**浙江理工大学2019年“挑战杯”大学生**

**课外学术科技作品竞赛作品申报书**

作品名称： 关于黑白图像预测实际色彩的研究

负责人姓名： 李智敏

联系方式： 15868159657

团队成员： 陆宇豪、舒丹璐、吴联磊、蔡淼、王秀仪、

沈红莉、杨思誉

指导教师： 冉宇瑶、刘海霞、铁治欣

作品所属领域： B．信息技术

□个人项目 ☑集体项目

类别：

□自然科学类学术论文

□哲学社会科学类社会调查报告和学术论文

☑科技发明制作A类

□科技发明制作B类

二○一八年九月

**说 明**

1、申报者应在认真阅读此说明各项内容后按要求详细填写。

2、申报者在填写申报作品情况时注明个人项目或集体项目。根据作品类别（自然科学类学术论文、科技发明制作A类、科技发明制作B类），分别填写作品所属领域包括：A．机械与控制（包括机械、仪器仪表、自动化控制、工程、交通、建筑等）；B．信息技术（包括计算机、电信、通讯、电子等）；C．数理（包括数学、物理、地球与空间科学等）；D．生命科学（包括生物、农学、药学、医学、健康、卫生、食品等）；E．能源化工（包括能源、材料、石油、化学、化工、生态、环保等）。根据作品类别（哲学社会科学类社会调查报告和学术论文），填写作品所属领域包括：A．哲学；B．经济；C．社会；D．法律；E．教育；F．管理。

3、作品所属领域请按作品的学术方向或所涉及的主要学科领域填写。

4、表内项目填写时一律用钢笔或打印，字迹要端正、清楚，此申报书可复制。

5、有关参赛事宜请向竞赛组委会秘书处咨询。

**浙江理工大学2019年“挑战杯”大学生课外学术**

**科技作品竞赛立项申报表**

申报单位：科技与艺术学院 填报日期：2018年9月9日

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 作品名称 | 关于黑白图像预测实际色彩的研究 | | | |
| 团队名称 | 缤纷机器学习团队 | | | |
| 团队情况 | 姓 名 | 性 别 | 班 级 | 手 机 |
| 负责人 | 李智敏 | 男 | 16计算机1班 | 15868159657 |
| 团队成员 | 陆宇豪 | 男 | 16计算机2班 | 15868168532 |
| 舒丹璐 | 女 | 16工商1班 | 15867192179 |
| 吴联磊 | 男 | 15计算机2班 | 15958006911 |
| 蔡淼 | 女 | 16人力1班 | 15867168345 |
| 王秀仪 | 女 | 16计算机2班 | 13857591503 |
| 沈红莉 | 女 | 16自动化1班 | 15868132611 |
| 杨思誉 | 男 | 16计算机1班 | 17348846507 |
| 指导老师 | 姓名 | 专业方向 | 职务（职称） | 联系电话 |
| 铁治欣 | 数据挖掘 | 教授 | 0575-82978008 |
| 冉宇瑶 | 计算机应用 | 副教授 | 13858006561 |
| 刘海霞 | 计算机应用 | 讲师 | 15868483883 |
| 项目  简介 | (包括选题背景、意义，研究内容、思路、方法，研究的重点、难点、创新点，预期目标，研究成果，经费预算等。不少于3000字，需另附纸) | | | |
| 立项作品关键问题  （不超过300字） | 在过去的算法中，主要通过传统的机器学习技术，他们用“回归”来预测每一个像素的颜色，有文献作者使用均方误差作为训练模型的损失函数。MSE将尝试“平均”出颜色，以获得最小的平均误差，这将导致平淡的外观(主要是偏暗黄色)。因此目前没有该项目较完美通用的解决方案，知网万方等数据库中相关论文也只有缪缪数篇，这项任务是非常具有挑战性的。  机器学习需要掌握C++，python等开发语言，需要掌握流行的各类框架技术，如tf，caffe。一个完整的项目需要web开发，数据库管理，Linux运维等。同时对算法编写者的代码调试能力有非常高的要求，以及较好的数学功底。 | | | |
| 指导教师意 见 |  | | | |
| 学院推荐意 见 | 公章  年 月 日 | | | |
| 竞赛评审委员会  审核意见 | 公章：  年 月 日 | | | |

本表一式二份，一份上交竞赛评审委员会，一份留各学院

**项目简介**

# 选题背景

黑白照片上色是进度极慢且复杂的工程，专业摄影师或者艺术家要花几个小时，甚至几天才能修复一张因泛黄丢失色彩，或原始是黑白的老照片，更不要说是视频了。这个过程需要花费极大的精力和高昂的费用。许多人都有需要修复的历史照片和视频，但只有极少人能承担雇佣专业人士来还原照片的费用，同时还伴有隐私泄露的风险。

对大部分人而言，黑白摄影似乎是上个世纪的事了，认为以前的人拍黑白照片是因为当时彩色胶卷没有发明出来。但事实上，即便是到了现在的数码时代，黑白图像从来没有退出历史舞台，而其地位也从来没有因为彩色胶卷、数码技术受到影响。故我们开发此项技术来对黑白图像进行快速修复。

1. **黑白照片的来源**
2. 不支持彩色。因彩色摄像头无法承受极限温度，如钢铁厂就采用热力探测(如红外技术)，需要以热力图的形式将其可视化。
3. 宇宙深空探测，拍摄光谱不在可见光范围内，需要通过彩色修复。
4. 生物细胞结构，如蛋白质结构等，设备读出的光谱转换成黑白照片之后需要根据专家经验上色。
5. **在特定领域彩色成本高昂**

夜间真彩色摄像头成本通常比黑白摄像头贵三倍。彩色需要存储多个通道信息，现今2K,4K技术对高密度存储技术提出了新一轮的挑战。在保证画质的情况下，有损的色彩通道转换压缩技术不可或缺。

1. **原本的色彩无法表达对应的情绪，需要更改**
2. 照片因为色彩的干扰，照片所要表达的意图被模糊。比如一张表达恐怖情绪的照片，主题色调应该设定为深红色(血腥的暗示)。比如一张表达温馨的家庭环境，我们就建议以米黄色作为主色调表达温馨的情绪。
3. 色彩会给人预设一种主观偏见，画面的色彩会让人更多去关注画面本身的形状、造型、事件和元素，所表达或看到的主题也会变得更加清晰。
4. **色彩缺失**

很多人家中都有需要被修复的老照片，博物馆拥有大量包含极高价值的老照片，比如一位妇人身着清末彩色衣服，专家知道其对应颜色，却只能用文字“洋红殷红”等描述，不直观也容易在文化传承中失去可证性。

# 意义

得益于科技的发展，依托于海量GPUs并行计算硬件加上设计良好的框架，如TensorFlow，caffe，mxnet，我们可以利用**深度学习技术(Deep Learning)**使**黑白图像在保证语义正确(Semantic correctness)的情况下将其还原为彩色**，且这个过程仅需几毫秒，成本低廉。

1. 具有历史意义的照片应该被转换为彩色，使其保留重要的历史价值。
2. 智能手机双摄像头中的黑白的长焦镜头也需要一种高效的算法将其还原为认知彩色，降低黑暗中的噪点。
3. 有损图片数据压缩，夜间红外摄像头还原真彩色，视频修复等专业领域提供新的思路和实践方法。
4. 分支训练，根据文物的形状无监督学习聚类并推测色彩，支持全局提示网络。

# 研究内容

给出无类标的黑白图片或视频，在保证语义正确的情况下，利用计算机视觉技术将其转换为彩色。

随着智能手机双摄像头(黑白+彩色)的方案使用越来越广泛，黑白的长焦镜头也需要一种高效的算法将其还原为认知彩色，降低黑暗中的噪点。同时，本项目还可以为有损图片数据压缩(Lossy picture data compression)，夜间红外摄像头还原真彩色，视频修复等专业领域提供新的思路和实践方法。许多具有历史意义的照片应该被转换为彩色，使其保留重要的历史价值。为物理，化学，生物，天文学，考古学中的黑白照片快速上色。

许多黑白的照片的电子扫描件，通过人工智能(Artificial intelligence)相关的算法，将其为人类可直觉认知的彩色图片。

# 研究思路

在过去的算法中，主要通过传统的机器学习技术(machine learning，ML)，他们用“回归”来预测每一个像素的颜色，然而结果却不尽如人意，根据黑白图片渲染出来的图像色彩平淡枯燥。有文献作者使用均方误差(Mean Squared Error ,MSE)作为训练模型的损失函数(Loss Function)，作者指出，MSE将尝试“平均”出颜色，以获得最小的平均误差，这将导致平淡的外观(主要是偏暗黄色)。因此目前没有该项目较完美通用的解决方案，知网万方等数据库中相关论文也只有缪缪数篇。

我们将设计不同的loss函数作为优化目标，对比不同loss在训练之后的结果。具体指标包括颜色还原准确率，性能，通用性等。之后我们将使用集成学习技术对各个模型预测的结果进行加权平均得到最后彩色图片。

小组内成员将进行一周两次的研讨，并对各个阶段的学习内容进行总结，交流学习心得。考虑经费预算，商讨并整理思路。与此同时我们将进行以下任务:数据集的下载及清理，编写模型代码。进行调试工作，进行模型优化以提高速度。进行网页制作，进行视频制作。

**对于本项目，目前我们已经有1篇关于自然光线下阴影处理技术的论文被省级刊物正式录用。**

# 研究方法

深入学习掌握机器学习算法(machine learning algorithm)和深度学习技术(deep learning technology)；使用Ubuntu16.04\_Server系统，并预装深度学习框架，方便此后代码的编写与测试。

随后确定统一的技术选型和编码风格，我们团队内有使用caffe和tensorflow的同学，但我们最后选择以tensorflow作为机器学习框架。

同时我们将获得来自老师们的耐心指导与帮助。**机电系系主任冉宇瑶老师和刘海霞老师**将对本项目进行耐心的指导

1. 在项目内容确定后，对项目的细节内容进行指导调整
2. 引导小组成员的具体学习方向并给予一定的协助
3. 在项目的测试阶段，提出项目中存在的问题，帮助我们项目更加完善。
4. 帮助我们进行算法的完善和论文的润色修改。

**浙江理工大学科技与艺术学院铁治欣副院长**研究**数据挖掘方向**，现正研究**计算机视觉方向**，将对本项目中遇到的问题，以及**论文和申报材料进行指导与修改。**

# 项目重难点

在过去的算法中，主要通过传统的机器学习技术(machine learning，ML)，他们用“回归”来预测每一个像素的颜色，然而结果却不尽如人意，根据黑白图片渲染出来的图像色彩平淡枯燥。有文献作者使用**均方误差(Mean Squared Error ,MSE)作为训练模型的损失函数(Loss Function)**，作者指出，MSE将尝试“平均”出颜色，以获得**最小的平均误差，这将导致平淡的外观**(主要是偏暗黄色)。因此目前没有该项目较完美通用的解决方案，知网万方等数据库中相关论文也只有缪缪数篇。

从认知的角度出发，分辨一个物体的色彩并不是一件难事。但假如给定的一朵花呢？这时候就会产生不一样的结果，一部分人认为花是黄的，另一部分人认为是红的，这就产生了**认知歧义(Cognitive Ambiguity)**。现阶段需要做的就是使计算机能够尽可能的还原事物本来的颜色，并将其运用在老照片颜色还原，智能手机摄像头拍摄照片颜色的增强，夜间红外摄像头拍摄的黑白视频色彩修复任务等。要完成这个项目，还要使用yolo, imageNet这些框架来感知物体类别，总的来说，这项任务是**非常具有挑战性的**。

**具体难点如下:**

1. **数据扩增**

在灰度图像复原成彩色RGB图像方面的代表性文章《全局和局部图像的联合端到端学习图像自动着色并且同时进行分类》中指出。稍作解释，黑白图像，实际上只有一个通道的信息，即灰度信息。彩色图像，则为RGB图像(也可以是CMYK)，有三个通道的信息。彩色图像转换为黑白图像极其简单，属于有损压缩数据；反之则很难，因为数据不会凭空增多。

1. **概率预测**

在待处理信息量不可扩充的前提下（即模糊的图像本身就未包含场景中的细节信息），可以借助海量的同类数据或相似数据训练一个神经网络，然后让神经网络获得对图像内容进行理解、判断和预测的功能，这时候，再把待处理的模糊图像输入，神经网络就会自动为其添加细节，尽管这种添加仅仅是一种概率层面的预测，并非一定准确。

1. **学习门槛高**

机器学习需要掌握C++，python等开发语言，需要掌握流行的各类框架技术，如tf，caffe。一个完整的项目需要web开发，数据库管理，Linux运维等。同时对算法编写者的代码调试能力有非常高的要求，以及较好的数学功底。

# 项目创新

我们将使用传统的机器学习算法(machine learning algorithm)和近几年才产生的深度学习技术(deep learning technology)来完成这一目标，具体内容包括:

1. **传统的机器学习算法(machine learning algorithm)**
2. 线性回归 (Linear Regression)
3. 逻辑回归 (Logistic Regression)
4. 贝叶斯分类 (Bayes Classification)
5. 决策树与随机森林 (Decision Tree)
6. k最近邻法 (KNN)
7. AdaBoost迭代
8. **无监督深度学习技术 (deep learning technology)**
9. 分别采用 Tensorflow,Caffe 训练两个模型
10. 使用集成学习技术
11. k折交叉验证
12. 运用遗传算法提高整个模型最后的准确度

# 预期目标

我们的项目前后端分类，主要借助服务器进行发布测试、运行。

前端部分，我们采用flask框架搭建，使用bootstrap，jquery进行页面的美化工作。

后端我们将使用MySQL/MariaDB管理数据库，使用Python搭建服务端，并采用tensorflow框架进行计算，将消息队列作为消息中间件。

当用户通过前端请求该功能时，后端会打包一个带鉴权的API供用户使用，此时前端和客户端请求API，读取并解析json中的地址并由前端加载处理完成的图片。如果存在需要修正的部分，可以通过全局提示网络进行进一步的修复，从而达到用户预期。

**项目发布模式**

计费形式为服务后付费模式。服务按照每月（自然月）识别总量采取阶梯到达的计费方式，当月总量在哪个阶梯内则按该阶梯单价计费，识别量越大即单价越低。

1. 在线调用API报价：

5万张及以下 0.0015元/张

5万张以上 0.0011元/张

1. 流量价格计算：

内网流入流出流量免费

公网流入流量免费

公网流出忙时0.5元/GB(8时-24时) 闲时0.25元/GB（0时-8时）

API的流控10次/秒/API/用户，每个用户每个API每秒限制调10次。公网流出流量按量计费，不限制最大带宽。

**对图像中物品的形状和大小要求：**

数据将会被转换为299x299的BGR图片，物品的尺寸建议大于 64\*64 像素，且图像清晰识别精度更佳。

**对图像格式的限制：**

目前支持JPEG、JPG、BMP、PNG、TIF、PBM等常见图像格式。

**对图片大小或请求内容是否有限制：**

图片大小要求5\*5像素以上，最大尺寸受传输body内容尺寸限制，请求body内容小于2M。

# 研究成果

1. **结题时提供成果如下**：

训练集全部图片资源，全部源代码以及训练的模型，展示网页和视频。

1. 训练集包括imageNet图片库资源，ins图片库资源。
2. 源代码包括dataset预处理部分，数据集清洗，tensorflow模型源代码，前端Flask框架源代码，API源代码，后端Kalfa或者RocketMQ编写的源代码，python微服务代码等。
3. 模型包括Auto Encoder预测模型，全局提示网络模型。
4. 网页包括上传页面和答辩页面。
5. 视频包括项目介绍，推广和展示等。
6. **项目预期成果**
7. 研究报告一份（包含实验记录、设计方案、tf计算模型等项目所产生的研究成果）
8. 论文1篇
9. 软件著作权1项
10. 详细的软件说明文档，包括安装环境，运行，修改编译等
11. 视频若干个 分别用于展示，答辩，项目推广等
12. 网页2个 用于在线API调用，项目展示

# 经费预算

**本项目将获得浙江理工大学科技与艺术学院特批资金约伍万元用于购买搭建深度学习环境所需的硬件环境，绍兴市科技创新项目立项资金赞助若干。**

**1.费用开支明细**

|  |  |
| --- | --- |
| **项目** | **金额** |
| 升级服务器显卡 | 3000元 |
| 升级服务器内存容量 | 1000元 |
| 论文版面费用 | 1000 |
| 软件著作权申请费用 | 1000 |
| 购买学习相关书籍(后附有详单) | 600元 |
| 前期中期后期相关材料打印复印 | 200元 |
| 文具类相关购买，笔记本，A4纸，笔等 | 200元 |
| **合计：** | 7000元 |

具体开发测试环境如下:

（1） 系统版本 Ubuntu16.04\_Server

（2） CPU: E5-2680 X 2 计 20C40T

（3） RAM: ECC 8G X 8 计 64GB

（4） SSD: Intel 256G + 三星256G

（5） 显卡: NVIDIA 1070 8G

（6） 预装深度学习框架如下：

cuda9, cudnn7, tensorflow1.8.0, tensorboard, jupyter, theano, chainer, Cython, graphviz, pydot, opencv-python, xgboost, cntk-gpu, mxnet-cu90, torch, torchvision, lightgbm, keras, tqdm, scipy, matplotlib, sklearn

**2.书目**

**所需书目如下，部分书目因图书馆已有馆藏将不再购买。**

|  |  |
| --- | --- |
| **书名(已馆藏部分\*，加粗为重要)** | **题名责任者** |
| **统计学习方法\*** | **李航** |
| 统计学习基础 | Trevor Hastie 等 |
| **机器学习(西瓜书)** | **周志华** |
| **模式分类** | **Richard O. Duda 等** |
| **Pattern Recognition and Machine Learning** | **Christopher Bishop** |
| 机器学习 | Tom Mitchell |
| **深度学习(花书)** | **Ian Goodfellow** |
| **数据挖掘导论** | **Pang-Ning Tan** |
| jQuery基础教程\* | (美)Jonathan Chaffer  (美)Karl Swedberg著 |
| 云存储技术:分析与实践:analysis and practice | 刘洋著 |
| Python机器学习\* | (美)塞巴斯蒂安·拉施卡(Sebastian Raschka)著 |
| 算法之美:隐匿在数据结构背后的原理:C++版\* | 左飞著 |
| Java虚拟机精讲\* | 高翔龙编著 |
| Python机器学习:预测分析核心算法 | (美)Michael Bowles著 |
| 机器学习 | (英)Peter Flach著 |
| 人工智能.3版\* | 朱福喜编著 |
| **机器学习导论** | **(美)米罗斯拉夫·库巴特(Miroslav Kubat)** |
| Ubuntu Linux操作系统\* | 张金石主编 |
| 函数式编程思维 | (美)Neal Ford著 |
| 重构:改善既有代码的设计\* | (美)Martin Fowler著 |
| 数据结构编程实验\* | 吴永辉，王建德编著 |
| 编写可读代码的艺术 | Dustin Boswell，Trevor Foucher著 |
| Python语言及其应用\* | (美)Bill Lubanovic著 |