch 设置为 256，epoch 设置为 50。

4.3 数据增强

神经网络需要大量数据进行学习，而且许多大型神经网络存在过拟合的问题，所以本文采用数据增强的方式来为神经网络提供更多图片数据进行训练。数据增强是通过剪切、翻转、随机剪裁等方式实现。本文方法采用在原图片基础上在 $(-10, 10)$ 度上进行随机旋转，在原图片基础上随机剪裁大小为 224×224 的图片及在原图片基础上进行水平翻转。



数据增强前后对比

5 实验结果

5.1 原始结果

Resnet18 网络训练得到到的构好结果分别是：RAF-DB 数据库准确率为 76.09%，CK+数据集准确率为 98.76%。RAF-DB 数据库的数据是真实场景下的面部表情，CK+数据集中的数据是实验室场景下的面部表情。所以 CK+数据集的准确率要高于 RAF-DB 数据库。实验原始数据如图表所示。

数据库	准确率
RAF-DB	76.09%
CK+	98.76%

原始结果

5.2 加入迁移学习后的实验结果

本实验采用在原网络模型的基础上加入迁移学习进行训练,该方法是将每个数据库在 Resnet 网络下训练得到的权重参数迁移到其他的数据库即将进行的训练中。本文实验将 CK+数据集在 Resnet18 上训练所得的准确率高的权重参数迁移至 RAF-DB 数据库的训练中,迁移后的实验结果比原始结果提高了 2.93%;将 RAF-Db 在 Resnet18 上训练所得的准确率高的权重参数迁移至 CK+数据集的训练中,迁移后的实验结果比原始结果提高了 0.16%。对比两次实验结果,证明了本文所使用的加入迁移学习的方法可以使训练得到丰定的提升。(注:由于 CK+数据库原始准确率已经超过 98%,所以迁移后结果无法大幅度提升)

数据库	准确率	进行迁移学习后的准确率	实验结果对比
RAF-DB	76.09%	79.02%	+2.93%
CK+	98.76%	98.92%	+0.16%

加入迁移学习后的结果

四、项目预期成果的实现情况

项目预期成果:

一套算法
一篇论文

项目取得成果:

一套算法
一篇论文

五、经费使用情况

经费合计 10000 元。

其中,国家资助 0 元,省级资助 10000 元,学院配套资助 0 元,其他经费 0 元。

经费支出情况:

租用服务器: 3354 元

各种硬件设备: 1708.15 元

U 盘: 249 元

软件服务费: 198 元

共 5509 元

六、指导教师及系审核意见

项目指导教师对结题的意见，包括对项目研究工作和研究成果的评价等。

该项目以卷积神经网络为基础，设计了一套可用于人脸面部表情识别的算法。该设计能对人脸图片及对应的标签进行学习，最后能识别人脸面部表情所表达的情绪，具有一定的实践意义。该项目不仅锻炼了学生的学术实践能力，还提升了学生的创新能力。项目组的成员从选题、开题、实施、研究到结题这一过程能够将自己的情况反馈给指导老师，小组成员之间积极配合，能较好地合作分工，基本达到研究目的，取得了较好的研究成果。

综上，该课题已完成既定的要求，同意该课题结项。

指导老师签章：

2021年6月15日

项目主持人所在系对结题的意见，包括对项目研究工作和研究成果的评价等

该项目设计了一套可用于人脸面部表情识别的算法，已完成项目的研究任务，达到既定的目标，同意结题。

负责人签章：

2021年6月15日

七、学院结题审核意见

同意结题

负责人签章：

2021年6月15日