

审定成绩: _____

重庆邮电大学
毕业设计（论文）

中文题目	基于粗粒度感知构建和细粒度交互学习 的盲图像质量评价系统设计与实现
英文题目	Design and Implementation of a Blind Image Quality Assessment System Based on Coarse-Grained Perception Construction and Fine-Grained Interaction Learning
学院名称	计算机科学与技术学院/人工智能学院
学生姓名	赵拓荀
专 业	数据科学与大数据技术
班 级	04082002
学 号	2020211843
指导教师	胡波 讲师
答 辩 组 负 责 人	栾晓 教授

二〇二四年六月
重庆邮电大学教务处制

计算机科学与技术/人工智能学院本科毕业设计(论文)

诚信承诺书

本人郑重承诺：

我向学院呈交的论文《基于粗粒度感知构建和细粒度交互学习的盲图像质量评价系统设计与实现》，是本人在指导教师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明并致谢。本人完全意识到本声明的法律结果由本人承担。

年级 2020

专业 数据科学与大数据技术

班级 04082002

承诺人签名

年 月 日

学位论文版权使用授权书

本人完全了解重庆邮电大学有权保留、使用学位论文纸质版和电子版的规定，即学校有权向国家有关部门或机构送交论文，允许论文被查阅和借阅等。本人授权重庆邮电大学可以公布本学位论文的全部或部分内容，可编入有关数据库或信息系统进行检索、分析或评价，可以采用影印、缩印、扫描或拷贝等复制手段保存、汇编本学位论文。

（注：保密的学位论文在解密后适用本授权书。）

学生签名：

指导老师签名：

日期： 年 月 日 日期： 年 月 日

摘要

随着互联网的快速发展，图像成为了信息交互过程中不可或缺的一部分。在社交软件中，数字图像可能面临着传输失真、压缩失真等问题，影响人们的观感体验。为了能够让失真图像更好地满足人类的主观感受，图像质量评价在实际生活中十分必要。而在科研领域，图像质量评价也是计算机视觉领域的重要研究内容。然而，现有的盲图像质量评价指标大多采用端到端的方式，不能充分模拟人的主观评价过程，限制了模型性能的进一步提高。在感知过程中，人们首先对图像的失真类型和相对质量给予初步印象，然后在两者的相互作用下给出具体的质量评分。虽然一些方法试图探索失真类型和相对质量的影响，但它们之间的关系一直被忽视，未能得到充分挖掘。

在本文中，我们提出了一个基于粗粒度感知构建和细粒度交互学习的盲图像质量评价方法，该方法的基本思想为：模拟人在主观图像质量评价中的学习过程，让模型从初步感知出发，更好地理解图像的质量。具体而言，在预训练阶段，骨干网络最初处理一对具有伪主观分数的合成失真图像，并且多尺度特征提取模块整合深度信息并将其传递给粗粒度感知构建模块，其执行失真判别和质量排序。在微调阶段，我们提出了一个细粒度交互式学习模块对两者信息进行交互，以进一步提高模型的性能。实验结果表明，该算法不仅在合成失真数据集上取得了较好的性能，而且在真实失真数据集上也取得了卓越的表现。

在提出算法的基础上，我们进一步设计了图像质量评价系统以更好地展示我们的算法和满足科研人员对质量评价算法的使用需求。具体地，系统中的质量评价模块可以对图像质量评价任务进行可视化展示并得出质量分数。在去模糊模块中，使用者可以选择两种不同的图像去模糊算法，生成其对应的去模糊图像，并利用图像质量评价算法对两者进行分数对比，以帮助使用者对图像去模糊算法的对比和认知。在模型框架模块中，使用者可以查询不同算法的框架图以便使用者对模型有进一步的了解。在美学板块中，使用者可以使用美学算法对图像进行美学评价。

关键词：盲图像质量评价，交互信息，细粒度感知，神经网络

Abstract

With the development of Internet interaction, image transmission has become an integral part of the information interaction process. In social software, digital images may face problems such as transmission distortion and compression distortion. To make the distorted images better meet the subjective feelings of human beings, image quality evaluation is necessary in real life. And in the field of scientific research, image quality assessment is also an important research content in the field of computer vision. However, most of the existing blind image quality assessment metrics use an end-to-end approach, which cannot fully simulate the human subjective evaluation process and limits the further improvement of the model performance. In the perception process, people first give an initial impression of the distortion type and the relative quality of an image, and then give a specific quality score under the interaction of the two. Although some approaches have attempted to explore the effects of distortion type and relative quality, the relationship between them has been neglected.

In this thesis, we propose a blind image quality assessment method based on coarse-grained perception construction and fine-grained interactive learning. The basic idea of this thesis is to simulate the human learning process in subjective image quality assessment, so that the model can understand the image better from initial perception. Specifically, in the pre-training phase, the backbone network initially processes a pair of synthetic distorted images with pseudo-subjective scores, and the multi-scale feature extraction module integrates the depth information and passes it to the coarse-grained perception building module, which performs distortion discrimination and quality ranking. In the fine-tuning stage, we propose a fine-grained interactive learning module to interact the two information to further improve the performance of the model. Experimental results show that the algorithm not only achieves better performance on synthetic distortion datasets, but also achieves excellent performance on real distortion dataset.

Based on the proposed algorithm, we further design the image quality assessment system to better demonstrate our algorithm and to meet the needs of researchers for using the quality assessment algorithm. Specifically, the quality assessment module in the system can visualize the image quality assessment task and derive the quality score. In

the deblurring module, the user can select two different image deblurring algorithms, generate their corresponding deblurred images, and compare the scores of the two using the image quality assessment algorithm to help the user's comparison and knowledge of the image deblurring algorithms. In the model framework module, the user can query the framework diagrams of different algorithms so that the user can have a further understanding of the model. In the Aesthetics section, the user can use the aesthetics algorithm to evaluate the image aesthetically.

Keywords: Blind image quality assessment, interaction information, fine-grained perception, neural network

目录

第 1 章 引言	1
1.1 研究背景和意义	1
1.2 国内外研究现状	2
1.2.1 主观图像质量评价	2
1.2.2 客观图像质量评价	7
1.2.3 评价指标	8
1.3 主要内容和工作安排	10
第 2 章 算法实现相关技术介绍	11
2.1 PyTorch 深度学习框架	11
2.2 卷积神经网络和 ResNet50 网络简介	12
2.2.1 卷积神经网络介绍	12
2.2.2 ResNet50 网络介绍	16
2.3 本章小结	18
第 3 章 基于粗粒度感知构建和细粒度交互学习的盲图像质量评价	21
3.1 引言	21
3.2 算法实现细节	23
3.2.1 数据准备	23
3.2.2 多尺度特征提取模块	24
3.2.3 粗粒度感知构建模块	26
3.2.4 细粒度交互式学习模块	27
3.2.5 深度质量回归模块	28
3.3 本章小结	29
第 4 章 算法实验结果展示	31
4.1 实验设置	31
4.2 单一数据集实验	31
4.3 跨库实验	33

4.4 消融实验.....	35
4.5 细粒度图像实验.....	37
4.6 细粒度交互学习的特征可视化.....	38
4.7 算法生成的失真评估可视化.....	39
4.8 本章小结.....	40
第 5 章 系统设计与实现.....	41
5.1.1 Flask 框架介绍.....	41
5.1.2 Vue 框架介绍.....	41
5.1.3 nginx 框架介绍.....	42
5.2 系统的功能介绍.....	43
5.2.1 图像质量评价功能.....	43
5.2.2 算法框架展示功能.....	46
5.2.3 去模糊算法评价功能.....	48
5.2.4 美学评价功能.....	49
5.3 本章小结.....	51
第 6 章 总结与展望.....	53
6.1 主要工作与创新点.....	53
6.2 后续研究工作展望.....	53
参考文献.....	55
致谢.....	61

第 1 章 引言

1.1 研究背景和意义

在多媒体技术快速发展的今天，图像和视频数据的质量显著影响用户的观看体验。通过有效的图像质量评价，我们可以确保用户接收到高质量的视觉内容，从而显著提升用户的满意度和互动体验。图像质量评价通常分为主观评价和客观评价两种方法。主观评价依赖于人类视觉系统的判断，而客观评价则通过算法模拟人眼对图像质量的感知。其中，客观评价方法可以进一步分为全参考（FR-IQA），半参考（RR-IQA），和无参考（NR-IQA）评价方法。

随着人工智能和机器学习技术的进步，图像质量评价的研究不断深入，已经能够更准确地模拟人类的视觉感知机制，为图像质量的自动化评价提供了新的方法和技术支持。这一领域的发展对于提升图像处理技术、优化用户体验以及推动相关技术的商业化应用都具有重要意义。具体地：首先，图像质量评价在图像和视频的传输及存储过程中尤为重要，因为这些过程需要对数据进行压缩以减少数据量。通过图像质量评价，可以帮助优化压缩算法，平衡压缩率和图像质量的关系，确保在最小的数据量下仍获得最佳的视觉效果；其次，图像处理技术如图像增强、修复、超分辨率等，都需要通过图像质量评价来提供性能评估的标准。客观的评价方法能够更准确地指导这些算法的改进和优化；最后，在特定应用领域，如医疗影像和卫星遥感图像，图像质量直接关系到诊断和分析的准确性。图像质量评价不仅能自动识别出质量不佳的图像，而且确保高质量的图像被用于进一步的分析和决策支持，从而提高整体的工作效率和准确度。随着人工智能和机器学习技术的进步，图像质量评价的研究不断深入，已经能够更准确地模拟人类的视觉感知机制，为图像质量的自动化评价提供了新的方法和技术支持。这一领域的发展对于提升图像处理技术、优化用户体验以及推动相关技术的商业化应用都具有重要意义。

在算法的基础上，一个图像质量评价的网站系统可以进一步地帮助普通用户以及科研工作者更好地了解和學習图像质量评价算法。普通用户可以利用这个网站系统评价自己日常生活中拍摄的真实图像的质量。例如，用户在旅行、家庭聚会或任何重要活动中拍照后，常常希望得知所拍摄的照片质量。通过网站提供的图像

质量评价工具，用户可以即时了解图像是否清晰、色彩是否准确、是否存在模糊或噪点等问题，从而帮助他们选择最佳的照片用于打印、分享或存档。

科研工作者在图像处理、计算机视觉领域的研究中，经常需要对图像质量进行严格评估。通过这个网站系统，他们可以轻松访问和使用多种先进的图像质量评价算法，对研究中使用的图像进行客观和标准化的评价。这不仅加快了研究过程，也提高了研究的精确性和有效性。此外，网站可能还提供算法比较、评分系统和用户反馈，这些都能帮助研究者优化自己的算法或了解当前图像质量评价领域的前沿技术。

网站系统可以使用户将图像质量评价应用于更广泛的实际场景，比如通过对比不同去模糊算法处理后的图像质量，用户不仅能评价这些算法的效果，还能深入理解去模糊技术的工作原理及其效果的实际应用。例如，摄影师可以使用这些工具来测试不同的图像处理软件对于提升模糊照片清晰度的效果，而科研人员则可以基于这些数据进一步研究图像处理算法的优化。

总之，这个图像质量评价网站系统不仅为普通用户提供了一个简便易用的工具来评估和改善个人图像的质量，也为科研人员提供了一个实用的平台，用于测试和研究最新的图像质量评估技术。这样的系统将架起科研与日常应用之间的桥梁，促进图像质量评价技术的普及和发展。

1.2 国内外研究现状

1.2.1 主观图像质量评价

主观图像质量评价依赖于观察者对图像的直观感受，涉及对图像清晰度、颜色准确性、视觉舒适度等多个因素的评估。在此过程中，评价者通常会同时观看多幅图像，并根据 1-5 分或 1-10 分的评分系统对它们进行排序或打分。评分维度包括图像的整体质量、颜色保真度、边缘清晰度等。

主观评价的重要性体现在几个方面。首先，它能直接反映用户对图像的实际感受，是评估用户满意度的直接方式。其次，在图像处理算法的开发过程中，如去噪、压缩、和增强，主观评价提供了关键反馈，帮助开发者改进算法以更好地满足用户

需求。此外，主观评价在图像和视频编码标准的制定过程中扮演了关键角色，它帮助决策者确定哪些技术能最优化用户的视觉体验。

主观图像质量评价包括了几个著名的数据库，分别为模拟失真数据库：LIVE 数据库，CSIQ 数据库，TID2013 数据库；真实失真数据库：LIVE Challenge 数据库，KonIQ-10K 数据库。下面分别介绍各个数据库的具体情况：

1) LIVE 数据库

LIVE 数据集是一个广泛用于图像质量评估的数据集。该数据集包括 29 幅高分辨率和高质量的彩色参考图像，来源于互联网和摄影光盘，涵盖了人脸、动物、特写、广角拍摄、自然场景、人造物体以及不同前景/背景配置的图像等多种类型。



图 1.1 LIVE 数据库 gblur 图

LIVE 数据集包含 779 幅经过降质处理的失真图像，这些图像模拟了实际应用中常见的损伤情况。这些图像通过 5 到 6 个不同等级的降质处理得到，涵盖了各类图像损伤类型，包括结构性损伤（例如高斯模糊）、图像压缩损伤（如 JPEG 压缩、JPEG2000 压缩及 JPEG2000 的快速尺度衰减失真）和随机噪声（如白噪声）。每种损伤类型分为四至五个不同的级别。

每幅失真图像都对应有平均主观得分差（Differential mean opinion score, DMOS），这些得分是由 161 个实验者共约 25000 次评分得出的。DMOS 值的范围在 0 到 100 之间，较高的 DMOS 意味着较低的图像质量。图 1.1 和 1.2 分别展示了 LIVE 数据库的 gblur 图和 fastfading 图。

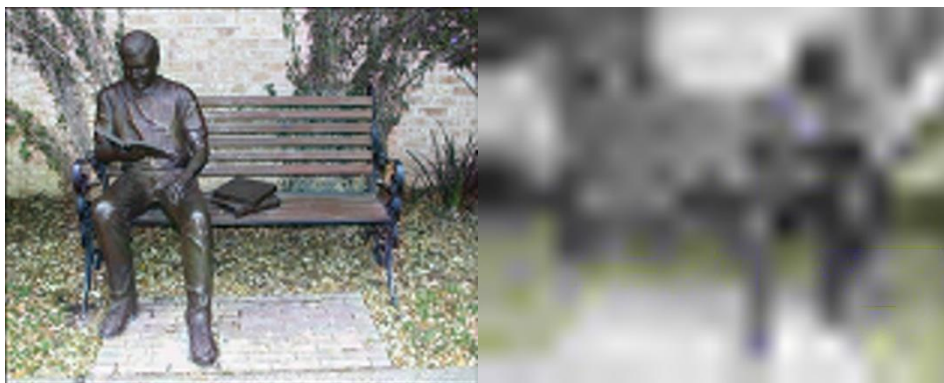


图 1.2 LIVE 数据库 fastfading 图

2) CSIQ 数据库

CSIQ 数据库包含 30 幅参考图像和 866 幅失真图像，涵盖六种失真类型，包括 JPEG 压缩、JPEG2000 压缩、整体对比度降低、加性高斯粉红噪声、加性高斯白噪声以及高斯模糊。CSIQ 的 DMOS 值基于 25 位观察者的大约 5000 个评分统计，取值在 0 到 1 之间，0 代表最优质量，1 代表最差质量。图 1.3 展示了 CSIQ 数据库的图像。



图 1.3 CSIQ 数据库图

3) TID2013 数据库

TID2013 数据集广泛应用于实验中，它包含 25 幅参考图像和超过 3000 幅测试图像。每幅参考图像被处理以展示 24 种不同的失真类型，每种失真各有 5 个级别。这些失真类型包括了常见的图像失真，例如 JPEG 压缩和模糊等，以及新增的颜色方面的失真。TID2013 数据集的建立是通过对来自五个不同国家的 985 名志愿者进行主观实验来收集图像的平均主观评分。

TID2013 数据集的建立是通过对来自五个不同国家的 985 名志愿者进行主观实验来收集图像的平均主观评分。



图 1.4 TID2013 数据库部分图像

此外，TID2013 数据集还包含了 25 张无失真彩色参考图像，其中 24 张来自柯达数据集，第 25 张参考图像是人工创建的，并添加到 24 张自然场景图像中。这些参考图像涵盖了各种不同的内容风格，有些包含丰富的纹理，而其他一些则包含较大的均匀区域，以确保对图像质量评估算法的全面评估。图 1.4 展示了 TID2013 数据库部分图像。

表 1.1 模拟失真数据库基本信息

名称	CSIQ	LIVE	TID2008	TID2013	Kadid-10K
发布时间	2009	2006	2008	2013	2019
观察者数量	35	161	838	971	30
图像类型	彩色	彩色	彩色	彩色	彩色
图像格式	PNG	BMP	BMP	BMP	PNG
主观分数形式	MOS	DMOS	MOS	MOS	MOS
失真种类	6	5	17	24	25
失真等级	4-5	5-6	4	5	5
主观分数范围	[0-1]	[0-100]	[0-9]	[0-9]	[1-5]
参考图像数量	30	29	25	25	81

失真图像数量	866	779	1700	3000	10125
模糊图像数量	150	145	100	125	—

3) LIVE Challenge 数据库

LIVE Challenge (LIVE-C) 数据集是一个真实失真图像质量挑战数据集，建立于 2016 年。该数据集包含了 1162 幅图像，这些图像是使用现代移动设备在真实世界中捕获的，并且受到了各种真实失真的影响。图 1.5 展示了 LIVE-C 数据库部分图像。

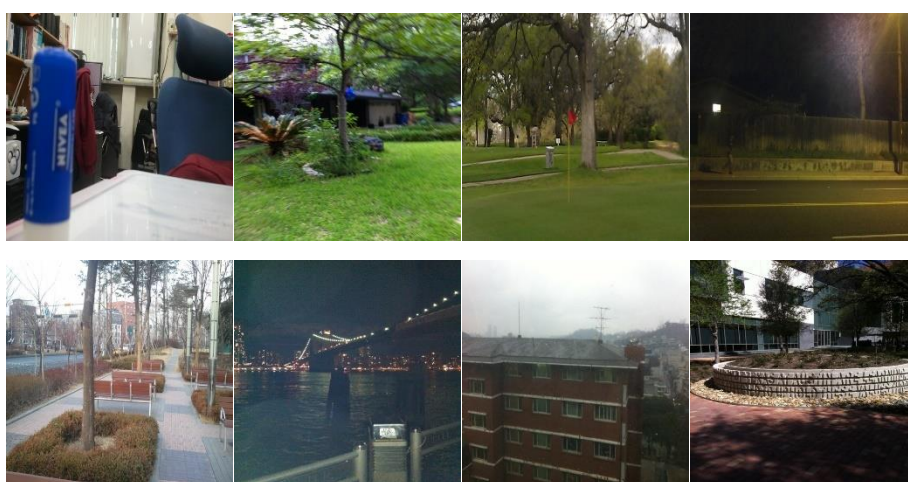


图 1.5 LIVE-C 数据库部分图像

4) KonIQ-10K 数据库

与 LIVEC 相比，Hosu 等人于 2020 年建立的 KonIQ-10K 数据库则更加广泛且均匀地分布在亮度、色彩、对比度和清晰度方面。该数据集从 YFCC100M 数据集中选择了 10073 张图片，并通过基于深度特征的内容指标和 7 个质量指标的采样过程，确保了图像内容和质量分布的多样性。



图 1.6 KonIQ-10K 数据库图像

对于每张图片，通过众包获得了 120 个可靠的质量评级，由 1459 名志愿者进行标注，收集的主观数据高达 120 万。在这个数据集中，MOS 的取值范围为 3.91 到 88.39。图 1.6 展示了 KonIQ-10K 数据库的图像。

表 1.1 和表 1.2 分别展示了模拟失真数据库基本信息和真实失真数据库基本信息。

表 1.2 真实失真数据库基本信息

名称	KonIQ-10k	Live-challenge
发布时间	2018	2015
测试者数量	1459	8100
图像格式	JPG	BMP/JPG
主观分数形式	MOS	MOS
主观分数范围	[0-100]	[0-100]
失真图像数量	10073	1169

1.2.2 客观图像质量评价

现有的客观图像质量评价主要分为传统和深度学习两大类，下面分别对其进行介绍。

1) 传统 NR-IQA

自然场景统计 (NSS) [1-3] 和码书 [4-13] 通常用于构建传统的质量评价模型。通过提取图像信息并对其退化的图像进行建模，可以有效地检测图像质量的退化程度。Moorthy 和 Bovik [1] 通过使用失真图像统计设计了一种盲 IQA 算法。另一方面，Mittal 等人 [2] 利用空间域的 NSS 特征来评估图像质量。Saad 等人 [3] 采用离散余弦变换 (DCT) 域特征来设计 NR-IQA 模型。所有这些方法都取得了良好的效果。在基于码本的度量中，Ye 等人 [4] 通过无监督的方式学习字典，然后通过最大池化的软分配编码获得视觉感知特征。Jiang 等人 [5] 通过优化多级判别词典解决了高维和语义鸿沟问题。传统的质量度量方法在合成失真数据集上可以取得显著的效果。然而，当处理真实的失真场景时，它们的性能往往受到限制。这主要是因为真实失真的图像往往更加复杂和不均匀，传统的机器学习模型很难提取复杂的失真信息。

2) 基于深度学习的 NR-IQA

随着深度学习技术的巨大成功,许多计算机视觉任务都取得了重大突破。因此,大多数 NR-IQA 方法主要依赖于深度学习技术,从而取得了显著的进步和令人满意的结果。在论文^[14]中,提出了一种双流网络架构,用于预测合成和真实图像失真,其中 VGG 16^[15,16]用于使用真实质量预测提取特征。Zhang 等人^[13]提出利用来自合成和真实数据集的图像对来训练他们的模型,该模型通过使用 ResNet-34 作为主干来提取特征。Fang 等人^[17]设计了一个基于基线和多任务深度神经网络的 BIQA 模型。Lv 等人^[18]提出了一种感知信息增强方法,用于提取图像特征并基于多尺度卷积、残差级联和通道注意力增强网络表示。在 Li 等人的工作中^[19],他们专注于图像内容变化对质量的影响,并提出了一种基于语义特征聚合的 NR-IQA 模型。在[20]中, Ma 等人利用动态参数隔离策略来更新网络中的参数子集,以帮助它更好地学习图像与分数之间的映射关系。为了描述图像的低级视觉特征和高级语义信息, Sun 等人^[21]引入了一种阶梯结构,其中 CNN 的特征层次被集成。这种集成使模型能够有效地利用低级和高级视觉信息进行全面的分析和理解。

有几个算法试图使他们的设计特定模块更好地获取有关失真类型或相对质量的信息,这些巧妙的想法已经取得了良好的效果。在文献[22]中,为了解决真实图像不足的问题,引入了一种预训练的方式,通过判断图像的相对优良度,使模型建立图像的初始知识。为了增强主观和客观分数的单调相关性, Golestaneh 等人^[23]利用图像之间的相对距离信息来生成相对排名。Wang 等人^[24]提取失真特征来捕获失真类型和失真程度,并利用失真特征和语义之间的交互信息来帮助生成图像分数。在[25]中, Zhou 等人认为人类大脑能够进行质量和类型区分,并且存在相互作用。他们通过扭曲辨别任务促进了质量评估任务的学习。即使他们都认识到存在的积极作用,这两个信息的质量评估任务,他们没有进一步探讨相对质量和失真类型之间的关系。因此,本文将重点放在模仿人类主观评价过程,重点研究失真类型和质量排名的影响,并建立一个高性能的模型。

1.2.3 评价指标

Spearman's Rank Correlation Coefficient(SRCC), Pearson Linear Correlation Coefficient(PLCC), Kendall's Rank Correlation Coefficient(KRCC) 和 Root Mean Square Error (RMSE)是四种常用的评价指标。下面分别介绍它们:

SRCC（斯皮尔曼等级相关系数）用于衡量两个变量等级之间的相关性。它通过比较变量的等级顺序，而非直接使用数据值，来评估关联度。**SRCC** 的值范围在 -1 到 1 之间，1 表示完全正相关，-1 表示完全负相关，0 则表示无相关性。值接近 1 或 -1 时，表明变量间等级顺序关联紧密，而接近 0 则表示关联性弱。在图像质量评估中，**SRCC** 常用于比较算法评分与人工评分的相关性。

PLCC（皮尔逊线性相关系数）衡量两个变量间的线性关系强度和方向。通过计算变量间协方差与其标准差的乘积，**PLCC** 的值也在 -1 到 1 之间，1 表示完全正相关，-1 表示完全负相关，0 表示无线性关系。值越接近 1 或 -1，表明变量间线性关系越强，接近 0 则线性关系弱。在图像质量评价算法中，通常用于评估算法评分与人工评分之间的线性相关性。

KRCC 是一种非参数统计方法，用来衡量两个序列间的序位相关性。**KRCC** 的值介于 -1 到 1 之间，其中 1 表示完全正相关，-1 表示完全反相关，0 则表明两者之间无相关性。**KRCC** 的计算基于序列中的一对数据项的一致性和矛盾性对数。具体公式如下：

$$KRCC = \frac{2(P - Q)}{n(n-1)} \quad (1.1)$$

其中， P 是一致对(concordant pairs)的数量， Q 是矛盾对(discordant pairs)的数量， n 是序列中的元素数量。

KRCC 不假设数据遵循任何特定的分布，使其适用于各种数据类型，对于异常值和非线性关系具有较好的鲁棒性。但是当数据集较大时，计算所有数据对的一致性和矛盾性可能变得计算成本高。对于数据集中的绑定排名(tied ranks)可能不够敏感。

RMSE 衡量数据预测模型或估计器准确性的常用指标之一。它表示预测值与实际值之间差异的平方的平均值的平方根。它直接反映了预测误差的大小，误差越小，**RMSE** 值越低，模型的预测性能越好。**RMSE** 的计算公式是：

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - Q_i)^2} \quad (1.2)$$

其中， P_i 表示预测值， Q_i 表示观测值或实际值， n 表示观测点的数量。

RMSE（均方根误差）是衡量数据预测模型或估计器准确性的常用指标。它计算的是预测值与实际值之间差异的平方的平均值的平方根，直接反映了预测误差的大小。误差越小，**RMSE** 值越低，表明模型的预测性能越好。

1.3 主要内容和工作安排

本文设计了基于粗粒度感知构建和细粒度交互学习的盲图像质量评价算法，并将该算法引用到展示平台中，以图像质量评价的现实用途和科研用途，并在系统中进一步展示算法细节和美学功能。

全文共 6 章，内容结构安排如下：

第一章为引言。第一节主要介绍了盲图像质量评价算法的研究背景和意义，可见该算法不仅可用于现实生活中的日常图像评价，还推进去模糊领域，图像超分辨率重建领域的研究。第二节介绍了目前图像质量评价领域的国内外研究现状。第三节介绍了本文的主要内容和工作安排。

第二章为图像质量评价算法的相关技术。主要包括 **PyTorch** 深度学习框架，卷积神经网络和 **Resnet50** 网络框架。

第三章为基于粗粒度感知构建和细粒度交互学习的盲图像质量评价，主要包括了算法实现的各个模块。

第四章为算法的实验结果展示。第一节为实验设置，第二节为单一数据集实验，第三节为跨库实验，第四节为消融实验，第五节为细粒度图像实验，第六节为细粒度交互学习的特征可视化，第七节为软件生成的失真评估可视化。

第五章为系统的设计与实现。主要包括系统设计实现相关技术细节介绍，以及系统的功能介绍和展示。

第六章为总结与展望。在该章节总结本文的工作内容，回顾了整个系统的设计与实现，并表明了今后的研究方向。

第 2 章 算法实现相关技术介绍

2.1 PyTorch 深度学习框架

PyTorch 在 2016 年推出后凭借其易用性、灵活性和强大功能迅速受到科研社区的青睐。其设计理念与 Python 相契合，注重易读性和简洁性，使用 Python 语言编写，扩展了 Python 的特性，使得其学习和应用更为便捷。官网如图 2.1 所示。

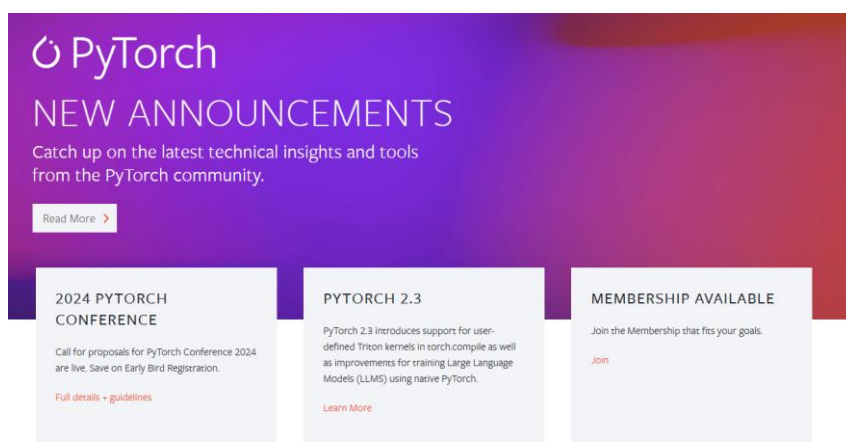


图 2.1 PyTorch 官网

PyTorch 区别于其他框架的一个显著创新是采用了动态计算图。与静态计算图不同，动态计算图允许用户在运行时修改计算图的结构，这增加了处理复杂模型时的灵活性。这种设计使得研究人员能够更加直观地理解和调试模型。随着时间的推移，PyTorch 在科研社区中迅速获得了广泛认可。2019 年发布的 1.0 版本引入了许多重要功能，包括支持 ONNX、新的分布式包以及对 C++ 的前端支持等，使得 PyTorch 在工业界应用更加广泛，同时保持了在科研领域的强劲地位。

PyTorch 是一个集成度高的机器学习库，它在 Python 环境中完全集成，支持所有 Python 库和工具。这使得开发者可以充分利用 Python 的功能来构建和训练模型。PyTorch 以其直观的 API 设计和模块化结构而受到新手和专业人士的青睐，这些特性使得用户可以灵活地设计和调整深度学习模型。其中一个核心特性是其动态计算图（自动微分系统），它允许在运行时动态构建图的结构，这大大简化了模

型的调试和优化过程。因此，PyTorch 特别适合于快速实验和原型设计，研究人员可以迅速测试新的想法或算法。

PyTorch 不仅在学术研究中被广泛使用，它的实用性也扩展到了生产环境。通过其子项目 TorchScript，PyTorch 允许模型转换成可以在非 Python 环境中运行的格式，从而支持模型的部署。此外，PyTorch 支持在多种硬件平台上运行，如 GPU 和 TPU，这有助于加速模型的训练和推理。PyTorch 社区的活跃也是其一大优势，拥有大量的教程、论坛和文档，为用户提供问题解决的支持。此外，PyTorch 支持多种高级功能，如多 GPU 训练和定制化的神经网络层，使得研究人员和工程师能够针对具体问题进行了优化。

2.2 卷积神经网络和 ResNet50 网络简介

2.2.1 卷积神经网络介绍

CNN 源自 Hubel 和 Wiesel 对猫视觉皮层的研究，是多层感知机（MLP）的一种特殊形式，如图 2.2 所示。CNN 的特性特别适用于图像处理，允许直接将图像用作网络输入，自动提取颜色、纹理、形状和拓扑结构等特征，并在处理图像的位移、缩放和其他形式的扭曲时显示出良好的鲁棒性和高效的计算性能。

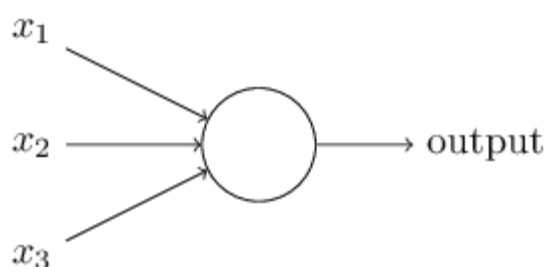


图 2.2 神经网络

CNN 是深度学习领域的一个重大突破，自从它的创立以来，对整个神经网络设计和科学研究产生了深远的影响。CNN 最初是为了解决图像识别问题而设计的，由于其独特的结构设计，使得它在处理图像、视频、声音和文本数据方面表现出色。

如图 2.3 所示，CNN 的设计核心是局部感知区和权值共享。这种结构反映了自然图像中的局部性原理，即图像的高层特征是由较低层的局部特征组合而成的。

权值共享则大大减少了模型的参数数量，使得训练更加高效。通过使用卷积层，CNN 能够学习到平移不变的特征，即使图像中的对象位置发生变动，依然可以被有效识别。这一点在图像和视频分析中尤其重要。通过多层卷积层和池化层的堆叠，能够抽象出从简单到复杂的特征，逐层提取图像中的有用信息，这种层次化的特征提取方式极大地提升了模型的表达能力。

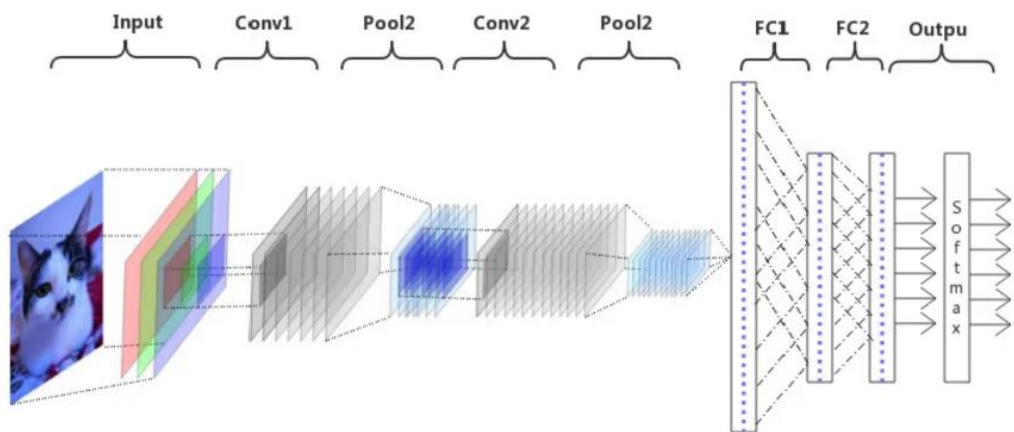


图 2.3 卷积神经网络

CNN 的成功不仅在视觉任务中表现突出，其基本概念和技术原理也被广泛应用于多个领域，包括自然语言处理、声音识别和无人驾驶等。这些应用都受益于 CNN 在层次化特征提取方面的能力，该技术能够从简单到复杂逐层提取数据的关键信息。

此外，CNN 的设计理念也启发了一系列新的神经网络架构的发展。例如，ResNet、Inception 和 DenseNet 等架构在 CNN 的基础上引入了残差连接和多尺度处理等创新技术，这些技术不仅提高了网络的性能，还增强了网络处理更复杂问题的能力。CNN 的应用并不限于处理视觉数据。在自然语言处理领域，通过使用词嵌入作为输入，CNN 能够有效捕捉文本中的局部相关性，展示了其在处理非视觉数据中同样的强大潜力。

由于 CNN 对计算资源的高需求，它也推动了 GPU 和 TPU 等专用硬件的发展。这些硬件的优化不仅加速了 CNN 模型的训练过程，还使得这些模型的部署更为高效，从而推动了整个人工智能领域的技术进步和应用普及。

1) 卷积操作

如图 2.4 所示，卷积操作是 CNN 中用于提取输入图像特征的基本方法。它涉及使用一个小的二维权重矩阵，即卷积核，这通常是 3×3 或 5×5 的正方形。卷积核通过与图像的局部区域逐元素相乘后求和，从而捕捉特征。在执行卷积时，卷积核在图像上滑动，每到一个新位置，都会重复这个逐元素相乘并求和的过程，生成输出特征图的一个像素值。

步长是卷积核移动的距离，决定了输出特征图的尺寸。较大的步长会缩小特征图尺寸，而较小的步长则可能增加尺寸。常用的步长是 1，但也可以根据需要进行选择。填充是在输入图像周围添加额外像素的做法，用来控制输出特征图的大小并减轻边缘效应。常见的填充策略包括“valid”和“same”

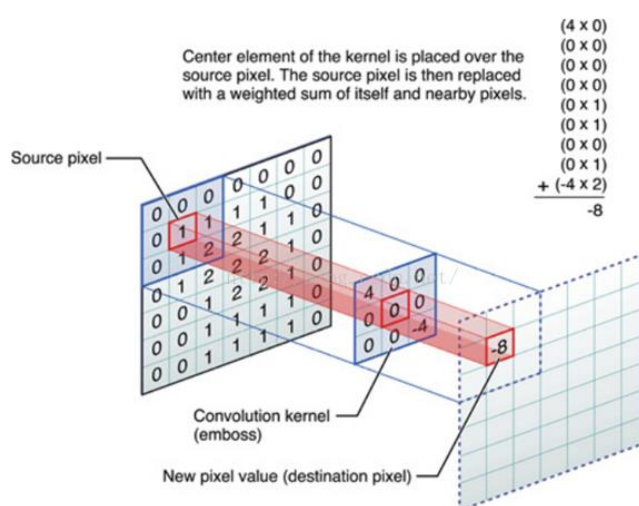


图 2.4 卷积操作

在卷积神经网络(CNN)中，处理多通道输入图像时，卷积核也具备多个通道，每个卷积核通道与输入图像的相应通道进行卷积操作。这些层级和操作的组合使得 CNN 在处理和识别图像数据方面表现出高效和强大的能力。

这种多通道卷积操作使得 CNN 能够有效地提取图像的局部特征，如边缘和纹理等。通过叠加多个这样的卷积层，加入非线性激活函数，CNN 能构建更为复杂的特征表达。每个通道的卷积核独立地与其对应的图像通道相乘后求和，然后将所有通道的结果累加起来，形成最终的输出特征图。

2) 池化操作

如图 2.5 所示，池化操作是卷积神经网络（CNN）中的一种关键技术，对于 CNN 的性能和效率具有重要意义。池化层通常在卷积层之后使用，主要功能是降低特征图的空间维度，从而减少计算量和模型中的参数数量。这一操作有助于提高网络的计算效率，并减轻过拟合的风险。池化层通过对邻近像素的下采样，可以在一定程度上提供对图像平移、旋转和尺度变化的不变性，增强了模型对于输入变化的鲁棒性。

池化的提出是为了模仿生物视觉系统中的感受野机制，通过局部区域的信息聚合来抽象更高层次的特征。这一设计哲学与其他一些神经网络架构不同，如全连接网络或传统的浅层学习模型，它们通常对所有输入特征进行全局处理，而不具有明显的层次化或局部性特征处理能力。

池化操作的提出和应用，显著增强了 CNN 在处理图像和其他高维数据时的性能，使其在多个领域的应用中表现出更高的效率和更好的效果。

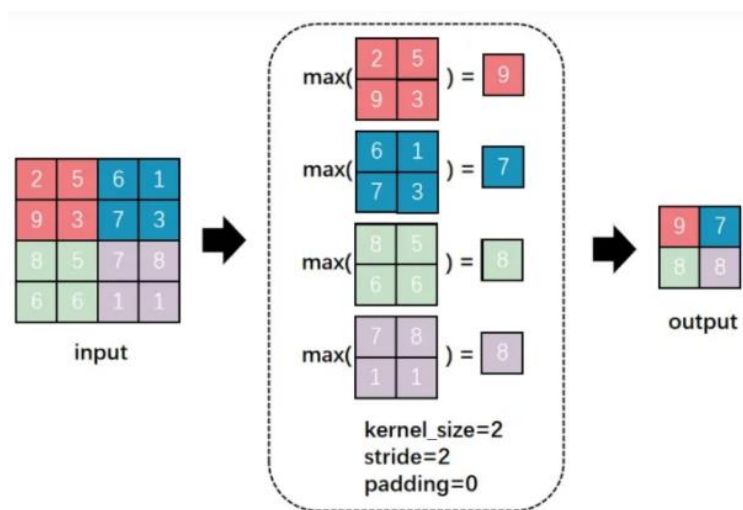


图 2.5 最大池化操作

3) 激活函数

激活函数是非线性函数，用于引入非线性特性并增加网络的表达能力。常用的激活函数包括 Sigmoid、ReLU（Rectified Linear Unit）、Tanh 和 Leaky ReLU 等。

如图 2.6 所示，Sigmoid 函数将输入值压缩到 0 到 1 之间，适合处理二分类问题，但容易出现梯度消失的问题。ReLU 函数通过将负值置零来提供 faster 的训练速度和更好的性能，但可能导致神经元死亡的问题。Tanh 函数将输入值压缩到 -1

到 1 之间，类似于 Sigmoid 但输出范围更广，适合处理多分类问题。Leaky ReLU 允许小于零的输入值通过一个小的斜率，从而解决了 ReLU 函数的神经元死亡问题。

选择合适的激活函数取决于任务的性质和网络的架构，通常在实践中会通过交叉验证等方法进行选择 and 调优。激活函数在卷积神经网络中扮演着至关重要的角色，影响着网络的收敛速度和性能表现。

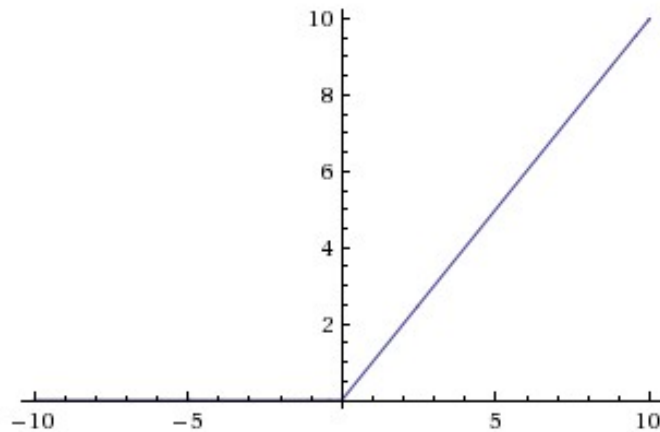


图 2.6 ReLu 激活函数

2.2.2 ResNet50 网络介绍

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

图 2.7 ResNet 网络图

ResNet-50 是由微软研究院提出的一种深度残差网络（Residual Network），是 ResNet 系列中的一员，如图 2.7 所示。它以其深度和性能而闻名，是 ImageNet 大规模图像识别挑战赛上的常用模型之一。

ResNet-50 的核心创新在于引入了残差连接（residual connections），这种连接方式使得网络可以更深，同时减轻了训练过程中的梯度消失问题。通过在网络的不同层之间添加残差连接，ResNet-50 能够训练出比以往更深的网络，同时减少了参数数量和计算复杂度，提高了模型的性能。

如图 2.8 所示，ResNet-50 的架构由一系列的卷积层、批量归一化层和残差块（residual blocks）组成，其中每个残差块包含多个卷积层和恒等映射或者维度匹配的 1×1 卷积层以适配输入和输出的维度。ResNet-50 在训练过程中通常使用带动量的随机梯度下降（SGD）优化器，并结合数据增强等技术来提高模型的泛化能力。

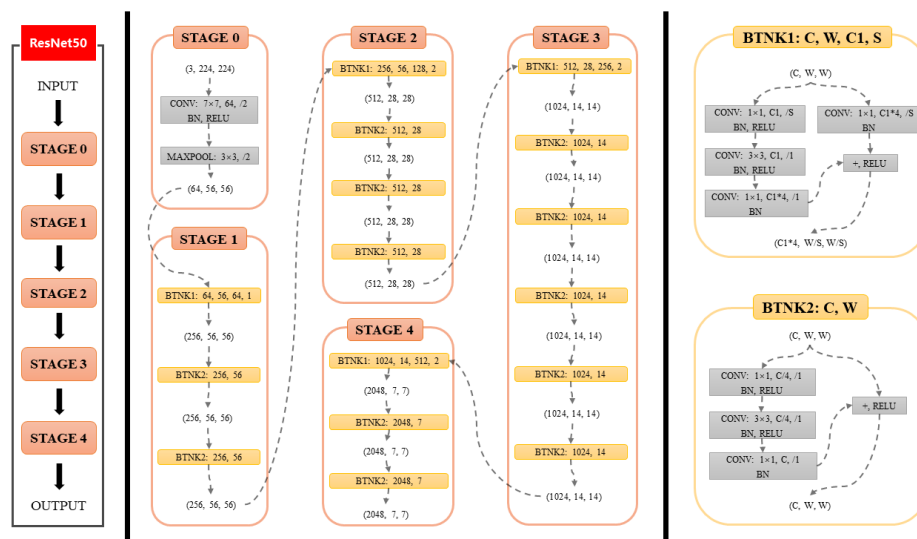


图 2.8 ResNet50 框架图

ResNet 的创立标志着深度学习领域的一个重大突破，尤其是在解决训练非常深的神经网络时遇到的梯度消失或梯度爆炸问题。在 ResNet 之前，随着网络层次的增加，这些问题常常限制了模型的深度和效能。ResNet 的核心创新在于引入了“残差学习”的概念，通过恒等映射（也称为跳跃连接）将输入直接传递到后面的层，从而解决了深度模型训练的难题。

通过引入跳跃连接，ResNet 允许梯度直接流过多个层次，有效地防止了在极深网络中的梯度问题。利用残差块，ResNet 能构建出比之前更深的网络，如 ResNet-

152 等，这在多个视觉任务中显著提高了性能。这些残差块不仅使得网络即使在深层结构下也能有效训练，还大大提高了训练的收敛速度和效率。

ResNet 的成功激励了研究者探索包含短连接的其他网络架构，如 DenseNet 等。这些架构在 ResNet 的基础上引入了更复杂的连接模式，进一步提高了信息和梯度的流动效率。残差学习的理念也被广泛应用于自然语言处理、语音识别等领域的深度学习任务中，显著改善了这些任务的模型性能和训练稳定性。随着更深的网络结构的开发和普及，对高效计算硬件和优化算法的需求也随之增长，推动了相关技术的发展。ResNet 的设计理念和架构在深度学习社区中广受接受，成为许多复杂任务的标准网络架构。

2.3 本章小结

在本章中，我们深入探讨了图像质量评价算法的实现技术，涉及了 PyTorch 深度学习框架、卷积神经网络（CNN）以及 ResNet50 模型。这些技术在图像质量评价领域发挥着重要作用，为用户提供了便捷准确的图像质量评价服务。

首先，PyTorch 作为一种流行的深度学习框架，为实现图像质量评价算法提供了强大的支持。其简洁明了的 API 设计以及动态计算图的特性使得算法的开发和调试更加高效。通过 PyTorch，我们可以轻松构建复杂的神经网络模型，并进行训练和推理。

其次，卷积神经网络（CNN）在图像处理领域具有突出的表现。其通过卷积层、池化层等结构有效地提取图像特征，具备对图像进行深度学习特征表示的能力。在图像质量评价算法中，CNN 可以帮助我们捕获图像的清晰度、对比度、噪声等方面的特征，从而实现对图像质量的评价。

最后，ResNet50 模型作为一种经典的深度卷积神经网络模型，在图像识别和特征提取方面表现优异。其通过残差学习结构有效解决了深层网络训练中的梯度消失和梯度爆炸问题，提高了网络的收敛速度和性能。在图像质量评价算法中，ResNet50 模型可以作为评价模型的基础，通过学习大量的图像数据来提高评价的准确性和鲁棒性。

综上所述，通过 **PyTorch** 深度学习框架构建基于卷积神经网络的 **ResNet50** 模型，可以有效实现图像质量评价算法。这种算法技术的应用，不仅可以为用户提供方便快捷的图像质量评价服务，还能够在图像处理领域推动技术的进步与创新。

第3章 基于粗粒度感知构建和细粒度交互学习的盲图像质量评价

3.1 引言

随着通信技术的进步,图像已日益成为用于各种目的的主要数据形式,例如数据传输,医学成像,监视等。然而,在采集、传输、存储的过程中,图像不可避免地会出现不同程度的失真,严重影响了人类对图像的感知和图像相关领域的发展。因此,确保对图像质量的精确评估至关重要。图像质量评价(Image Quality Assessment)可分为主观 IQA 和客观 IQA。虽然主观 IQA 可以获得准确的图像评分,但人类直接评估图像的成本非常昂贵,无法处理大量图像^[26],并且快速高效地评估图像也会带来挑战。相反,在客观的 IQA 中,预先设计的算法可以自动有效地执行 IQA 任务,这是近年来的主要研究方向。而根据方法是否需要参考图像,客观 IQA 可分为三类:全参考 IQA(Full Reference IQA)^[27,28],半参考 IQA(Reduced Reference IQA)^[29]和无参考 IQA(No Reference IQA)^[26,30-32]。在现实中,由于大多数图像没有参考图像,因此 NR-IQA 的研究是 IQA 领域的重点^[33-35]。

近来,随着深度学习技术的进步和成功,领域内已经提出了许多基于深度学习的算法来准确测量和评估图像质量^[14,16,18,19]。与传统的依赖手工设计特征的通用模型相比,深度学习具有更强大的表达能力,可以实现比传统方法更好的评估结果。大多数现有的深度学习算法采用卷积神经网络(CNN)的架构,添加全连接网络,来细化深度语义并产生质量分数。这种从图像端直接到分数端的端到端方式,利用了流行的 CNN 框架作为两者之间的桥梁,获得了图像更深层次的语义。部分算法通过修改 CNN 框架进一步利用图像的深层语义特征,以便网络可以更好地学习和理解图像细节^[18,21]。从目前的算法上看,端到端的方式比手工设计的特征更适合理解图像质量,而这种方法目前在 BIQA 领域占据主导地位。

尽管前文的方法取得了很大的进步,但这些方法仍然具有以下缺点。首先,端到端的方式强调直接从语义信息为每个图像生成分数,因此忽略了失真类型对质量的影响^[22,23]。如果失真类型可以为模型学习提供辅助信息,BIQA 度量的性能将进一步提高^[1,24,25]。其次,大多数度量都将单一图像作为网络的输入^[36,37]。一些指

标的确会选择两个图像^[24,38]，但以类似于一个图像作为输入的方式，它们忽略了不同图像之间的相对质量对比。从主观 IQA 上看，由于缺乏对比度信息，受试者很难对单个失真图像进行评估。事实上，受试者在评估前往往会被提供不同的图像，他们需要通过比较这些图像，在大脑中建立模糊的初始感知。这在很大程度上指导他们评估随后的所有图像。图 3.1 说明了这种构建两个阶段的评分方法。因此，和主观 IQA 中建立模糊感知的过程一样，模型也可以通过图像之间的比较来建立初始感知，这将进一步提高模型的性能。最后，在模型的构建过程中如何综合考虑图像表层信息和深层语义等方面，将进一步影响模型的性能。

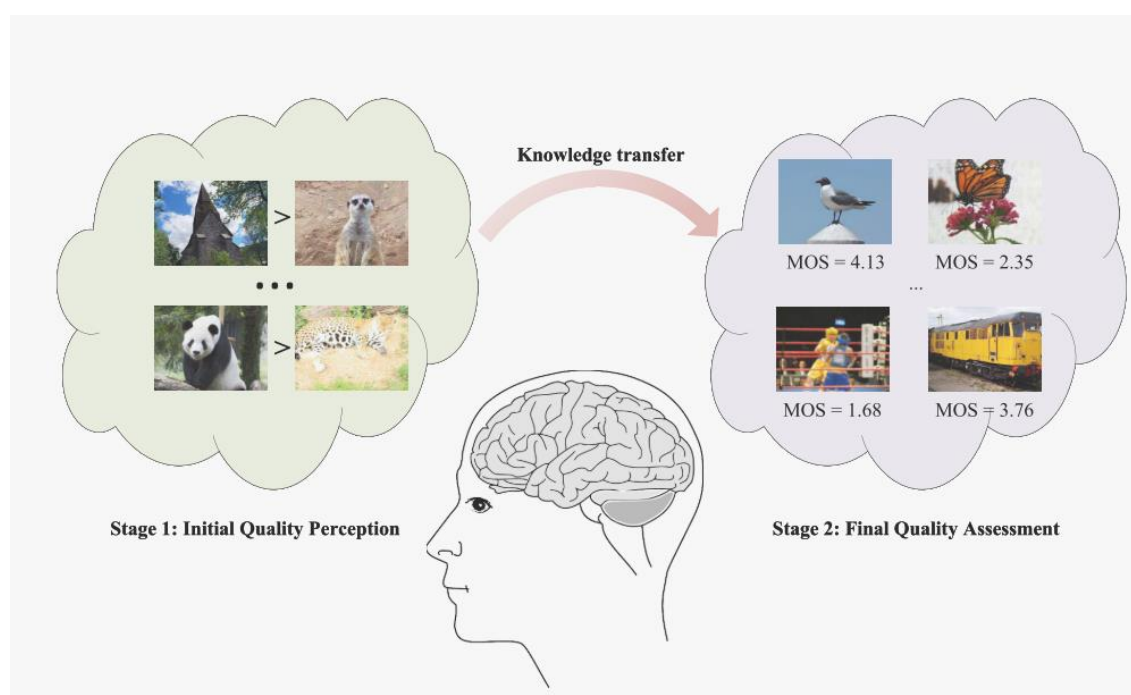


图 3.1 初始感知在质量评价中的作用

受此启发，本文提出了一种基于粗粒度感知构建和细粒度交互学习的盲图像质量评价方法。该算法的基本思想是模仿人类在实际情况下的两阶段评分方法，即同时考虑失真类型和相对质量，然后利用两者之间的关系进行交互学习，从而构建高性能的模型。具体地，首先，基于低层视觉特征和高层语义信息对感知视觉质量的影响，提出了一种阶梯结构的语义特征提取模块。然后，在预训练阶段，通过粗粒度感知构建模块建立模型的失真类型和相对质量的初始感知。在微调阶段，基于注意力机制，进一步提出了一个交互式学习模块，以充分整合相对质量和失真类型

之间的信息。最后，利用一个简单的回归网络将学习到的特征映射到质量分数。实验结果证实了该算法的优越性。

3.2 算法实现细节

在本节中，我们的目标是开发一个能够模拟人类的感知过程来评估图像质量模型，即首先有一个粗粒度的感知（如失真判别和质量排名），然后通过组合各种因素形成最终的评价。本算法包括了三个模块：多尺度特征提取模块、粗粒度感知构建模块和细粒度交互学习模块。图 3.2 为算法框架图，本章将并详细描述了每个模块。

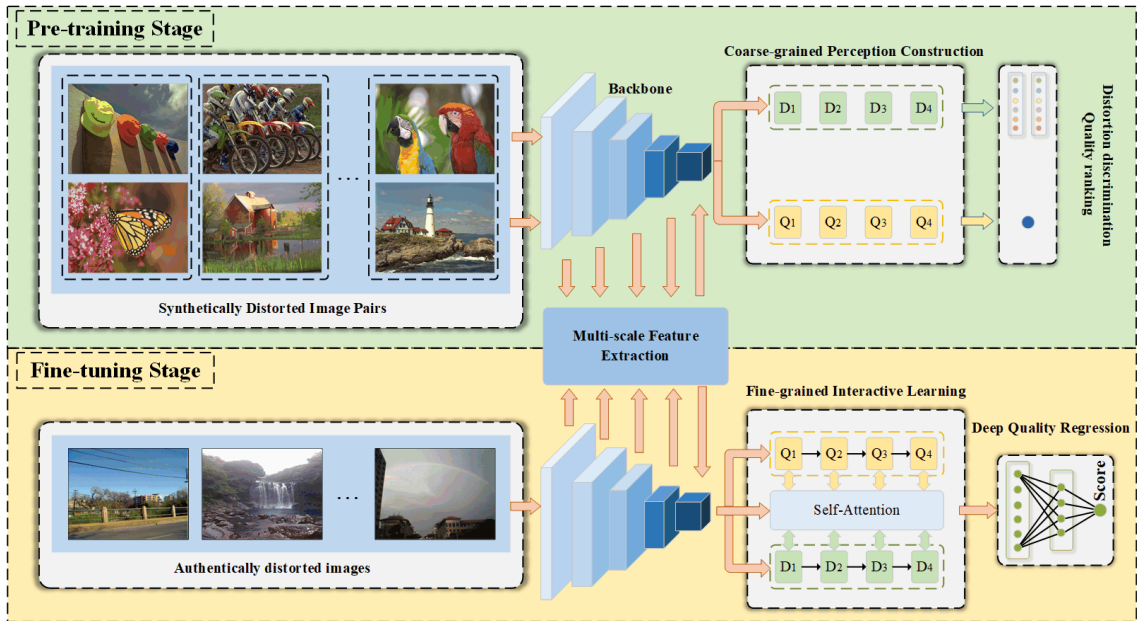


图 3.2 算法网络框架图

3.2.1 数据准备

本算法选择使用一个包含了 1000 张原始图像大型数据集对算法进行预训练^[22]。首先，我们利用传统的合成失真方法对原始图像进行了 6 种不同程度的图像失真处理和 6 种不同类型的失真处理，包括高斯模糊、运动模糊、JPEG、JPEG2000、椒盐噪声以及白色噪声，这些都是合成图像中常见的失真类型。利用上述合成失真，

一张原始图像对应于 36 个不同的失真图像，每个失真图像都拥有其分数和失真类型。每一张生成图像的分数由 SSIM^[39] 算法生成，如图 3.3。在预训练过程中，算法会随机选择两幅得分差异大于 0.1 的图像作为粗粒度感知构建模块的输入。而在 SSIM 算法精度高于 0.1 级的情况下，将可以得到两幅图像的近似相对质量。与方法^[22]不同，我们模块的输入是得分差异大于 0.1 的任意两个图像，它们不需要来自于同一张参考图像。这种策略不仅可以获得更多的图像匹配对，而且可以让模型在匹配过程中感知包含不同内容的图像之间的差异。



图 3.3 大型数据集部分高清图像

3.2.2 多尺度特征提取模块

在真实环境下，真实的失真类型往往更加复杂多样。因此，从图像中提取多尺度特征的方法近年来逐渐流行。在 HVS 中，人眼首先会对整幅图像进行全局观察和信息理解，然后再关注图像的局部区域。在信息处理过程中，全局信息和局部信息会相互影响。因此，本章设计了多尺度特征提取模块来获取图像的多尺度特征。该模块背后的核心概念是，为了进一步提取 CNN 每个特征中的信息，CNN 的特

征层次结构被集中到分层网络中。与算法^[21]中的多尺度方法类似，本模块采用分层结构来有效地从图像中提取局部失真信息。从 CNN 的每一层经过多次深度卷积后形成的特征块只能简单地表示图像的高层语义信息，但丢失了重要的低层语义信息。该模型从各个维度出发，尽可能地将各个层次之间的语义信息结合起来，以保证不丢失图像的失真信息。

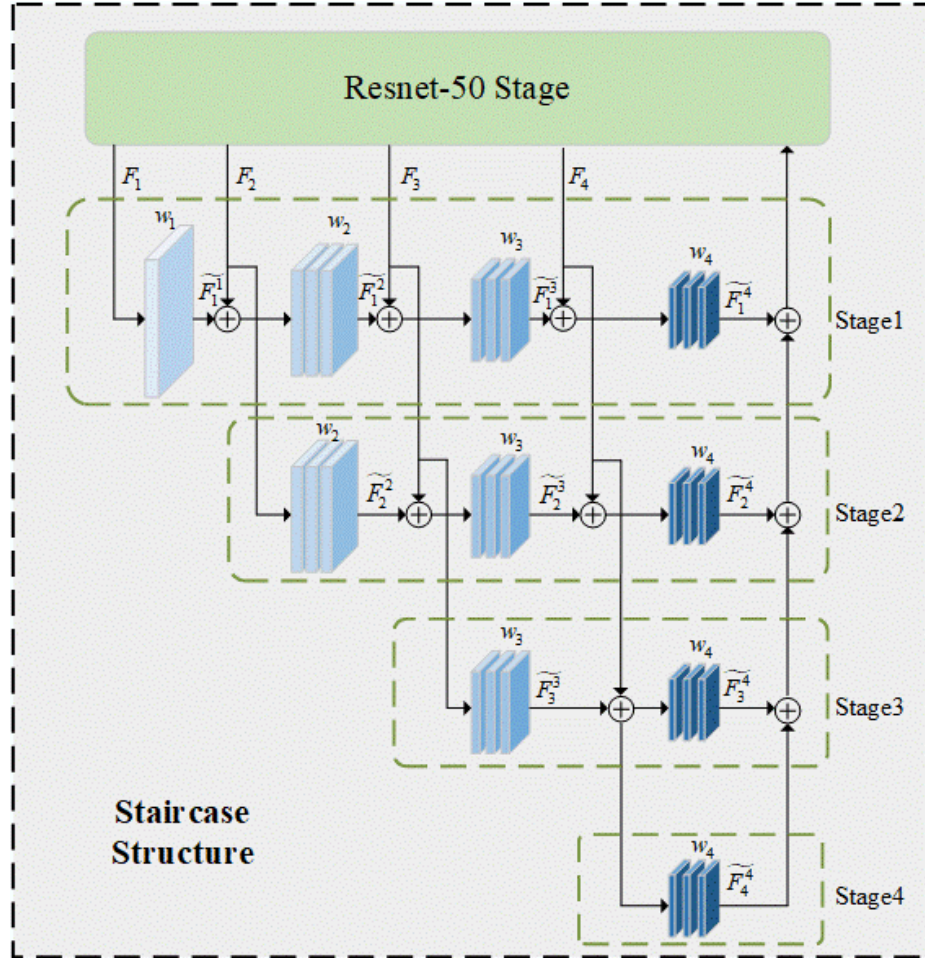


图 3.4 多尺度特征提取模块框架图

如图 3.4 所示，CNN 网络下方是一个广义的局部失真信息提取器，它可以有效地理解图像内容并整合局部和全局信息。多尺度特征提取模块和失真感知模块中使用的骨干网络是 Resnet-50。从层 conv1、conv2_10、conv3_12 和 conv4_18 中提取特征块 F_1, F_2, F_3, F_4 作为多尺度特征提取模块的输入。这四个特征块是每个阶段的原始特征和中间学习特征。例如，特征块 F_1 将是第一阶段的原始特征，并且特征块 F_2, F_3, F_4 将被用作中间特征以帮助 F_1 整合信息。

由于特征块 F_i 的维度和大小在每个 CNN 层次之后是不同的，为了使前一个特征块 F_i 和下一个特征块 F_{i+1} 具有相同的维度和大小，它们必须在可以组合之前被统一。特征块 F_i 的维数通过卷积层 1×1 降低到四分之一，降低了整个过程的计算复杂度。然后使用卷积层 3×3 将 F_i 的大小减少到一半。最后，通过卷积层 1×1 将特征块 F_i 的维度放大 8 倍，以获得其维度和大小与 F_{i+1} 相同的 \tilde{F}_i 。通过进行上述操作，卷积块 F_i 的尺寸可以扩大两倍，并且大小可以减小到原始大小的一半。特征映射的表达式为：

$$F_i = F_i W_{1 \times 1} W_{3 \times 3} W_{1 \times 1} = F_i W_i, \quad i = 2, 3, 4 \quad (3.1)$$

其中， $W_{1 \times 1}$ 和 $W_{3 \times 3}$ 对应于 1×1 卷积层和 3×3 卷积层的权重矩阵，并且 W_i 表示通过将 $W_{1 \times 1}$ 、 $W_{3 \times 3}$ 和 $W_{1 \times 1}$ 相乘而获得的结果。在这种情况下，由于特征块 F_1 与 F_2 具有相同的大小，因此它只需要通过 1×1 的卷积层，以确保相同数量的空间通道。

$$F_1 = F_1 W_{1 \times 1} = F_1 W_1 \quad (3.2)$$

在维度和大小相同之后，可以通过添加相应的元素来融合两个特征 \tilde{F}_i 和 F_{i+1} 。以第一阶段为例， F_1 在通过 F_1 大小的 W_1 后变为 \tilde{F}_1^1 ，与 F_2 融合后，在通过 F_2 大小的 W_2 后变为 \tilde{F}_1^2 。以此类推，最后得到的结果是 \tilde{F}_1^4 。然后，在 j 阶， i 特征的 \tilde{F}_i^j 为：

$$F_j^i = (F_j^{i-1} + F_i) W_i, \quad i = 2, 3, 4, j = 1, 2, 3, 4 \quad (3.3)$$

最后，将每个阶段的融合特征 \tilde{F}_i^j 与 CNN 的最终输出特征 F_5 合并，表示为：

$$F = F_5 + F_1^4 + F_2^4 + F_3^4 + F_4^4 \quad (3.4)$$

3.2.3 粗粒度感知构建模块

传统的 IQA 通常使用单个图像作为直接输入，并端到端地输出质量分数。然而，在实践中，受试者将被给予一些图像信息，以建立对整个图像数据集的全局感知，而不直接评估图像。而且，传统方法在 BIQA 中只能直接使用深度学习模型来拟合图像进行尽可能多的学习，由于缺乏图像失真信息，这使得模型对图像的理解和思考过程变得残缺，不能像人类一样完整地整合图像失真信息。在主观 IQA 中，将向受试者呈现随机选择的一组图像，以允许他们建立初始感知，而不是直接对图像进行评级。类似地，粗粒度感知构建模块的目的是允许模型构建 IQA 任务的初始感知。

如图 3.4 所示,具体地,粗粒度感知构建模块可以收集相对质量和失真类型以构建双任务学习网络。ResNet50^[40]被选为骨干网络,并继承了 ImageNet^[41]的模型参数。ResNet50 的最后两层被删除,并替换为质量提取器和失真类型提取器,它们具有相同的网络结构,具有四个 3×3 卷积层,一个 Relu 层和一个 7×7 全局平均池化层。在两个提取器之后,有三个完全连接的层用于图像好坏的二进制分类任务和图像失真类型的 6 个分类任务。

$$F_q = Q(F), F_d = D(F) \quad (3.5)$$

其中 F 是输入特征。 $Q(\cdot)$ 和 $D(\cdot)$ 分别是质量提取器和失真类型提取器。 F_q 表示由质量提取器提取的特征,而 F_d 表示由失真类型提取器提取的特征。

通过对具有合成失真的图像的学习,该模块可以初步感知图像的失真信息,并对整个 IQA 任务有全局的理解,这更接近于人们如何感知图像。在粗粒度感知构建模块中,质量提取器参数和失真类型提取器参数将在微调任务中继承。

3.2.4 细粒度交互式学习模块

根据 HVS,当感知图像时,人类试图将图像失真类型的视角作为质量评估的重要参考因素。与人类相似,这两类信息的生成在模型学习过程中应该是相辅相成的,但传统方法直接将辅助任务的特征输入到质量提取模块中。与此相反,本算法将更关注整合和引导来自两者的信息。如果被两类提取器异化的特征相互影响,它们也会有对方的引导信息,作为下一次特征进化的引导因子,引导特征进化得更接近当前图像^[42-47]。

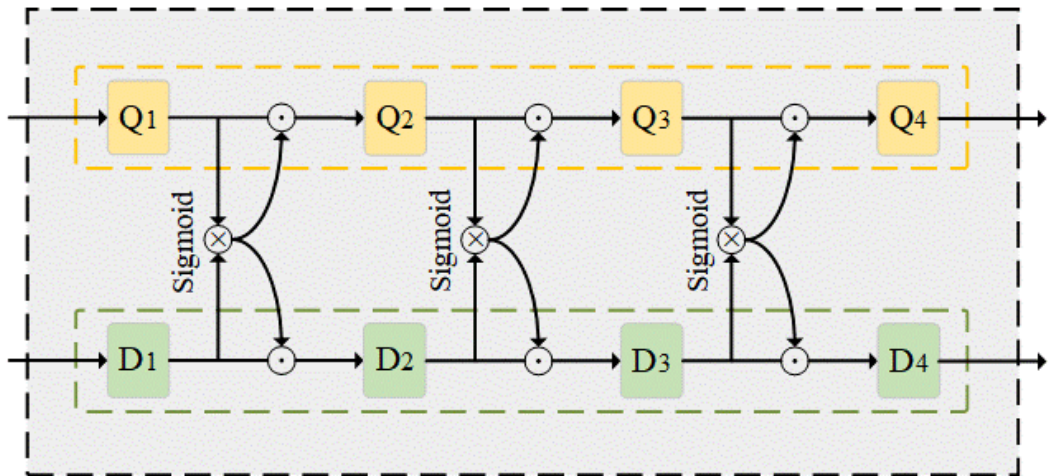


图 3.5 细粒度交互式学习模块框架图

如图 3.5 所示，通过学习整合的失真图像对，质量提取器和失真类型提取器将对图像有初步的感知。首先，通过从预训练继承的两种类型的特征提取器筛选初始特征 F 以获得信息为 F_q^1 和 F_d^1 。

$$F_q^1 = Q_1(F), F_d^1 = D_1(F) \quad (3.6)$$

然后，为了让两者相互学习，实现了一个类似于自我注意机制的策略，如图 3.4 所示。将失真类型信息 F_d^1 投影到 F_q^1 的向量空间中，得到基于 F_q^1 空间的两个特征之间的相似度，归一化后成为准则因子。类似地， F_q^1 将被投影到 F_d^1 中以获得 F_d^1 的引导因子。公式如下：

$$F_q^n = \text{sigmoid}(F_d^n \times F_q^n) \times F_q^n \quad (3.7)$$

$$F_d^n = \text{sigmoid}(F_q^n \times F_d^n) \times F_d^n \quad (3.8)$$

$$F_q^{n+1} = Q_{n+1}(F_q^n), \quad n = 1, 2, 3 \quad (3.9)$$

$$F_d^{n+1} = D_{n+1}(F_d^n), \quad n = 1, 2, 3 \quad (3.10)$$

其中 sigmoid 的函数由下式给出：

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (3.11)$$

3.2.5 深度质量回归模块

在穿过细粒度交互式学习模块之后，全连接层将特征映射到图像质量。通过质量提取器和失真类型提取器将特征块配置在一起。对于质量回归任务，在应用全局平均池化后选择了三个完全连接的层。此任务的损失函数是平均绝对误差 (MAE)，其定义为：

$$L = \frac{1}{N} \sum_i^N |\varphi(i) - q(i)| \quad (3.12)$$

其中 N 表示数据集中的样本数，而 $\varphi(i)$ 和 $q(i)$ 表示第 i 个样本的预测值和对应的真值。

3.3 本章小结

在本节中，我们详细介绍了我们提出的算法的实现细节，包括了算法实现过程中的各个关键模块。首先，我们介绍了数据准备模块，该模块负责对输入数据进行预处理和准备工作，确保数据的质量和适用性。接着，我们介绍了多尺度特征提取模块，该模块旨在从输入数据中提取多尺度的特征，以捕获图像中不同尺度下的信息。然后，我们介绍了粗粒度感知构建模块，该模块利用提取的特征构建粗粒度的感知模型，用于捕获图像的整体失真信息。接下来，我们介绍了细粒度交互学习模块，该模块通过学习图像中不同区域之间的交互关系来进一步优化失真的估计和预测。最后，我们介绍了深度质量回归模块，该模块利用深度学习技术对图像的失真程度进行回归预测，从而得到最终的失真评分。

在接下来的章节中，我们将进一步展示算法的实验结果和性能分析，以验证算法的有效性和可靠性。

第 4 章 算法实验结果展示

4.1 实验设置

为了增加训练样本，根据^[40]中的训练策略，我们用 25 个补丁随机裁剪每个训练图像，并水平翻转 25 个大小为 224×224 像素的图像块进行增强，这些补丁继承了原始图像的质量分数。我们利用 ADAM 优化器来优化模型参数，并应用 5×10^{-5} 的权重衰减。我们使用的批量大小为 32，epoch 为 64。实验是在 Intel Core i9-10980 XE CPU@3.00 GHz 上进行的，沿着两个 NVIDIA GeForce RTX 3090 GPU，通过 Pytorch 实现。

该算法在两种不同类型的数据集上进行了验证：合成失真数据集和真实失真数据集。其中，合成失真数据集包括了 LIVE^[55]，CSIQ^[56]和 TID 2013^[57]。真实失真数据集包括 LIVE Challenge (LIVE-C)^[58]和 KonIQ-10 K^[59]。

在每个数据集中，80% 的图像用于训练目的，而剩余的 20% 用于验证模型的性能。使用这种方法，我们在每个数据集上重复进行 10 次实验，并取最终 SRCC 和 PLCC 的平均值。

4.2 单一数据集实验

我们选择了最先进的基于图像质量评价 (BIQA) 方法，与拟议方法进行性能比较。这些方法包括五种传统方法和十五种基于深度学习的方法。

根据表 4.1 的结果，我们的方法在两种类型的数据集上都表现出更好的性能。首先，在合成失真方面，我们的方法在三个合成图像数据集上优于其他模型。具体来说，在大型数据库 TID 2013 中，我们的方法的 SRCC 达到了 0.877，比第二名高出 0.012，并且 PLCC 高出 0.01。我们的方法采用了基于语义特征的融合，将失真类型信息与相对质量相结合，从而提取出图像中更为复杂和多变的信息，有助于提升图像质量。这解释了为什么我们的方法能在合成数据集上实现 SOTA 结果。

其次，在真实失真的情况下，我们的模型在 LIVE-C 数据集上表现出卓越性能，达到了 0.872 的分数，超过了类似 Zhou 等人^[26]提出的 SOTA 方法。需要注意的是，

由于不均匀性和预测难度增加等因素，在真实失真情况下，一般的评估度量往往表现较差。尽管在真实图像中缺乏原始图像信息和失真类型信息，但通过失真感知网络训练的模型参数仍然可以帮助模型理解和整合图像失真信息^[48-54]。

总之，我们提出的方法通过结合粗粒度感知构建和细粒度交互学习取得了最先进的结果，证明了其优越性。

表 4.1 单一数据集实验结果

Method	Published	CSIQ		LIVE		TID2008	
		SRCC	PLCC	SRCC	PLCC	SRCC	PLCC
DIIVINE ^[47]	TIP11	0.777	0.743	0.913	0.923	0.535	0.664
CORNI ^[4]	CVPR12	0.695	0.768	0.937	0.938	0.657	0.734
BRISQUE ^[2]	TIP12	0.747	0.819	0.940	0.942	0.5589	0.676
HOSA ^[35]	TIP16	0.741	0.812	0.946	0.947	0.735	0.815
NIQE ^[48]	SPL13	0.821	0.865	0.902	0.908	0.521	0.648
RAN4IQA ^[49]	AAAI18	0.914	0.931	0.961	0.962	0.820	0.859
WaDIQaM-NR ^[50]	TIP18	/	/	0.954	0.963	0.761	0.787
Two-stream ^[51]	TIP19	/	/	0.969	0.978	/	/
DBCNN ^[14]	TCSVT20	0.946	0.959	0.968	0.971	0.816	0.865
CaHDC ^[52]	TIP20	0.903	0.914	0.965	0.964	0.862	0.878
HyperIQA ^[53]	CVPR20	0.923	0.942	0.962	0.966	/	/
UNIQUE ^[54]	TIP21	0.902	0.927	0.969	0.968	/	/
IEIT ^[55]	TCSVT22	/	/	/	/	/	/
CONTRIQUE ^[56]	TIP22	/	/	/	/	/	/
VCRNet ^[57]	TIP22	0.943	0.955	0.973	0.974	0.846	0.875
MetaIQA+ ^[58]	TCSVT22	/	/	/	/	/	/
M.D. ^[37]	TIP23	0.903	0.925	0.969	0.978	0.835	0.859
Lan et al. ^[30]	TBC23	0.953	0.956	0.969	0.971	0.865	0.883
Zhou et al. ^[26]	TBC23	0.931	0.959	0.969	0.972	0.849	0.861
Gao et al. ^[59]	TCSVT23	0.933	0.948	/	/	0.844	0.852
Ours		0.956	0.964	0.976	0.978	0.877	0.893

表 4.1 单一数据集实验结果（续）

Method	Published	KonIQ-10K		LIVE-C	
		SRCC	PLCC	SRCC	PLCC
DIIVINE	TIP11	0.579	0.632	0.523	0.551
CORNI	CVPR12	0.545	0.569	0.617	0.636
BRISQUE	TIP12	0.665	0.681	0.578	0.602
HOSA	TIP16	0.791	0.761	0.640	0.678
NIQE	SPL13	/	/	0.594	0.589
RAN4IQA	AAAI18	0.752	0.763	0.586	0.612
WaDIQaM-NR	TIP18	0.797	0.805	0.671	0.680
Two-stream	TIP19	/	/	0.595	0.613
DBCNN	TCSVT20	/	/	0.851	0.869
CaHDC	TIP20	/	/	0.738	0.744
HyperIQA	CVPR20	0.906	0.917	0.859	0.882
UNIQUE	TIP21	0.896	0.901	0.854	0.890
IEIT	TCSVT22	0.892	0.916	0.833	0.865
CONTRIQUE	TIP22	0.894	0.906	0.845	0.857
VCRNet	TIP22	0.894	0.909	0.856	0.865
MetalQA+	TCSVT22	0.909	0.921	0.852	0.872
M.D.	TIP23	/	/	/	/
Lan et al.	TBC23	/	/	/	/
Zhou et al.	TBC23	0.911	0.921	0.871	0.883
Gao et al.	TCSVT23	0.908	0.921	0.830	0.863
Ours		0.912	0.927	0.872	0.884

4.3 跨库实验

为了评估算法的泛化能力，本文不仅分析了在单个数据集上的表现，还进行了跨数据集的实验。具体来说，在表 4.2 中，我们在四个数据集上进行了跨数据集实验，即 LIVE, LIVE-C, CSIQ 和 TID 2013，其中的最佳性能用粗体标记。具体地，我们选择在其中一个数据集中训练数据，并在其余三个数据集中评估其性能。比较方法方面，我们选择在单个数据集评价中表现良好的七种方法进行比较。

从结果来看，该算法在整体跨数据集评估上取得了良好的结果，这证明了该方法具有很强的泛化性能。具体来说，它在大多数时间都处于前两名。通过具体的实

验可以发现，当模型在合成的失真数据集上进行训练，然后在真实的失真数据集上进行测试时，模型的质量指标会显著下降。这是因为真实失真的评估更具挑战性，并且模型很难通过学习简单的合成失真来掌握复杂的真实失真。值得一提的是，尽管 LIVE、CSIQ 和 TID 2013 中包含不同类型的失真图像，该模型仍能获得良好的结果。这些结果都表明，该算法具有很好的泛化能力。

表 4.2 跨库实验

Train \ Test	Method	CSIQ	LIVE-C	TID2013
LIVE	BRISQUE	0.573	0.326	0.354
	HOSA	0.619	0.455	0.470
	WaDIQaM-NR	0.601	0.151	0.396
	HyperIQA	0.744	/	/
	DIIVINE	0.602	0.296	0.342
	Two-Stream	0.621	0.273	0.431
	VCRNet	0.768	0.615	0.502
	Ours	0.799	0.592	0.532
Train \ Test	Method	CSIQ	LIVE	TID2013
LIVE-C	BRISQUE	0.236	0.244	0.275
	HOSA	0.334	0.522	0.372
	WaDIQaM-NR	0.323	0.323	0.141
	HyperIQA	/	/	/
	DIIVINE	0.419	0.353	0.327
	Two-Stream	0.200	0.081	0.049
	VCRNet	0.566	0.746	0.416
	Ours	0.647	0.647	0.355

Train \ Test	Method	LIVE	LIVE-C	TID2013
CSIQ	BRISQUE	0.823	0.106	0.433
	HOSA	0.786	0.277	0.341
	WaDIQaM-NR	0.813	0.106	0.315
	HyperIQA	0.922	0.169	0.484
	DIIVINE	0.817	0.366	0.417
	Two-Stream	0.848	/	/
	VCRNet	0.886	0.463	0.542
	Ours	0.931	0.482	0.503

Train \ Test	Method	CSIQ	LIVE	LIVE-C
TID2013	BRISQUE	0.568	0.724	0.109
	HOSA	0.609	0.844	0.244
	WaDIQaM-NR	0.683	0.805	0.009
	HyperIQA	/	/	/
	DIIVINE	0.583	0.714	0.235
	Two-Stream	0.613	0.566	0.014
	VCRNet	0.721	0.822	0.907
	Ours	0.727	0.860	0.304

4.4 消融实验

在本节中，我们进行了消融研究，以检查每个模块的贡献。消融实验在 LIVE 和 LIVE-C 数据集上进行了训练和测试。首先，我们通过使用预训练参数验证了我们的粗粒度感知构建（CPC）模块的有效性，Resnet 50 继承了该参数以直接生成分数。然后，在骨干网中增加了多尺度特征提取模块进行测试。在 CPC 和 MFE 的基础上，通过修改骨干网络，增加了一个细粒度的交互式学习模块（FIL）模块进行测试。实验结果可以在图 4.1 中观察到。

从结果来看，一个简单的预训练任务(CPC)可以使模型达到相对较好的效果。这表明继承到分类任务明显优于将 Resnet50 中的参数直接应用于回归任务。在

CPC+MFE 中,通过利用多尺度方法,模型可以更深入地学习图像的局部失真信息,从而显著提高模型的整体性能。具体而言, LIVE 数据集上的 SRCC 值从 0.926 增加到 0.952。进一步, CPC+MFE+FIL 模型表明,细粒度交互式学习模块可以更渐进地联合收割机结合图像的失真信息,有效地将先前学习到的多尺度特征分为质量特征和失真特征,并让两者相互学习。添加 FIL 模块后, LIVE 数据集中的 SRCC 值达到 0.976。以上分析表明,这三个模块都有助于提高性能,并证实了它们的有效性和将它们结合起来的必要性。

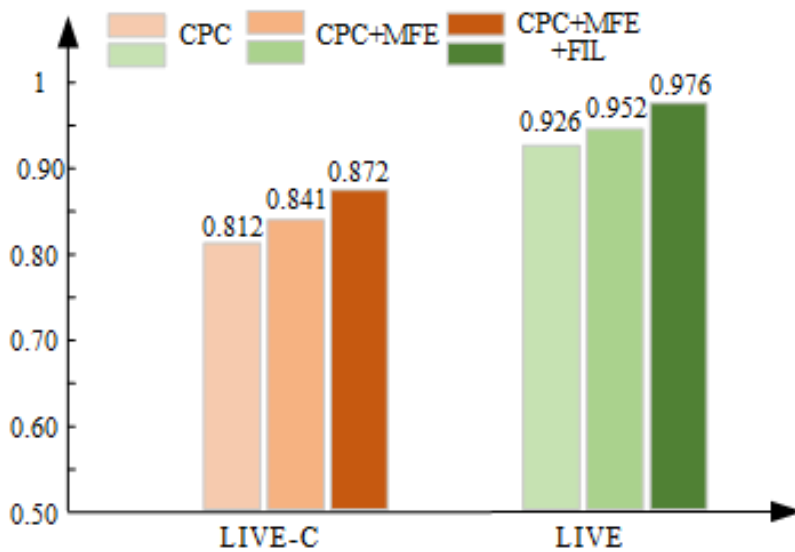


图 4.1 消融实验

此外,为了进一步探索 FIL 模块中自注意操作的数量,我们设置了不同的操作时间(即, $times = 1, 2, 3, 4$)。如表 4.3 所示,随着运算次数的逐渐增加,失真类型与质量排名之间的交互学习效果越来越好。当时间=3 时,模型的性能是最佳的,之后性能下降。

表 4.3 交互学习效果比较表

Times	CSIQ	LIVE	TID2013	KonIQ-10K	LIVE-C
1	0.937	0.968	0.863	0.905	0.850
2	0.930	0.971	0.853	0.904	0.860
3	0.956	0.976	0.877	0.912	0.872
4	0.931	0.968	0.827	0.909	0.862

4.5 细粒度图像实验

细粒度图像质量评估是对具有相同失真水平的图像质量的预测^[60]。一个好的模型不仅应该在具有较大质量变化的数据集上表现良好，而且应该在具有相同失真水平的数据集上表现良好。因此，本实验的目的是评估该算法在细粒度数据集中的性能。我们根据每个数据集的平均主观得分将五个数据集（LIVE，CSIQ，TID 2013，LIVE-C 和 KonIQ 10 K）划分为高级子数据集和低级子数据集。为了比较，我们选择了性能更好的 HyperIQA^[53]和 DBCNN^[14]。

表 4.4 细粒度图像实验

Quality	Method	CSIQ		LIVE		TID2008	
		SRCC	PLCC	SRCC	PLCC	SRCC	PLCC
High-quality	DBCNN	0.859	0.862	0.889	0.899	0.528	0.537
	HyperIQA	0.850	0.853	0.898	0.904	0.605	0.602
	Ours	0.856	0.862	0.921	0.920	0.611	0.612
Low-quality	DBCNN	0.823	0.845	0.760	0.743	0.776	0.840
	HyperIQA	0.862	0.886	0.652	0.685	0.821	0.881
	Ours	0.956	0.946	0.896	0.925	0.812	0.862

Quality	Method	KonIQ-10K		LIVE-C	
		SRCC	PLCC	SRCC	PLCC
High-quality	DBCNN	0.819	0.816	0.679	0.677
	HyperIQA	0.789	0.782	0.671	0.676
	Ours	0.799	0.797	0.726	0.734
Low-quality	DBCNN	0.705	0.751	0.713	0.752
	HyperIQA	0.719	0.767	0.691	0.724
	Ours	0.784	0.832	0.732	0.767

如表 4.4 所示，我们的算法性能优于比较的方法。在高质量数据集中，算法的性能略低于 CSIQ 和 KonIQ-10 K 上的 DBCNN 方法。在低质量数据集中，我们的算法在 TID 2013 上获得了 0.812 的 Sc 分数，略低于 HyperIQA 获得的 0.821 分数。与^[63]类似，与单个数据集的实验结果相比，这三个指标在细粒度图像质量评估中

的性能下降。与 HyperIQA 和 DBCNN 相比，本文的算法不仅融合了全局和局部特征，而且在后续任务中更注重相对质量和失真类型的重要性，这恰恰有助于模型在细粒度数据集评估中取得更好的效果。

4.6 细粒度交互学习的特征可视化

经过预训练阶段，该模型可以达到约 99% 的失真类型的分类结果。可以观察到，如果仅在对失真类型进行分类的任务中，则识别失真类型不是具有挑战性的任务。此外，重要的是要知道这种识别能力是否也适用于微调阶段。在这个实验中，我们将研究模型在微调阶段是否仍然可以使用与预训练阶段相同的失真提取器提取失真信息。

在微调阶段，模型通过失真类型提取器提取失真类型信息，并在细粒度交互学习模块中不断细化。因此，如果模型能够真正提取出合理的信息，那么在细粒度交互学习阶段之后的特征应该包含明显的失真类型信息。



图 4.2 高清图像与五种不同的失真图像

在这个实验中，我们从 LIVE 数据集中选择了不同类型的失真图像（包括快速衰落（FF），高斯模糊（GB），JPEG 2000（JP2K），JPEG 和白色噪声（WN）），如图 4.2 所示。然后，我们使用在 LIVE 数据集上训练的网络测试了这些图像，并提取了测试期间细粒度交互式学习模块生成的参数。这使我们能够观察其区分失真类型的能力。图 4.3 显示了五种不同失真类型的最后一层特征提取器的输出值分

布。虽然不同类型的特征值有一些重叠，但仍然可以区分主要的数据聚类。这表明细粒度交互式学习模块中的失真提取器成功地区分了不同的失真类型，并为它们分配了不同的权重。这表明在预训练阶段训练的失真类型提取器确实有效地在微调阶段整合失真类型信息，并有助于质量评估任务。

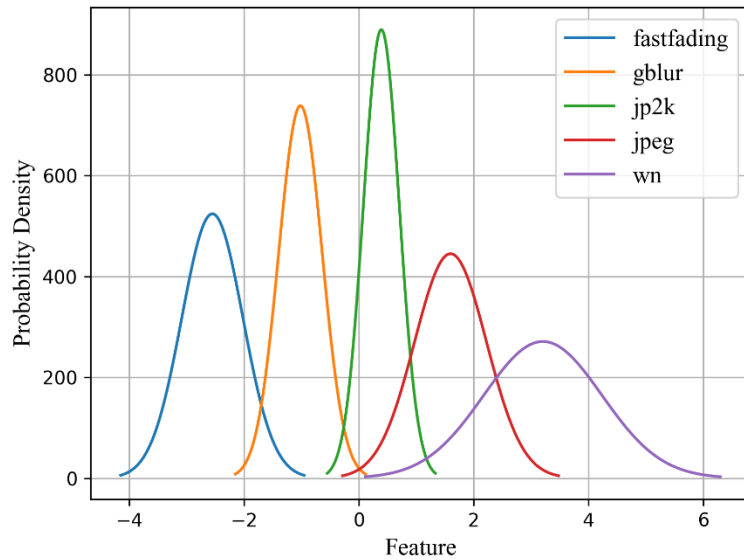


图 4.3 细粒度交互学习的可视化结果

4.7 算法生成的失真评估可视化

对于 BIQA，最重要的应用之一是评估算法生成的失真^[61]。为此，在本实验中评估算法生成的图像。由于图像去模糊是低层视觉的一个主导方向，本节以图像去模糊为例。具体来说，我们从不同的场景中拍摄了一组运动模糊图像。这些图像取自各种实际拍摄情况，包括移动物体，不稳定的相机镜头等。然后，选择两种代表性算法^[62,63]进行去模糊。最后，利用该算法预测这些去模糊图像的质量。

图 4.4 显示了实验结果。第一列显示模糊图像，而第二列和第三列显示相应的去模糊图像。通过仔细比较去模糊图像和模糊图像，可以观察到这两种去模糊算法在提高模糊图像的质量方面是有效的。同时，本文的算法可以提供关于与人类视觉感知一致的算法生成的图像的判断。实验结果全面展示了本文的算法在评估算法产生的失真的有效性。

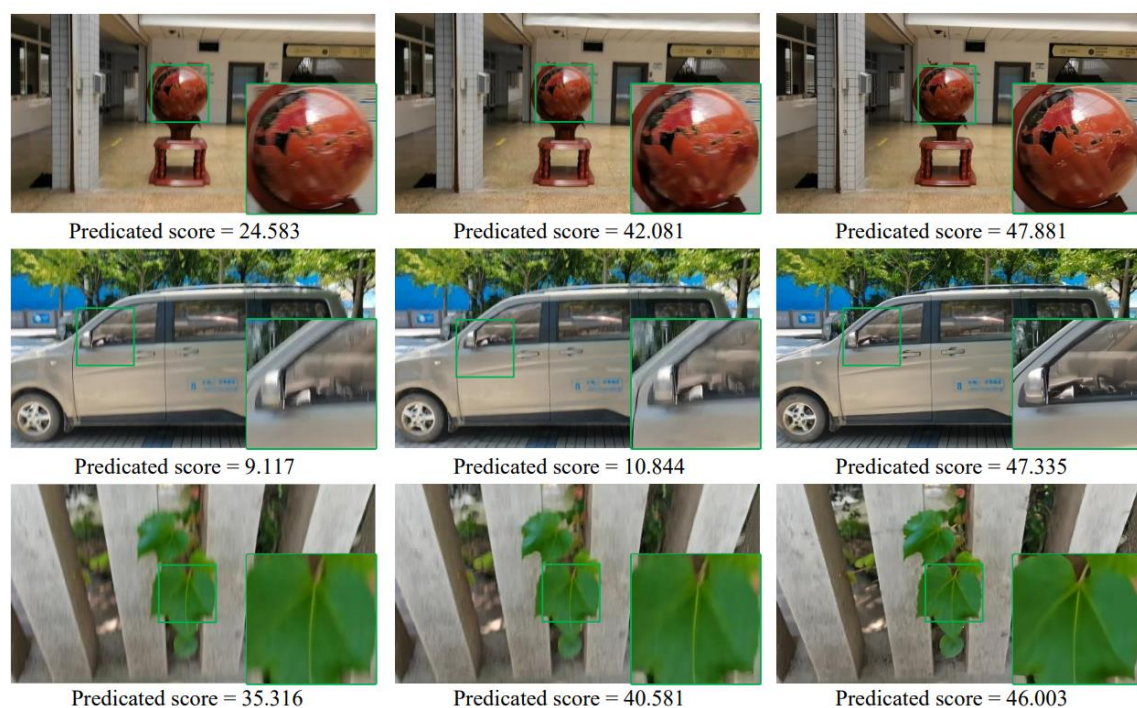


图 4.4 去模糊算法生成图像的质量分数对比

4.8 本章小结

在本章中，我们全面展示了我们提出的图像质量评价算法在多个数据库上的实验结果。通过在单一数据库（CSIQ、LIVE、TID2013、KonIQ-10k、Live-C）上的实验，我们验证了算法在不同数据集上的性能表现。此外，我们进行了跨库实验，进一步验证了算法的泛化能力。消融实验揭示了算法各部分对最终性能的贡献，而细粒度图像实验则展示了算法在处理各种图像类型时的鲁棒性。通过细粒度交互学习的特征可视化实验，我们深入探讨了算法的内在机制，为其性能提升提供了解释。最后，通过算法生成的失真评估可视化实验，我们直观地展示了算法对图像失真的有效评估能力。综合这些实验结果，我们强有力地证明了我们提出的算法在图像质量评价领域的优越性，为进一步的研究和应用提供了坚实的基础。

第 5 章 系统设计与实现

当今社会，随着信息技术的迅猛发展和智能化应用的日益普及，图像处理技术作为其中的重要组成部分，正日益受到广泛关注与应用。本章旨在深入探讨和实现基于 Flask、Vue 和 Nginx 的系统设计与开发，特别关注于图像质量评价、算法框架展示、去模糊算法评估和美学评价等核心功能。

5.1 系统设计实现相关技术介绍

5.1.1 Flask 框架介绍

Flask 是一个由 Armin Ronacher 开发的轻量级 Python Web 应用框架，自 2010 年发布以来，因其简单、易学和易用而广受欢迎。这个框架特别适合构建小到中型的 Web 应用程序和 RESTful API。Flask 提供了基础功能如路由、模板渲染、以及请求和响应的处理，同时也支持开发者通过各种插件和扩展来增强其功能。Flask 的核心设计理念是保持简洁和灵活，它提供了高度的自由度，允许开发者根据项目需求选择合适的工具和组件。Flask 不仅拥有丰富的文档资源，还有一个活跃的开发社区，使其在 Python Web 开发领域中非常受欢迎，广泛应用于多种 Web 应用和服务的构建。如图 5.1 所示。



图 5.1 Flask

5.1.2 Vue 框架介绍

Vue.js 是一个广受欢迎的前端 JavaScript 框架，专门用于构建用户界面。其核心特点包括响应式数据绑定、组件化开发、虚拟 DOM 技术，以及简洁的 API

设计。响应式数据绑定确保了数据变化能够即时反映在视图上，而组件化开发则支持页面结构的清晰划分、便于维护和复用。利用虚拟 DOM 技术，Vue.js 提升了页面渲染的效率，并减少了不必要的 DOM 操作，从而优化了应用的性能。整体来看，Vue.js 是一个轻量级、灵活且易于掌握的框架，非常适合用于开发各种规模的 Web 应用和跨平台应用。如图 5.2 所示。



图 5.2 Vue

5.1.3 nginx 框架介绍

Nginx 是一个高性能的开源 Web 服务器。它的设计主要是为了解决 C10k 问题，即如何有效处理成千上万的并发连接。如图 5.3，Nginx 在处理高并发情况时能够展现出卓越性能。其特点包括占用资源少、高并发支持、高度可靠性和灵活的配置选项。由于这些特性，Nginx 已经成为许多大型网站和应用的首选服务器软件之一。

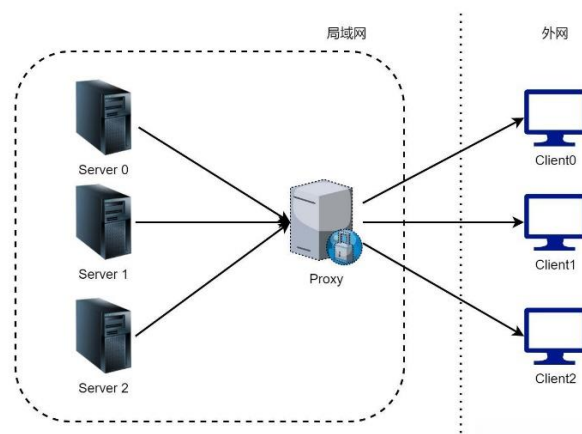


图 5.3 Nginx 服务器

5.2 系统的功能介绍

该框架为用户提供了一体化的图像处理解决方案，包括图像质量评价、算法框架展示、去模糊算法评价和美学评价等功能。用户可以通过上传图像，快速获取质量评价结果，并深入了解我们的算法框架及其应用。同时，我们提供了先进的去模糊算法评价工具，帮助用户评估和优化图像去模糊效果。美学评价功能则为用户提供了专业的图像审美建议，助力用户提升图像的艺术品质。该框架旨在为用户提供全方位、高效的图像处理服务，满足个人和专业用户的各种需求。

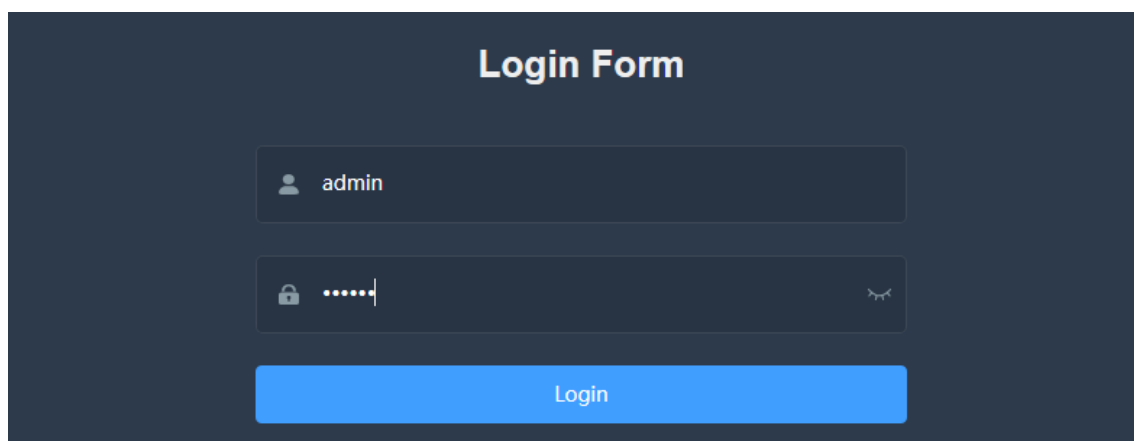


图 5.4 登录页面

5.2.1 图像质量评价功能

图像质量评价模块允许用户上传图像，并根据其选择的评价算法（可供选择的多个算法之一）获取详尽的图像分析结果。如图 5.5，用户首先选择图像并确定所需的评价算法，然后通过点击“预测”按钮，系统将呈现图像的质量分数、边缘检测结果以及可能的综合失真因素的直方图，如图 5.6。这种详细的分析为用户提供了全面的图像质量信息，包括定量评价和视觉分析，帮助他们更好地理解图像特征并做出相应的处理决策。



图 5.5 图像质量评价页面

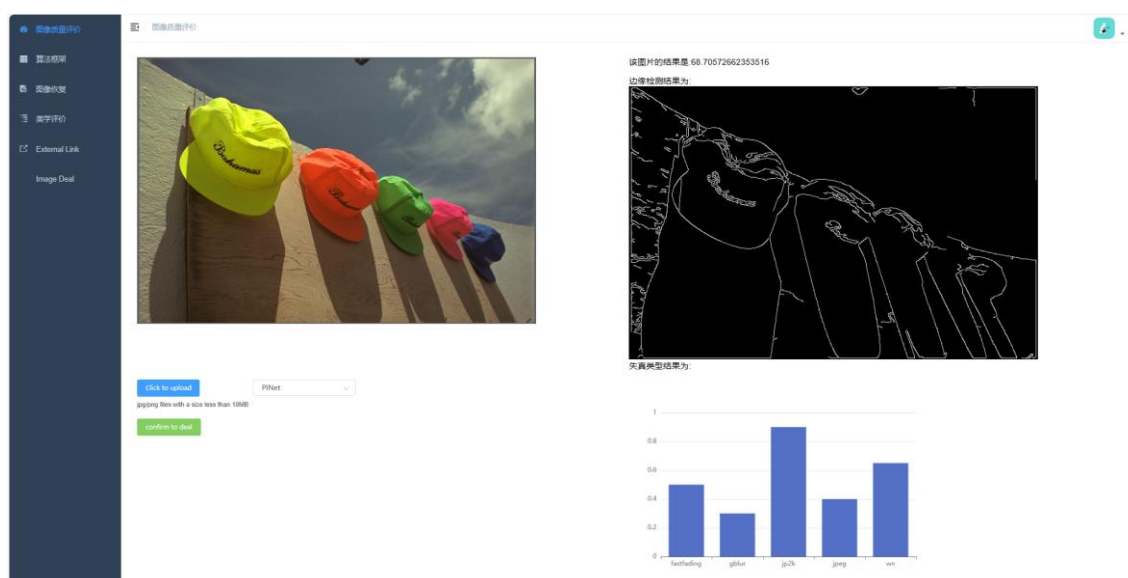


图 5.6 图像质量评价结果展示

边缘检测在图像处理中具有多重作用。首先，它可以用于物体识别和分割，通过找到图像中物体的边界，进而实现物体的区分和定位。其次，边缘也是图像中重要的特征之一，因此边缘检测可用于特征提取，帮助识别和描述图像中的内容。此外，对图像进行边缘检测还可以用于图像增强，即强调图像中的轮廓和细节，使图像更加清晰和易于理解，如图 5.7 所示。



图 5.7 边缘检测上传的原始图像

当进行边缘检测时，算法通常会计算图像中每个像素的灰度级别的梯度，并寻找梯度变化最大的地方，这些地方被认为是边缘的位置。常见的边缘检测算法包括 Sobel、Prewitt 和 Canny。Sobel 和 Prewitt 算子主要基于图像像素灰度的一阶导数来识别边缘，它们通过分别计算图像在水平和垂直方向的梯度，然后合并这两个方向的梯度来确定边缘的位置。相比之下，Canny 算子是一种更为复杂的边缘检测方法，它首先通过高斯滤波器对图像进行平滑处理，以消除噪声，然后计算图像的梯度来识别边缘方向和强度，接着通过非极大值抑制技术细化边缘，最后应用双阈值方法来确定真实的边缘，从而更准确地检测和链接边缘。检测结果如图 5.8 所示



图 5.8 边缘检测结果

当处理真实失真图像时，算法需要考虑到多种失真因素的影响，而这些图像并没有自己的高清原始版本，也不是由特定的失真方法生成的。因此，为了更好地模拟这些真实失真，我们设计利用综合失真来拟合真实失真图像的效果。

失真类型结果为:

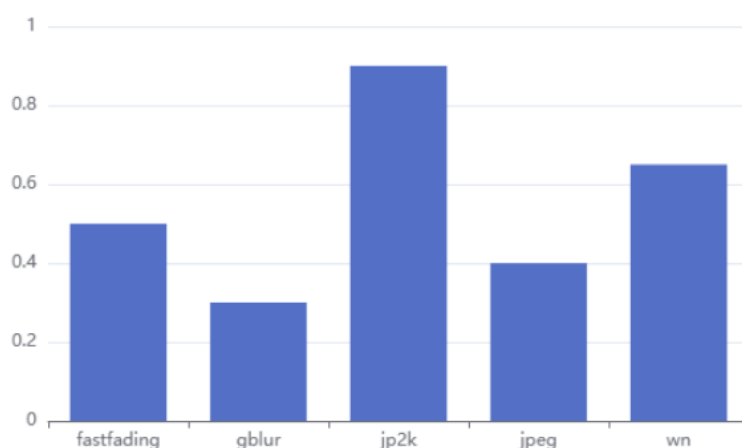


图 5.9 失真类型甄别结果

这项功能旨在帮助用户识别图像中可能存在的五种综合失真类型及其占比程度。通过直方图，用户可以清晰地了解图像中各种失真类型的相对影响。这种方法使得算法能够更有效地处理真实场景中的失真图像，因为它考虑了多种可能的失真因素，并尝试在算法中进行综合模拟。这样一来，算法在处理真实失真图像时可以提高准确性和效率。

5.2.2 算法框架展示功能

随着人工智能和机器学习的发展，算法的应用越来越广泛。然而，理解和掌握这些算法的框架结构对于学习者来说可能是一项挑战。因此，为了帮助用户更好地理解不同算法的结构和工作原理，某网站实现了算法框架展示功能。

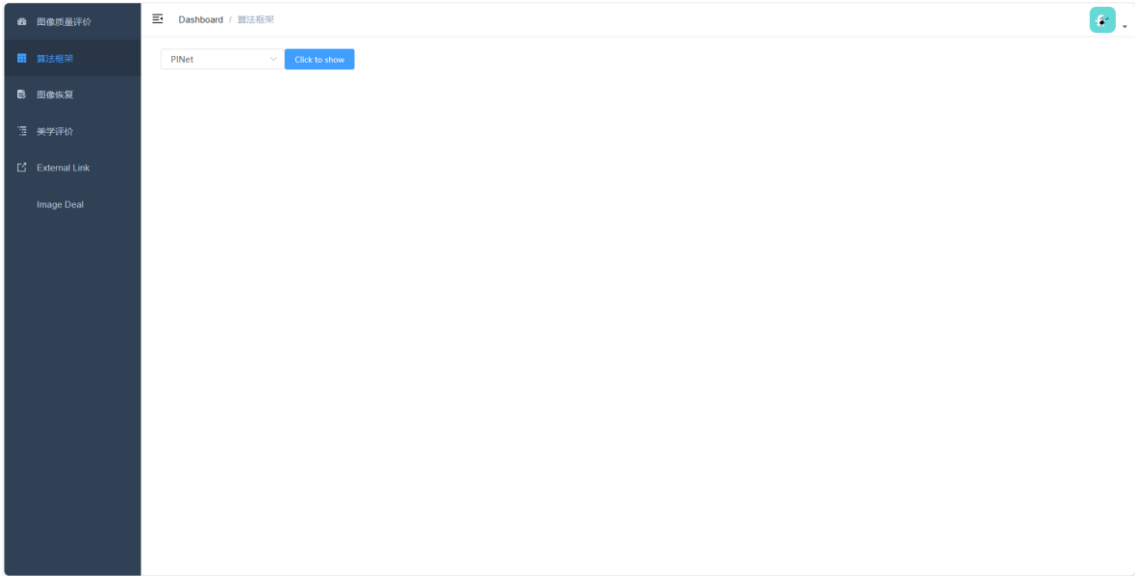


图 5.10 算法框架选择

该功能的目的是为用户提供一个直观的方式来了解不同算法的框架结构，帮助他们更有效地学习和掌握这些算法。通过查看算法的框架图，用户可以更清晰地理解算法的组成部分和其运作方式。

用户可以在网站上浏览各种算法的列表，并选择他们感兴趣的算法。一旦选择了特定的算法，系统将显示该算法的框架图。用户可以通过缩放、拖动等方式来查看框架图的不同部分，以便更详细地了解算法的结构。

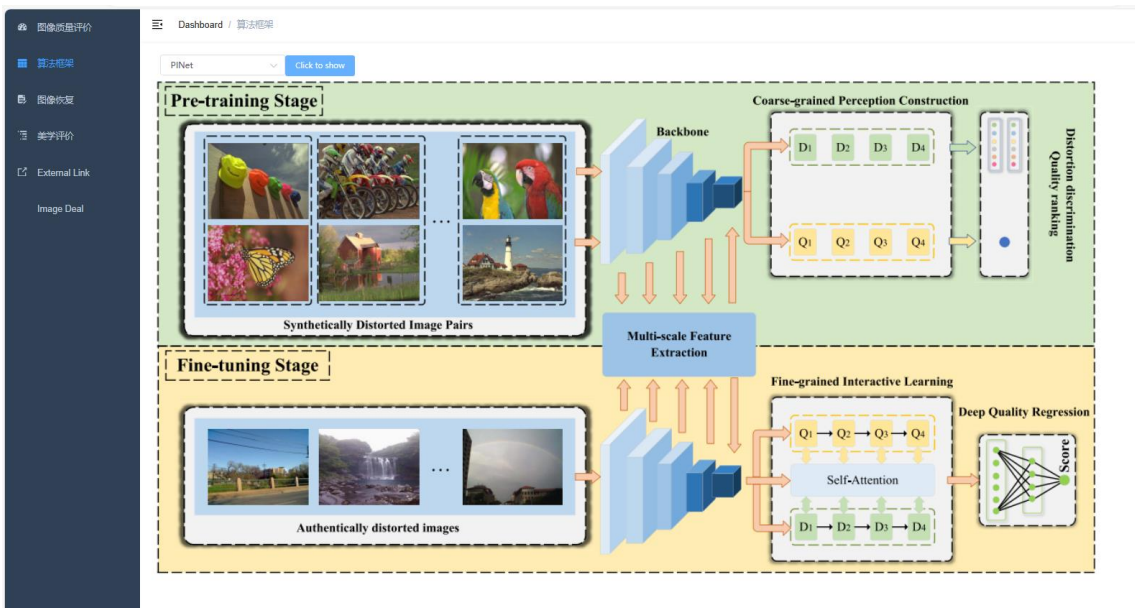


图 5.11 算法框架展示

通过使用这个功能，用户可以更加直观地理解不同算法的框架结构，从而加深对算法原理的理解。这有助于他们在学习和应用算法时更加游刃有余，并且有助于提高他们在实践中的应用能力。

5.2.3 去模糊算法评价功能

随着数字图像处理领域的发展，去模糊算法成为了重要的研究方向。然而，不同的去模糊算法可能会产生不同的效果，而图像质量评价则可以帮助用户判断去模糊结果的好坏程度。

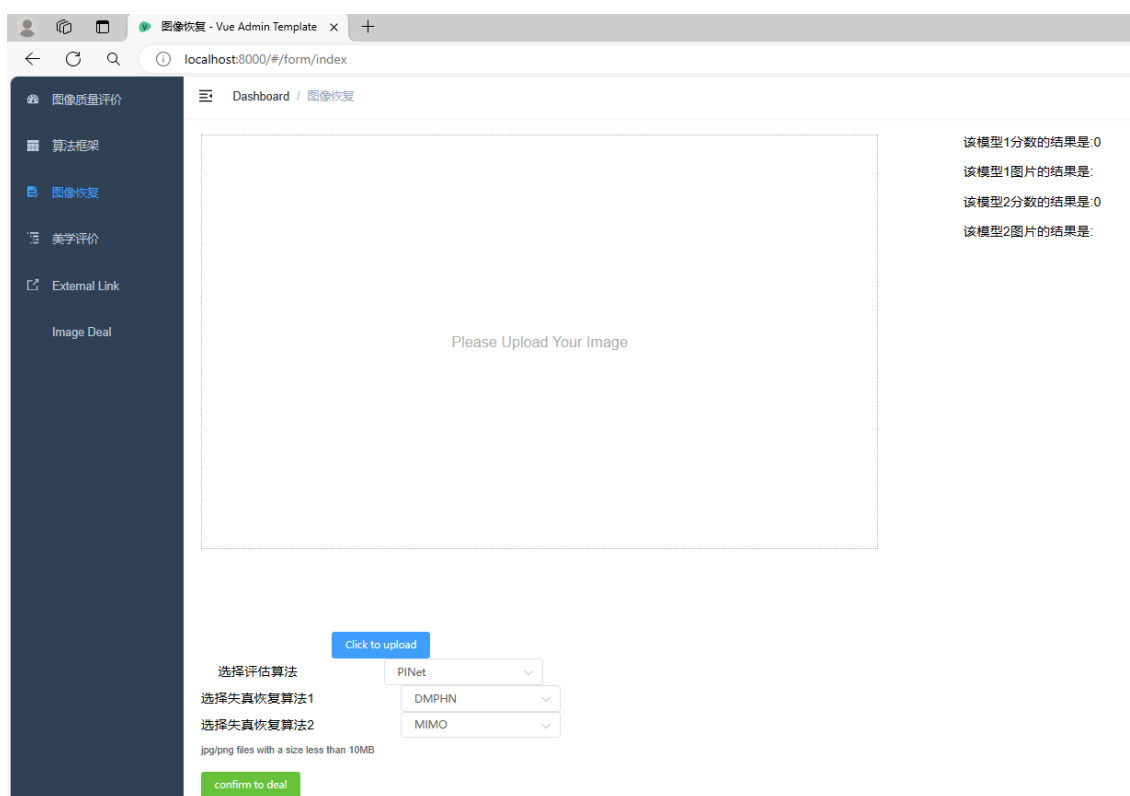


图 5.12 去模糊页面

该功能的目的是让用户能够直观地比较不同去模糊算法生成的图像质量，以便选择最适合其需求的算法。通过图像质量评价算法的分数评价，用户可以更好地了解 and 比较不同算法的效果。

用户首先需要选择一张图像进行上传，这张图像应该是需要进行去模糊处理的原始图像。然后在提供的算法列表中选择两种不同的去模糊算法，以及一种图像质量评价算法。之后，用户点击“预测”按钮后，网站将会对选择的原始图像使用用户选择的两种不同去模糊算法进行处理，生成两张去模糊后的图像。同时，选定的图像质量评价算法会对这两张去模糊图像进行评分。

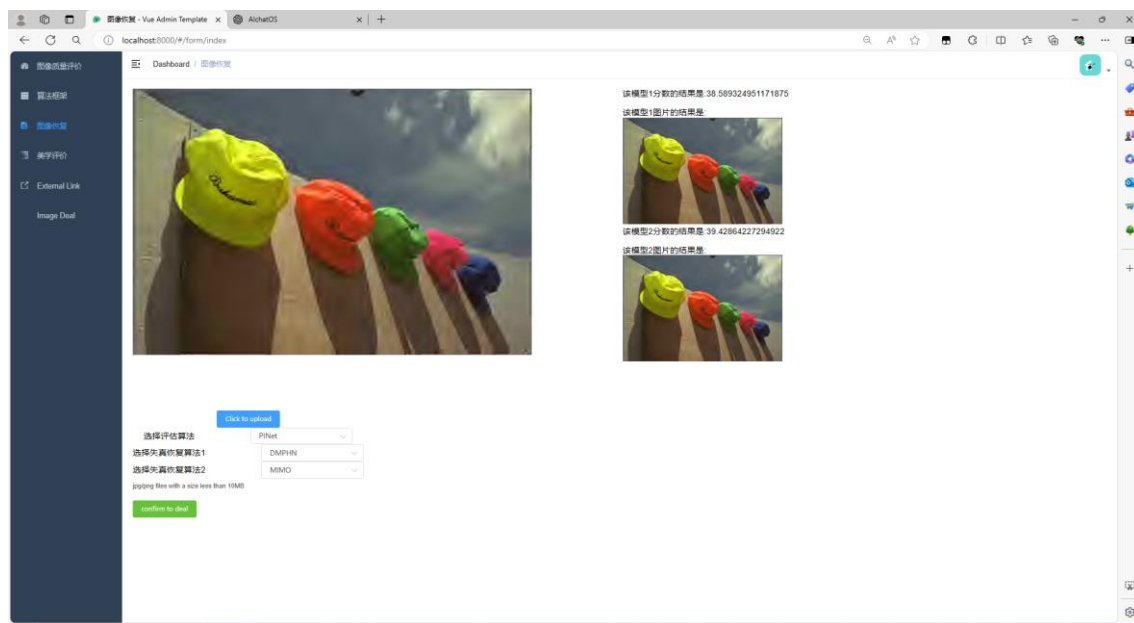


图 5.13 去模糊功能展示

网站会将两种去模糊算法生成的两张去模糊图像以及它们的图像质量评价分数展示给用户。用户通过对比两张去模糊图像的分数评价，可以直观地了解两种算法的优劣程度。这有助于他们在实际应用中选择合适的去模糊算法，从而获得更好的图像处理效果。

5.2.4 美学评价功能

该功能旨在让用户能够通过上传图像并选择美学算法，快速了解图像的美学质量。通过生成图像的美学分数，用户可以客观地评价图像的美感，为后续的处理或决策提供参考。

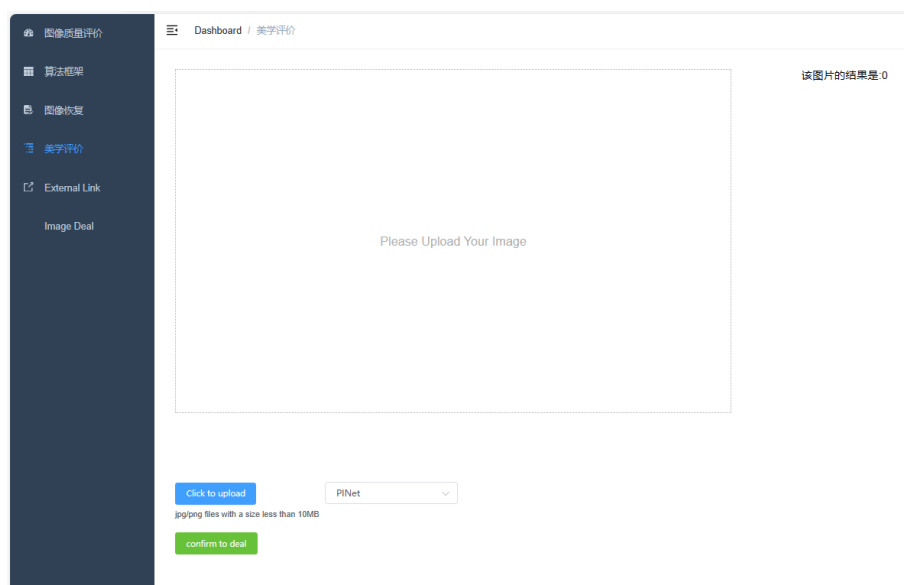


图 5.14 美学图像选择

用户首先选择一张图像进行上传，这张图像是需要进行美学评价的原始图像。然后从提供的美学算法列表中选择一种用于评价图像美学质量的算法。这些算法可能包括基于颜色、构图、纹理等方面的美学评价指标。之后点击“预测”按钮后，系统会对上传的图像使用选定的美学算法进行分析，然后生成图像的美学分数。网站将会显示生成的图像美学分数，让用户直观地了解图像的美学质量。分数越高代表图像美学质量越好，反之则代表美学质量较低。用户通过该功能可以快速、直观地了解上传图像的美学质量，而无需专业知识。这有助于用户在美学分析、图像处理等领域做出更加准确的决策，并提升图像处理的效率和质量。

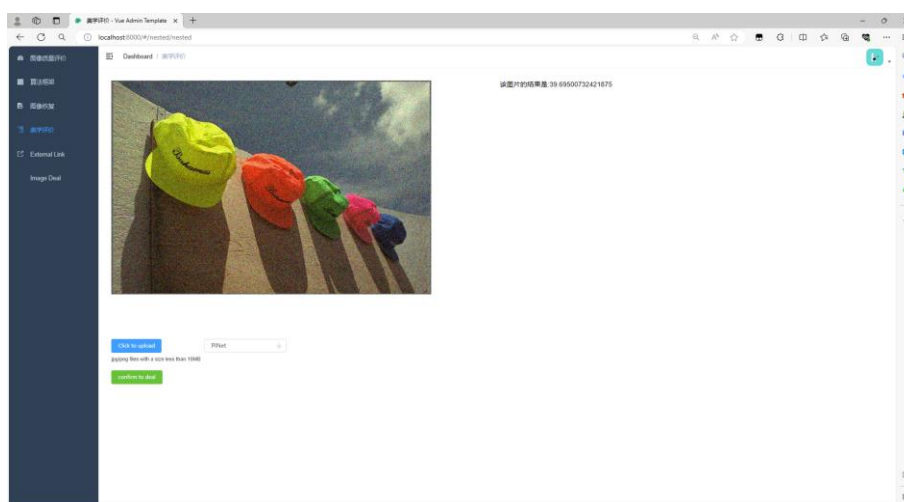


图 5.15 美学评价功能展示

5.3 本章小结

在本章中，我们详细介绍了系统的设计与实现，着重讨论了所采用的技术和系统功能。首先，我们深入探讨了系统的实现相关技术，使用的框架包括但不限于如 Flask、Vue 和 Nginx。这些框架不仅在技术上得到了广泛的应用和认可，更重要的是它们为系统的搭建提供了高效、灵活的支持，为后续功能的实现奠定了坚实的基础。

其次，我们详细介绍了系统的功能模块。首要功能是图像质量评价，其中包括图像质量评价、图像边缘检测以及图像真实类型分类。这些功能的实现不仅提供了对图像质量的全面评估，还为用户提供了更深层次的图像分析。除此之外，我们还实现了算法框架展示功能，通过直观的展示方式向用户展示算法的工作原理和流程。另外，去模糊算法评价功能为用户提供了针对模糊图像的评价与改进方案，增强了系统的实用性和用户体验。最后，美学评价功能为用户提供了对图像美学质量的评估，使得系统在艺术领域的应用更加全面。

通过本章的设计与实现，我们不仅构建了一个功能强大的系统，还为用户提供了多样化的图像处理和评价功能。系统的设计考虑了用户需求和技术实现的可行性，旨在为用户提供便捷、高效的图像处理与评价服务。本网站不仅可以展示本文中的算法，还能够用于以后的学术展示。

第6章 总结与展望

6.1 主要工作与创新点

在核心算法方面，本研究首次提出了一种结合粗粒度感知构建和细粒度交互学习的盲图像质量评价方法。主要工作包括开发了一个多尺度特征提取模块，有效地从不同层次捕获图像的本质特征。在此基础上，我们设计了粗粒度感知构建模块，能够执行失真歧视和质量排名任务，为图像质量评价提供了初步的直观感知。此外，细粒度交互式学习模块的引入，模拟了人脑处理图像失真信息的复杂机制，通过注意力机制深化了特征之间的互动，从而实现了更加精确和细致的图像质量判断。这一创新结构不仅提高了评价的准确性，也增强了模型对于不同图像失真类型的适应能力和泛化性。

在系统设计与实现方面，通过运用现代 Web 技术栈——Flask 和 Vue，我们成功构建了一个功能全面的图像质量评价网站。该网站不仅支持基本的图像质量评估，还展示了算法框架，并包含了去模糊算法评价和图像美学评价等高级功能。这种实践的应用不仅使得图像质量评价更加便捷和直观，也为研究人员和普通用户提供了一个互动平台，用于探索和评估不同的图像处理算法。此外，该系统的开发体现了理论与实际应用的有效结合，为图像处理技术的推广和进一步的研究提供了一个实验和展示的平台。

6.2 后续研究工作展望

在未来的研究中，我们将致力于进一步改进粗粒度感知构造模块和细粒度交互式学习策略，以提高模型对图像失真的感知能力。这将包括探索更有效的方法来捕捉人类感知过程中的细微差异，以及优化注意机制和特征空间映射方法。此外，我们还计划扩展研究范围，通过使用更多样化的数据集和进行跨领域的数据增强技术，进一步增强模型的泛化能力。最终，我们将尝试将这些技术应用于实际场景中，并评估其在图像编辑软件、图像质量改善系统等实际应用中的效果和性能。通过这些努力，我们期望能够为图像质量评价领域的发展做出更大的贡献。

参考文献

- [1] Moorthy AK, Bovik AC. A two-step framework for constructing blind image quality indices. *IEEE Signal Process. Lett.*, 2010, 17(5): 513–516.
- [2] Mittal A, Moorthy A, Bovik A. No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.*, 2012, 21(12): 4695–4708.
- [3] Saad MA, Bovik AC, Charrier C. Blind image quality assessment: A natural scene statistics approach in the DCT domain. *IEEE Trans. Image Process.*, 2012, 21(8): 3339–3352.
- [4] Ye P, Kumar J, Kang L, et al. Unsupervised feature learning framework for no-reference image quality assessment. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2012, 1098–1105.
- [5] Jiang Q, Shao F, Lin W, et al. Optimizing multistage discriminative dictionaries for blind image quality assessment. *IEEE Trans. Multimedia*, 2018, 20(8): 2035–2048.
- [6] 方玉明, 睦相杰, 鄢杰斌, 等. 无参考图像质量评价研究进展[J]. *中国图象图形学报*, 2021, 26(2): 265-286.
- [7] 王成, 刘坤, 杜砾. 全参考图像质量指标评价分析[J]. *现代电子技术*, 2023, 46(21): 39-43.
- [8] 冯晨曦, 张地, 叶龙. 基于多视口自适应融合的全景图像质量评价[J/OL]. *北京航空航天大学学报*, 2024, 1(14): 1001-5965.
- [9] 史晨阳, 卞仕磊, 江本赤, 等. 基于局部色貌的无参考彩色模糊图像质量评价方法[J]. *光学技术*, 2023, 49(05): 615-622..
- [10] 刘义鹏, 吕雅俊, 钟琦, 等. 基于双分支加权卷积神经网络的视网膜图像质量评价方法[J]. *高技术通讯*, 2023, 33(4): 352-359.
- [11] 陆绮荣, 丁昕, 梁雅雯. 基于双线性特征融合方法的畸变失真图像质量评价[J]. *中国科技论文*, 2023, 18(03): 259-291.
- [12] 于天河, 柳梦瑶. 基于人眼视觉系统的图像质量评价方法[J]. *北京邮电大学学报*, 2023, 46(02): 129-136.
- [13] 惠子薇, 何坤, 冯犇, 等. 基于视觉特性的图像质量评价[J]. *计算机工程*, 2023, 49(07): 189-195.

- [14] Zhang W, Ma K, Yan J, et al. Blind image quality assessment using a deep bilinear convolutional neural network. *IEEE Trans. Circuits Syst. Video Technol.*, 2020, 30(1): 36–47.
- [15] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, 1–14.
- [16] Zhang W, Ma K, Zhai G, et al. Learning to blindly assess image quality in the laboratory and wild, in *Proc. IEEE Int. Conf. Image Process.*, 2020, 111–115.
- [17] Fang Y, Zhu H, Zeng Y, et al. Perceptual quality assessment of smartphone photography. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2020, 3674–3683.
- [18] Lv X, Xiang T, Yang Y, et al. Blind Dehazed image quality assessment: A deep CNN-based approach, *IEEE Trans. Multimedia*, 2023.
- [19] Li D, Jiang T, Lin W, et al. Which has better visual quality: The clear blue sky or a blurry animal? *IEEE Trans. Multimedia*, 2019, 21(5): 1221–1234.
- [20] Ma R, Wu Q, Ngan KN, et al. Forgetting to remember: A scalable incremental learning framework for cross-task blind image quality assessment. *IEEE Trans. Multimedia*, 2023, 25, 8817–8827.
- [21] Sun W, Min X, Tu D, et al. Blind quality assessment for in-the-wild images via hierarchical feature fusion and iterative mixed database training, *IEEE J. Sel. Topics Signal Process*, 2023.
- [22] Liu X, van de Weijer J, Bagdanov AD. RankIQA: Learning from rankings for no-reference image quality assessment, *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, 1040–1049.
- [23] Golestaneh S, Dadsetan S, Kitani K. No-reference image quality assessment via transformers, relative ranking, and self-consistency, *Proc. IEEE Wint. Conf. Appl. Comput. Vis. (WACV)*, 2022, 1220–1230.
- [24] Wang X, Xiong J, Lin W, et al. Visual interaction perceptual network for blind image quality assessment. *IEEE Trans. Multimedia*, 2023, 25, 8958–8971.
- [25] Zhou Y, Sun Y, Li L, et al. Omnidirectional image quality assessment by distortion discrimination assisted multi-stream network. *IEEE Trans. Circuits Syst. Video Technol.*, 2022, 32(4): 1767–1777.

-
- [26] Zhou M, Lan X, Wei X, et al. An end-to-end blind image quality assessment method using a recurrent network and self-attention, *IEEE Trans. Broadcast.*, 2023, 69(2): 369–377.
- [27] Zhu H, Li L, Wu J et al. A full-reference image quality assessment method via deep meta-learning and conformer, *IEEE Trans. Broadcast.*, 2023.
- [28] Liao Xingran, Wei Xuekai, Zhou Mingliang et al. Full-reference image quality assessment: Addressing content misalignment issue by comparing order statistics of deep features, *IEEE Trans. Broadcast.*, 2023.
- [29] Paudyal P, Battisti F, Carli M. Reduced reference quality assessment of light field images, *IEEE Trans. Broadcast.*, 2019, 65(1): 152-165.
- [30] Lan X, Zhou M, Xu X, et al. Multilevel feature fusion for end-to-end blind image quality assessment, *IEEE Trans. Broadcast.*, 2023, 69(3): 801-811.
- [31] Jiang Q, Gu Y, Li C, et al. Underwater image enhancement quality evaluation: Benchmark dataset and objective metric, *IEEE Trans. Circuits Syst. Video Technol.*, 2022, 32(9): 5959–5974.
- [32] Jiang Q, Shao F, Gao W, et al. Unified no-reference quality assessment of singly and multiply distorted stereoscopic images, *IEEE Trans. Image Process.*, 2019, 28(4): 1866–1881.
- [33] Li L, Zhu H, Yang G, Qian J, et al. Referenceless measure of blocking artifacts by Tchebichef kernel analysis. *IEEE Signal Process. Lett.*, 2014, 21(1): 122–125.
- [34] Zhang L, Zhang L, Bovik A, et al. A feature-enriched completely blind image quality evaluator. *IEEE Trans. Image Process.*, 2015, 24(8): 2579–2591.
- [35] Xu J, Ye P, Li Q, et al. Blind image quality assessment based on high order statistics aggregation. *IEEE Trans. Image Process.*, 2016, 25(9): 4444–4457.
- [36] Zhang W, Li D, Ma C, et al. Continual learning for blind image quality assessment. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2023, 45(3): 2864–2878.
- [37] Liu M, Huang J, Zeng D, et al. A multiscale approach to deep blind image quality assessment. *IEEE Trans. Image Process.*, 2023, 32(1): 1656–1667.
- [38] Liu Y, Ni Z, Wang S, et al. High dynamic range image quality assessment based on frequency disparity. *IEEE Trans. Circuits Syst. Video Technol.*, 2023, 33(8): 4435–4440.
- [39] Wang Z, Bovik A, Sheikh H, et al. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.*, 2004, 13(4): 600–612.

- [40] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, 770–778.
- [41] Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2009, 248–255.
- [42] Sheikh H, Sabir M, Bovik AC. A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Trans. Image Process., 2006, 15(11): 3440–3451.
- [43] Larson EC, Chandler DM. Most apparent distortion: Full-reference image quality assessment and the role of strategy. J. Electron. Imag., 2010, 9(1): 11006.
- [44] Ponomarenko N, et al. Color image database TID2013: Peculiarities and preliminary results. Proc. Eur. Workshop Vis. Inf. Process. (EUVIP), 2013, 106–111.
- [45] Ghadiyaram D, Bovik AC. Massive Online crowdsourced study of subjective and objective picture quality. IEEE Trans. Image Process., 2016, 25(1): 372–387.
- [46] Hosu V, Lin H, Sziranyi T, et al. KonIQ-10k: An ecologically valid database for deep learning of blind image quality assessment. IEEE Trans. Image Process., 2020, 29, 4041–4056.
- [47] Moorthy AK, Bovik AC. Blind image quality assessment: From natural scene statistics to perceptual quality. IEEE Trans. Image Process., 2011, 20(12): 3350–3364.
- [48] Mittal A, Soundararajan R, Bovik AC. Making a ‘completely blind’ image quality analyzer. IEEE Signal Process. Lett., 2013, 20(3): 209–212.
- [49] Ren H, Chen D, Wang Y, et al. RAN4IQA: Restorative adversarial nets for no-reference image quality assessment. Proc. AAAI Conf. Artif. Intell., 2018, 7308–7314.
- [50] Bosse S, Maniry D, Müller KR, et al. Deep neural networks for no-reference and full-reference image quality assessment. IEEE Trans. Image Process., 2018, 27(1): 206–219.
- [51] Yan Q, Gong D, Zhang Y, et al. Two-stream convolutional networks for blind image quality assessment. IEEE Trans. Image Process., 2019, 28(5): 2200–2211.
- [52] Wu J, Ma J, Liang F, et al. End-to-end blind image quality prediction with cascaded deep neural network. IEEE Trans. Image Process., 2020, 29, 7414–7426.

-
- [53] Su S, et al. Blindly assess image quality in the wild guided by a self-adaptive hyper network. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2020, 3664–3673.
 - [54] Zhang W, Ma K, Zhai G, et al. Uncertainty-aware blind image quality assessment in the laboratory and wild. *IEEE Trans. Image Process.*, 2021, 30, 3474–3486.
 - [55] Song T, Li L, Chen P, et al. Blind image quality assessment for authentic distortions by intermediary enhancement and iterative training. *IEEE Trans. Circuits Syst. Video Technol.*, 2022, 32(11): 7592–7604.
 - [56] Madhusudana PC, Birkbeck N, Wang Y, et al. Image quality assessment using contrastive learning. *IEEE Trans. Image Process.*, 2022, 31, 4149–4161.
 - [57] Pan Z, Yuan F, Lei J, et al. VCRNet: Visual compensation restoration network for no-reference image quality assessment. *IEEE Trans. Image Process.*, 2022, 31, 1613–1627.
 - [58] Zhu H, Li L, Wu J, et al. Generalizable no-reference image quality assessment via deep Meta-learning. *IEEE Trans. Circuits Syst. Video Technol.*, 2022, 32(3), 1048–1060.
 - [59] Gao Y, Min X, Zhu Y, et al. Blind image quality assessment: A fuzzy neural network for opinion score distribution prediction. *IEEE Trans. Circuits Syst. Video Technol.*, 2023.
 - [60] Zhang X, Lin W, Huang Q. Fine-grained image quality assessment: A revisit and further thinking. *IEEE Trans. Circuits Syst*, 2022.
 - [61] Yang W, Wang S, Fang Y, et al. From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2020, 3060–3069.
 - [62] Chen L, Chu X, Zhang X, et al. Simple baselines for image restoration. *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2022, 17–33.
 - [63] Zhang H, Dai Y, Li H, et al. Deep stacked hierarchical multi-patch network for image deblurring. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2019, 5978–5986.

致谢

行文至此，落笔为终，往事千帆过，才发现学生生涯已行至此。时光仍然，回首望去，还曾记得 2020 年的那个雨天，初次踏入校园的时刻。纵使心有不舍，但仍感激万千。

首先，我要感谢母校，为我提供的学习环境和学术资源。学校的图书馆、实验室设施以及各类学术活动为我的研究工作提供了必要的条件和支持。此外，特别感谢学校对我的研究项目的经费资助，使我能够顺利开展研究工作。

在此，我要特别感谢我的导师胡波老师。导师在整个研究过程中给予我悉心的指导和无私的帮助。导师的专业知识、严谨治学的态度以及耐心指导为我解决了许多困难，使我能够在学术道路上稳步前行。从论文的选题，到论文的创新，我们无数次的讨论与沟通，您无数次的指点与纠正，成就了今天的这篇文章。在此感谢您的付出与帮关注，祝愿您生活顺利，身体健康。

同时，感谢我的室友们。他们在我繁忙的学术生活中给予我理解、支持和鼓励，让我能够保持积极的心态和高效的工作状态。特别感谢室友邓洪和室友岳展翼在论文写作过程中对我的帮助和理解。

之后，我衷心感谢孙文琦在我大学四年中所给予的精神支持和帮助。在论文写作的过程中，她的鼓励和支持让我能够坚定地克服种种困难。她的理解和耐心帮助我克服了写作中的障碍，使我能够顺利完成这篇论文。再次感谢她在我学术生涯中的重要贡献和无私支持。

此外，我还要感谢所有在研究工作中提出建议和提供帮助的同学、朋友和家人。他们的意见和支持对我论文的完善起到了重要作用。感谢他们始终在我身后默默支持和鼓励。

总之，没有你们的支持和帮助，我无法完成这篇论文。在此向所有支持者表示最衷心的感谢和敬意。