

# 浙江大学校级大学生创新创业训练计划

## 中期检查表

项目编号： X20200292

项目名称： 水印检测与去除

项目负责： 徐晓丹

学号： 3180103422

院（系）： 计算机科学与技术学院

联系电话： 18069959665

电子邮件： 2682724004@qq. co

指导教师： 宋明黎

职 称： 教授

浙江大学本科生院教务处

2020年11月18日

|      |            |                      |                |                       |                   |
|------|------------|----------------------|----------------|-----------------------|-------------------|
| 项目名称 |            | 水印检测与去除              |                |                       |                   |
| 立项经费 |            | 1000                 | 起止时间           | 2020-03-03至2020-05-31 |                   |
| 负责人  | 学号         | 姓名                   | 所在院系、专业        | 联系电话                  | E-mail            |
|      | 3180103422 | 徐晓丹                  | 计算机科学与技术学院、混合班 | 18069959665           | 2682724004@qq.com |
| 参加成员 | 3180104396 | 周寒靖                  | 计算机科学与技术学院、混合班 | 15306591758           | 15889652@qq.com   |
| 参加成员 | 3180103781 | 茅一宁                  | 竺可桢学院、混合班      | 18888910511           | 1477045246@qq.com |
| 导师   | 姓名         | 宋明黎                  | 院系:            | 职称                    | 教授                |
|      | E-mail     | brooksong@zju.edu.cn |                | 联系电话                  | 13750888255       |

## 一、项目研究进展情况（含项目研究已取得阶段性成果和收获）（800字内）

### （1）项目研究进展情况

项目通过GitHub仓库来管理涉及的所有代码文件及其他资料，通过“一周一会”来保持项目组成员的进度同步，讨论解决疑问，并及时记录未能解决的问题。

项目起步阶段，经过与导师及实验室学长会谈、广泛查阅去水印技术相关论文等过程，确立课题的内容为“水印检测与去除”，选定谷歌的论文《On the Effectiveness of Visible Watermarks》作为初步研究对象；目标是对论文中提及的“水印检测与去除”算法进行复现、改进和创新，并考虑将该算法应用于实际场景中，并以产品或论文的形式呈现研究结果。

项目准备阶段，认真阅读论文原文，爬取网页中的图片、下载coco数据集并构建实验所需的图像集，为后续研究打下基础。使用anaconda安装相关库（opencv、scipy等），搭建算法运行所需的python环境。用python编写了图片爬取及批量加水印的程序，构建了150张图片的固定位置水印图像集、150张图片的随机位置水印图像集，图像的大小均为500\*500，格式为png；使用coco数据集，将图片数量扩展至3000张。此外，还构建了不透明水印的测试集。

算法研究阶段，精读论文原文，思考可能的改进和创新方向。通过做阅读笔

记的方式加深对论文的理解，学习前景蒙版、Canny边缘检测、图像梯度、倒角距离等相关知识，通过查阅官方手册、教程等方式学习相关python库函数用法，通过阅读论文的参考文献[9]学习alpha matte的求法。基本复现论文算法，以自建的数据集作为输入，运行算法并评估结果，提出算法可改进之处。从本算法的特点出发，提出了创新思考方向“利用水印模式重复性攻击满屏水印”。

中期整理阶段，通过制作“中期汇报PPT”的方式，总结了立项至今取得的成果、仍未解决的疑问、未来的改进及创新方向，并向项目导师做汇报，以解决疑问并确立下一阶段的目标。

## **（2）项目研究已取得阶段性成果和收获**

1. 使用anaconda安装相关库（opencv、scipy等），搭建了算法运行所需的python环境；通过阅读官方手册等方法，学习了相关python库函数的使用方法，提高了python语言应用能力。
2. 使用python语言编写了图片爬取及批量加水印的程序，构建了150张图片的固定位置水印图像集、150张图片的随机位置水印图像集，图像的大小均为500\*500，格式为png，水印半透明；使用coco数据集，将图片数量扩展至3000张，以便进一步测试。此外，还构建了不透明水印的测试集。
3. 通过查阅文献、做阅读笔记的方式，理解了论文《On the Effectiveness of Visible Watermarks》提出的水印检测与去除算法的流程，掌握了代码复现的思路。在反复阅读论文的过程中，项目组成员提高了英文论文阅读的能力，学习了论文复现的方法和技巧。
4. 项目组成员分工协作，基本实现了对论文中算法的代码复现。
5. 提出了创新思考方向“利用水印模式重复性攻击满屏水印”，即将满屏水印图像分割成小块，从而构建具有相同水印模式的图像集以评估并去除水印。

## 二、项目研究存在的主要问题分析及应对思路与措施（500字内）

1. **问题：**参考文献中的数学推导过程过于复杂，涉及到未学过的数学知识，而这些数学公式恰好是代码复现的核心部分。

**应对：**项目组成员花了很多时间查阅相关资料和书籍，补充学习论文中公式推导所需要的数学知识。

2. **问题：**本算法具有一定的局限性，它只适用于水印半透明的情况，若水印不透明（完全覆盖原图），则无法实现去水印效果。

**应对：**近年来，越来越多的实验证明基于深度学习的去水印方法具有较好的效果（例如circleGAN等），这有待于我们下一步尝试和验证。

3. **问题：**算法运行的效果不是非常理想，还留有一定的水印痕迹。

**应对：**后期继续改进和完善复现代码；考虑增加输入图片的数量；考虑将图像集进行分类，对不同类别的子图像集分别进行训练和评估。

## 三、项目研究下阶段主要任务及时间进程安排（500字内）

- （1） 算法改进与完善：2020年12月至2020年1月
- （2） 探索尝试基于GAN/神经网络的方法：2020年1月至2020年2月
- （3） 算法创新与应用：2020年2月至2020年3月
- （4） 撰写论文或研究报告：2020年3月至2020年4月
- （5） 填写结题表和参加答辩：2020年4月至2020年5月
- （6） 论文发表及成果推广：2020年5月至2020年6月

#### 四、项目组成员个人分工所承担和完成研究内容情况（100字内）

负责人所承担和完成研究内容情况汇报：1. 组织每周会议，明确目标及分工。 2. 阅读论文原文，理解算法思想，复现“图像重建”部分。 3. 使用coco数据集拓展素材库，运行算法并评估去水印效果。 4. 提出算法实际应用的相关思考。

周寒靖所承担和完成研究内容情况汇报：1. 阅读论文原文，理解算法思想，复现 $\alpha_n$ 求解部分。 2. 阅读原论文的参考资料，重点理解和整理closed\_form\_matte算法和数学知识。 3. 爬取图片，加随机水印和固定位置水印，初建素材库

茅一宁所承担和完成研究内容情况汇报：1. 阅读论文原文，理解算法思想，复现“水印估计”部分。 2. 查找资料，理解和整理部分细节的算法和数学知识。 3. 参与答辩ppt制作，绘制算法流程图。

#### 五、项目经费使用情况（说明购置材料、资料、调研、交通等已开支经费数额）（100字内）

购买移动硬盘以存储图像集数据

#### 六、指导教师意见（从研究内容和进展、阶段性成果、存在问题等方面加以评价）（180字内）

同意继续开展工作

签 名：宋明黎  
2020年11月18日

#### 七、院（系）评审意见（100字内）

签名盖章  
年 月 日