**关于灰度图像预测实际颜色的研究与实现**

李智敏

浙江理工大学科技与艺术学院

**摘要**

为灰度图像预测其实际颜色是计算机视觉领域的难题，在历史照片修复，有损图片数据压缩，夜间红外摄像头还原彩色，电影制作等不同领域发挥着重要的作用。其主要任务为给定无类标特征的灰度图片或视频，在保证图像语义正确的情况下，利用计算机视觉技术将其转换为彩色。目前已有的还原方法主要色彩迁移和颜色扩展两种，它们都需要在专家帮助下进行交互式着色。给灰度图像还原色彩无疑是进度极慢且复杂的工程，专业的摄影师，艺术家，都要花几天才能修复一张因泛黄丢失色彩的旧照或灰度图。这个过程需要花费极大的精力和高昂的费用。得益于科技的发展，我们能够以低成本收集大量的彩色高清图片用于深度学习模型的训练。依托于GPUs并行计算硬件和设计良好的模型，我们可以在几十毫秒至几秒的时间范围内为灰度图像预测出符合人类认知的颜色。因此关于灰度图像预测实际颜色的研究与实现具有非常重要的研究价值与实际意义。本文基于深度学习技术对该问题进行研究，具体内容如下:

(1) 采用深度学习公用数据库，如imageNet，Instagram等，其中包括建筑，风景，动植物，人类生活相关的照片。

(2) 针对传统灰度图像预测算法的弊端，如交互成本高，需要专家辅助等，本文提出了一种基于深度学习的图像着色算法，选用不同的损失函数取得不同的效果并与其他方法进行比较。通过实验证明，本文提出的算法在颜色预测的质量(包含准确度，预测风格)和速度上好于其他方法。

(3) 本文提出利用集成学习技术，将不同模型预测的结果进行多数投票和加权计算，进一步提高模型预测的质量。

(4) 本文针对在视频领域的灰度预测问题进行了研究与分析。

(5) 我们将模型应用到生活中，大批量还原灰度图片，证明本文提出算法的有效性和实际意义。

**关键字：色彩预测；卷积神经网络；深度学习；集成学习；图像处理**

Abstract

Predicting the actual color of gray image is a difficult problem in the field of computer vision. It plays an important role in historical photo restoration, lossy image data compression, night infrared camera color restoration, film production and other fields. Its main task is to give gray images or videos without class label features. Under the condition of ensuring the correct semantics of images, the computer vision technology is used to convert them into colors. At present, there are two main restoring methods, color migration and color expansion, which need to be interactive coloring with the help of experts. Restoring color to a grayscale image is a slow and complex project, and it takes professional photographers and artists a few days to repair an old photograph or grayscale image that has lost color due to yellowing. This process takes a lot of energy and high cost. Thanks to the development of technology, we can collect a large number of color and high-definition pictures at a low cost for the training of deep learning model. Depending on the GPUs parallel computing hardware and well-designed model, we can predict the color of gray image in the time range of tens of milliseconds to a few seconds. Therefore, the research and implementation of grey image prediction of actual color has very important research value and practical significance. This paper studies the problem based on deep learning technology. The details are as follows:

(1) The use of deep learning common databases, such as imageNet, Instagram, etc., including buildings, landscapes, animals and plants, human life-related photos.

(2) Aiming at the drawbacks of traditional gray-scale image prediction algorithms, such as high interactive cost and expert assistance, this paper proposes an image coloring algorithm based on deep learning. Different loss functions are selected to achieve different results and compared with other methods. Experiments show that the proposed algorithm is superior to other methods in quality (including accuracy, prediction style) and speed of color prediction.

(3) In this paper, ensemble learning technology is used to conduct majority voting and weighted calculation on the results of different models to further improve the quality of model prediction.

(4) the gray prediction problem in the video field is studied and analyzed in this paper.

(5) We apply the model to life and restore gray images in large quantities, which proves the effectiveness and practical significance of the proposed algorithm.

**Key words: color prediction; image coloring; convolution neural network; deep learning; ensemble learning; image processing**

**目录**

[一、 绪论 5](#_Toc526516841)

[(一) 选题依据，研究背景和意义 5](#_Toc526516842)

[1. 选题依据 5](#_Toc526516843)

[2. 研究背景 5](#_Toc526516844)

[(1) 灰度图像的来源 6](#_Toc526516845)

[(2) 彩色图像在特定领域彩色成本高昂 6](#_Toc526516846)

[(3) 原本的色彩无法表达对应的情绪，需要更改 6](#_Toc526516847)

[(4) 因色彩缺失造成没有彩色 6](#_Toc526516848)

[3. 研究意义 7](#_Toc526516849)

[(二) 国内外研究现状 7](#_Toc526516850)

[1. 基于传统方法的研究 7](#_Toc526516851)

[2. 基于深度学习进行图像着色的研究现状 7](#_Toc526516852)

[(三) 本文研究内容 8](#_Toc526516853)

[1. 研究内容 8](#_Toc526516854)

[2. 论文结构 8](#_Toc526516855)

[(四) 本章小结 8](#_Toc526516856)

[二、 图像领域深度学习相关流程 8](#_Toc526516857)

[三、 基于深度学习技术为灰度照片预测实际颜色算法的研究 8](#_Toc526516858)

[(一) 实验环境 8](#_Toc526516859)

[(二) 数据集处理 8](#_Toc526516860)

[1. 数据清洗 9](#_Toc526516861)

[2. 大小格式转换 9](#_Toc526516862)

[(三) 色彩模型转换 9](#_Toc526516863)

[1. LAB色彩模型 9](#_Toc526516864)

[2. BGR转换LAB的过程 9](#_Toc526516865)

[(四) 划分训练集 测试集 验证集 9](#_Toc526516866)

[(五) 网络结构设计 9](#_Toc526516867)

[(六) 损失函数的设计 9](#_Toc526516868)

[1. 模型训练和调参 10](#_Toc526516869)

[2. 模型对比 10](#_Toc526516870)

[3. 结果分析 10](#_Toc526516871)

[4. 本章小结 10](#_Toc526516872)

[(七) 集成学习技术 10](#_Toc526516873)

[1. 集成学习简介 10](#_Toc526516874)

[2. 不同模型进行集成学习 10](#_Toc526516875)

[3. 本章小结 10](#_Toc526516876)

[四、 着色在视频中的应用 11](#_Toc526516877)

[(一) 基于场景不变的视频着色 11](#_Toc526516878)

[1. 相邻帧 11](#_Toc526516879)

[2. 颜色融合 11](#_Toc526516880)

[(二) 变化的场景着色 11](#_Toc526516881)

[(三) 本章小结 11](#_Toc526516882)

[五、 总结与展望 11](#_Toc526516883)

[(一) 关于视频领域的展望 11](#_Toc526516884)

[(二) 全文总结 12](#_Toc526516885)

# 绪论

## 选题依据，研究背景和意义

### 选题依据

为灰度图像预测其实际颜色是计算机视觉领域的难题，在历史照片修复，有损图片数据压缩，夜间红外摄像头还原彩色，电影制作等不同领域发挥着重要的作用。其主要任务为给定无类标特征的灰度图片或视频，在保证图像语义正确的情况下，利用计算机视觉技术将其转换为彩色。目前已有的还原方法主要色彩迁移和颜色扩展两种，它们都需要在专家帮助下进行交互式着色。给灰度图像还原色彩无疑是进度极慢且复杂的工程，专业的摄影师，艺术家，都要花几天才能修复一张因泛黄丢失色彩的旧照或灰度图。这个过程需要花费极大的精力和高昂的费用。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。

得益于科技的发展，我们能够以低成本收集大量的彩色高清图片用于深度学习模型的训练。依托于GPUs并行计算硬件和设计良好的模型，我们可以在几十毫秒至几秒的时间范围内为灰度图像预测出符合人类认知的颜色。因此关于灰度图像预测实际颜色的研究与实现具有非常重要的研究价值与实际意义。

### 研究背景

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，黑白摄影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。在数字时代，灰度图像仍然来源广泛。

#### 灰度图像的来源

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。

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。

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。

#### 彩色图像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。

#### 原本的色彩无法表达对应的情绪，需要更改

(1) 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。

(2) 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。

#### 因色彩缺失造成没有彩色

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。

### 研究意义

## 国内外研究现状

### 基于传统方法的研究

在过去，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，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。

### 基于深度学习进行图像着色的研究现状

有文献作者使用均方误差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。目前没有该项目较完美通用的解决方案。

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。

我们在待处理信息量不可扩充的前提下，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。

## 本文研究内容

### 研究内容

### 论文结构

## 本章小结

# 图像领域深度学习相关流程

**待补充**

# 基于深度学习技术为灰度照片预测实际颜色算法的研究

## 实验环境

服务器……

## 数据集处理

### 数据清洗

### 大小格式转换

## 色彩模型转换

### LAB色彩模型

光照强度指单位面积所接受可见光的光通量，简称照度，用于指示光照的强弱和物体表面被照明程度的量。

颜色模型LAB基于人对颜色的感觉。Lab 中的数值描述正常视力的人能够看到的所有颜色。 因为 Lab 描述的是颜色的显示方式，而不是设备生成颜色所需的特定色料的数量，所以 Lab 被视为与设备无关的颜色模型。色彩管理系统使用 Lab 作为色标，以将颜色从一个色彩空间转换到另一个色彩空间。Lab色彩模型是由亮度（L）和有关色彩的a, b三个要素组成。L表示亮度（Luminosity），a表示从洋红色至绿色的范围，b表示从黄色至蓝色的范围。L的值域由0到100，L=50时，就相当于50%的黑；a和b的值域都是由+127至-128，其中+127 a就是红色，渐渐过渡到-128。a的时候就变成绿色；同样原理，+127 b是黄色，-128 b是蓝色。所有的颜色就以这三个值交互变化所组成。Lab色彩模型除了上述不依赖于设备的优点外，还具有它自身的优势：色域宽阔。它不仅包含了RGB，CMYK的所有色域，还能表现它们不能表现的色彩。人的肉眼能感知的色彩，都能通过Lab模型表现出来。另外，Lab色彩模型的绝妙之处还在于它弥补了RGB色彩模型色彩分布不均的不足，因为RGB模型在蓝色到绿色之间的过渡色彩过多，而在绿色到红色之间又缺少黄色和其他色彩。如果我们想在数字图形的处理中保留尽量宽阔的色域和丰富的色彩，最好选择Lab。

### BGR转换LAB的过程

## 划分训练集 测试集 验证集

## 网络结构设计

## 损失函数的设计

### 模型训练和调参

### 模型对比

### 结果分析

### 本章小结

## 集成学习技术

### 集成学习简介

### 不同模型进行集成学习

### 本章小结

# 着色在视频中的应用

## 基于场景不变的视频着色

### 相邻帧

### 颜色融合

## 变化的场景着色

## 本章小结

# 总结与展望

## 关于视频领域的展望

## 全文总结