

图文信息技术

基于底层特征和高级语义的真实失真图像质量评价

王晓红^{1,2}, 庞云杰¹, 麻祥才²

(1.上海理工大学, 上海 200093; 2.上海出版印刷高等专科学校, 上海 200093)

摘要: **目的** 由于现有无参考质量评价方法无法准确判断真实失真图像的质量, 提出一种基于图像底层特征和高级语义提取的真实失真图像质量评价方法。 **方法** 首先根据真实失真图像的底层特征指标进行 k-means 聚类, 在每一类图像中利用深度卷积神经网络 (Deep Convolutional Neural Network, DCNN) 的方法提取图像的一级高级语义特征, 采用多种特征函数对一级高级语义特征进行聚合, 得到图像的二级高级语义特征, 并建立了高级语义特征与平均意见主观分 (Mean Opinion Score, MOS) 的高容量回归器。 **结果** 提出的算法在 KonIQ-10k 图像库预测出的质量分数与对应 MOS 值能达到很高的一致性, Spearman 秩序相关系数 (SROCC) 和 Pearson 线性相关系数 (PLCC) 分别能达到 0.95 和 0.97。 **结论** 提出的算法能够快速且准确地对真实失真图像质量作出评价。

关键词: 真实失真; k-means; 深度卷积神经网络 (DCNN); 高级语义; 图像质量评价

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1001-3563(2020)01-0134-09

DOI: 10.19554/j.cnki.1001-3563.2020.01.021

Real Distorted Images Quality Assessment Based on Image Underlying Features and High-Level Semantics

WANG Xiao-hong^{1,2}, PANG Yun-jie¹, MA Xiang-cai²

(1.University of Shanghai for Science and Technology, Shanghai 200093, China;

2.Shanghai Publishing and Printing College, Shanghai 200093, China)

ABSTRACT: The paper aims to propose a real distortion IQA method based on image underlying features and high-level semantics in view of the fact that the existing non-reference IQA methods can not accurately evaluate the quality of the real distortion image. Firstly, k-means clustering algorithm was performed according to the underlying feature index of the image. Secondly, the deep convolutional neural network (DCNN) was used to extract the first-grade high-level semantics in each group. Then, second-grade high-level semantics that can provide better representation of image features were obtained by performing multiple statistical functions on first-grade high-level semantics. Besides, we established an effective high-capacity regressor with high-level semantics and subjective mean opinion scores (MOS) values of the human eyes. The experimental results showed that the proposed model on the KonIQ-10k image database can predict the quality score effectively and achieve a high consistency with the corresponding MOS value. Spearman order correlation coefficient (SROCC) and Kendall order correlation coefficient (KROCC) can reach 0.95 and 0.97, respectively. The proposed method can quickly and accurately evaluate the quality of real distorted image.

收稿日期: 2019-06-26

基金项目: 上海市教育发展基金会和上海市教育委员会“晨光计划”(18CGB09); “柔版印刷绿色制版与标准化”国家新闻出版署重点实验室资助项目

作者简介: 王晓红 (1971—), 女, 博士, 上海理工大学教授, 上海出版印刷高等专科学校特聘研究员, 主要研究方向为数字图像处理、计算机视觉、图像质量评价、图像增强。

KEY WORDS: real distortion; k-means; deep convolutional neural network (DCNN); high-level semantics; image quality assessment

现实生活中人们无时无刻不在接受着来自四面八方的信息,其中有 80%的信息是以图像的形式呈现的。数字图像在传输过程中,都会在一定程度上产生失真,因此如何有效评价数字图像的质量已经成为数字图像处理领域、计算机视觉领域的热点研究课题之一^[1]。客观的 IQA (Image Quality Assessment) 主要是利用数学和工程的方法,其目的在于寻找一个理想的数学模型对图像进行质量评价,模拟人类视觉系统感知图像的质量,目前主流的 IQA 主要集中于人工合成的失真图像。MITTIAL 等^[2]提出了一种无参考的空间域图像质量评价算法 (Blind/Referenceless Image Spatial Quality Evaluator, BRISQUE),该质量评价方法通过对图像进行预处理,然后提取图像的形状和方差等统计特征来评价图像质量。SAAD 等^[3]基于无参考评价方法,提出了一种基于 DCT 的无参考图像完整性统计质量评价算法 (Blind Image Integrity Notator Using DCT Statistics, BLINDS) 及其改进算法 (BLINDS-II),并通过实验研究及分析建立了支持向量回归模型;MOORTHY 等^[4]研究发现图像的失真会有规律地影响图像小波域的自然统计特性,并以此为理论基础提出了无参考图像质量评价算法 (Blind Image Quality Assessment, BIQA)。这些传统图像特征统计的质量评价方法实现的瓶颈在于特征参数的选择以及特征参数之间如何结合,所以近些年来研究重点渐渐转移到了用特征学习方法来进行无参考图像质量评价上。由于深度学习是基于特征学习的典型方法,结合深度学习技术对图像质量进行评价的方法也越来越得到研究人员的重视。基于深度学习的质量评价最大的优点在于它不需要丰富的先验知识和手动提取的图像特征,而是学习图像自身的信息。KANG 等^[5]提出了基于卷积神经网络 (CNN) 的无参考图像质量评价模型 (Image quality assessment using Convolutional Neural Network, IQA_CNN),将特征提取和回归集成到 CNN 框架加深网络深度,但是并未考虑对比度引起的质量差和亮度引起的质量差异。文献[6]中提到的无参考图像质量评价算法 (A Deep Neural Network for Image Quality Assessment, BosICIP) 通过对大小为 32×32 的 RGB 局部图像子块进行 CNN 训练,对图像子块进行 MOS 值平均,得到整个图像的 MOS 值,但是局部图像的 MOS 值并不能反映出整个图像的 MOS 值,从而影响了质量评价的精度。

以上基于人工合成的失真图像的质量评价研究方法,其混合失真图像的失真类型都是理想化的几种失真叠加而成,然而,真实世界中的图像的失真往往

是由多种未知类型的失真相互叠加作用的结果。基于此,近几年学者们研究了一些基于传统特征组合方法来解决真实失真图像的质量评价问题。杨蓝平^[7]结合图像的自然场景统计特征和人类视觉审美感知,提出了一种无参考图像质量评价模型,该模型主要分为图像特征提取、特征数据处理和支持向量机学习回归三部分。GHADIYARAM 等^[8]提出了一种基于特征图的无参考图像质量评估 (Feature maps based Referenceless Image Quality Evaluation Engine, FRIQUEE) 算法,该方法通过综合利用广义高斯分布 (generalized Gaussian distribution, GGD) 数学模型提取在多种色彩空间下的图像自然场景统计特征,并且利用支持向量机回归 (Support Vector Regression, SVR) 得到图像的质量预测得分。随着学者们对机器学习技术的不断深入,越来越多的基于深度学习的真实失真图像质量评价算法受到研究者的关注。唐敏^[9]提出了一种 DCNN 来捕获图像的局部和全局信息,并在 LIVEWild^[10]图像库和 CID2013^[11]模糊失真图像库上进行了有效验证。LI 等^[12]提出了一种基于高级语义特征的无参考图像质量评价方法 (Semantic Feature Aggregation using PLSR, SFA),提取图像子块的高层语义信息并将其输入到线性回归模型,实现图像质量的预测,该方法只针对真实模糊图像而未充分考虑到现实失真图像的其他失真特性。DOMONKOS VARGA 等^[13]利用迁移学习的概念,提出了一种无参考真实失真图像质量评价方法 (A content preserving deep architecture for blind image quality assessment, DeepRN),该算法运用迁移学习的方法,并采用大规模真实失真图像库 KonIQ-10k^[14]对算法性能进行测试,该方法在 MOS 值的预测精度上有了显著提高,但是该模型计算成本和时间成本过高。

基于上述真实失真图像质量评价算法研究现状以及目前存在的问题,文中提出了一种基于图像底层特征和高级语义提取的真实失真图像质量评价方法。由于图像的底层特征能够更充分地表达图像的内容,所以首先采用 KonIQ-10k 图像库提供的图像底层特征指标为基础研究数据,建立了真实失真图像的预分类机制;然后利用 DCNN 模型能够提供丰富的图像表示的特性,提取出图像的一级语义特征;接着考虑到有效的特征统计函数能够更全面地表示图像失真信息,对一级语义特征进行特征聚合,形成真实失真图像的二级语义特征;最后,为了解决图像高级语义特征的维度灾难问题,提出高容量回归器将二级高级语义特征回归到图像的 MOS 值,并在 KonIQ-10k 这个真实失真图像库上对算法进行验证。

1 相关工作

由于图像的信息十分丰富,人眼视觉系统通常运用多种过滤和分解的方式对图像进行处理,以提取出视觉信息中的不同视觉感知特征。客观的质量评价方法主要目的在于寻找一个理想的数学模型对图像进行质量评价,模拟人类视觉系统感知图像质量的过程^[15-16]。

1.1 图像的底层特征

虽然主观评价是最准确可靠的表征图像直觉感受的方法,但是常用的 MOS 值只是针对图像质量整体的数值概述,它不能够全面地描述图像所表达的意义;图像的底层特征一定程度上能够更全面反映图像的特性,综合描述人眼对图像的感知信息^[17-18]。对于图像质量评价中的底层特征,不仅要满足区分能力强、抗干扰能力强、鲁棒性好等特点,最重要的就是要做到无线接近于人的视觉感知。图像可以通过自身特征多样性来表达不同的内容,这些特征的多样性表现在多个方面,并且构成了人类视觉认知的基础。研究者通过大量主观测试实验统计得出了与人眼视觉感知密切相关的 5 项图像底层特征指标:亮度、颜色、对比度、噪声和清晰度。

节选 KonIQ-10k 图像库中 MOS 值相同的 5 张图,见图 1,底层指标数据见表 1,从表 1 可以看出 MOS 值相同的图像对应着不同的图像底层特征分布,说明图像的底层特征能够更针对性地反映图像的信息变化。

1.2 高级语义特征和特征聚合

由于图像的底层特征是基于人眼视觉感知的,人们难以选取特定的底层特征或者对其进行特征组合

来实现对图像质量进行评价,所以需要高级语义特征来量化地表示图像信息。图像质量评价领域所涉及的关于高级语义特征的描述,指的是通过深度学习框架进行训练学习得到的能够表示图像失真信息的高维特征。研究表明,从深度学习训练中提取的用于图像质量评价的高级语义特征,能够提供丰富的图像失真信息表示^[19-21]。

高级语义特征一般都是维数较高的向量,但是这些高维的特征数据不足以准确表达图像内在的复杂的信息,需要对这些高维的特征数据进行进一步的统计处理。特征聚合指的是对高级语义特征数据进行特征函数处理,实现特征重构的过程。利用高级语义特征对图像进行描述过程时,需要运用不同的数据归纳分析方法,这样才能从不同的角度与侧面反映表达图像的失真特征。在图像质量评价问题中,高级语义特征数据统计函数的准确性,直接决定着图像特征数据能否代表图像的失真信息,所以高级语义特征的重构会对失真图像的准确描述形成重要的影响,使得重构后的特征更能准确可靠地表示图像失真信息。

1.3 高容量回归

为了解决图像高级语义特征的维度灾难问题,需要高容量的回归器将图像高级语义特征回归到图像的主观质量分数。高容量回归分析是指在大量样本数据的基础上,研究高维自变量和低维因变量之间关系的一种预测模型技术。传统的单计算层感知器主要用于解决线性可分问题,而大数据背景下的回归预测问题是线性不可分的。为了克服单计算层感知器这一局限性,人们在输入层与输出层之间引入隐藏层作为感知过程的内部表示,即形成了多层前馈神经网络。结合深度学习的相关知识,全连接层的实质就是个矩阵乘法,相当于一个特征空间变换,可以把前面所有有



图 1 KonIQ-10k 图像库中的图像示例 (MOS 值相同的 5 张图像)

Fig.1 Examples in the KonIQ-10k image library

表 1 图像底层特征与 MOS 值的关系

Tab.1 Relationship between image underlying features and MOS values

图像	MOS	亮度	颜色	对比度	噪声	清晰度
图 1a	3.036	0	11	85	15	0
图 1b	3.036	1	28	47	34	0
图 1c	3.036	5	18	57	28	2
图 1d	3.036	1	14	100	21	0
图 1e	3.036	1	15	74	21	0

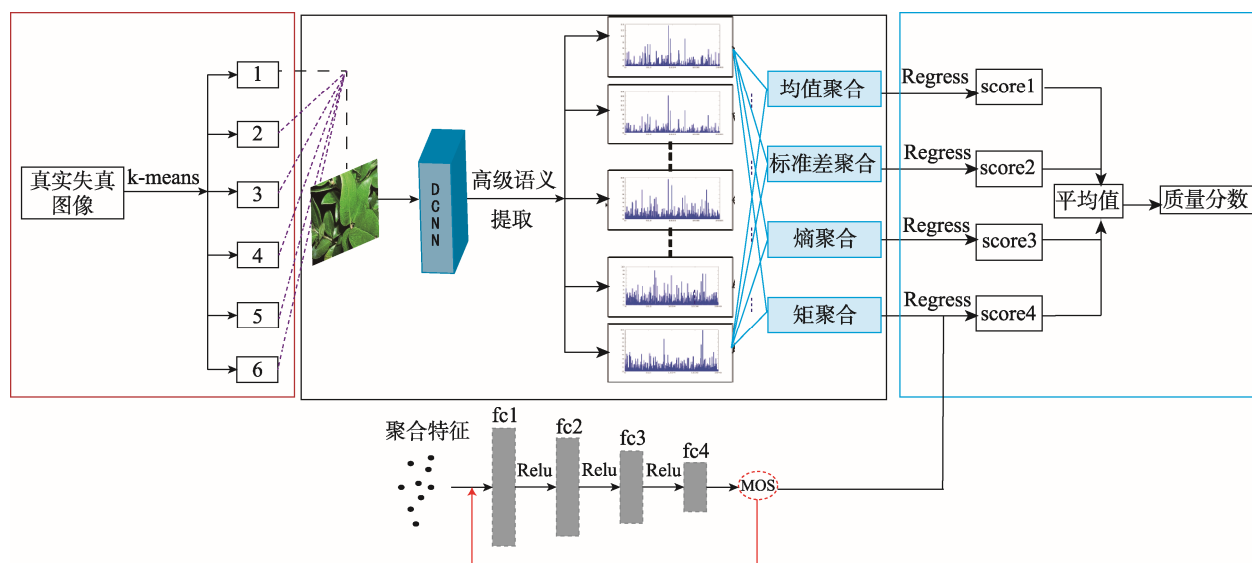


图 2 文中算法的框架流程

Fig.2 Framework of the proposed algorithm

用的信息提取整合。再加上激活函数的非线性映射，由输入层、多层全连接层、输出层组成的多层前馈神经网络（Multilayer Feed-forward Neural Network, MFNN）理论上可以模拟任何非线性变换，且相关研究证明该模型能够有效解决高维数据背景下的回归预测问题^[22]。

2 真实失真图像质量评价模型

提出了基于图像底层特征和高级语义的真实失真图像质量评价研究方法，根据图像底层特征通过对 KonIQ-10k 图像库进行 k-means 聚类实现预分类，结合 DCNN 模型的高效图像描述性能以及特征统计函数对高级语义分析能力的提升，利用深度卷积神经网络提取出图像的一级语义特征，然后通过均值聚合、标准差聚合、熵聚合和矩聚合 4 种聚合方式，将一级语义特征聚合成二级语义特征，将聚合后的二级语义特征输入到由多层前馈神经网络构成的高容量回归器中，以解决真实失真图像的质量分数回归问题，提出的真实失真图像质量评价的具体算法流程见图 2。

2.1 基于底层特征的图像分类

由于真实失真图像之间内容千差万别，并且存在多种多样的失真，为了避免由于多种失真叠加而导致的无法准确预测出图像质量分数的问题，提出了基于 KonIQ-10k 真实失真图像库进行 k-means 聚类，实现图像库的预分类，基于 KonIQ-10k 真实失真图像库提供的 5 个底层指标（亮度、颜色、对比度、噪声和清晰度）的绝对类别评级指标（Absolute Category Rating, ACR）值为聚类样本，将图像库划分为不同子类别，便于后续对真实失真图像的处理。

2.2 基于深度学习的高级语义特征提取器

在利用深度学习研究图像质量评价算法模型时，为了有效解决图像数据量不足的问题，一般研究者都会采用 2 种方法来实现数据扩增：将图像划分为许多图像子块，用来表示一幅图像，再通过图像子块进行特征的训练学习，最后平均子块的质量分数得到一幅图像的质量得分；原始图像做各种旋转、镜像和尺寸变换，然后将扩增后的图像输入到质量评价算法模型进行训练学习。在图像质量评价领域，图像局部的 MOS 值并不能与图像整体的 MOS 值保持一致，而图像质量评价的目的是模拟代替人类视觉系统感知图像的质量，所以为了解决分块的数据扩增方法无法准确描述图像质量的问题，对图像进行 90°, 180°, 270° 翻转来实现数据扩增，然后考虑到图像的全局信息，直接将大小为 224 × 224 的原图作为 DCNN 的输入，提取出真实失真图像的一级高级语义特征，再对一级高级语义特征进行聚合得到二级高级语义特征，使得聚合后的特征更能表达图像失真的信息。

2.3 一级高级语义特征提取

从 DCNN 模型中提取的图像高维特征可以用来表示图像的高级语义信息。文献[23]中提到残差图像能够提供重要的图像失真信息，文中使用预训练的 ResNet-50 模型作为图像的特征提取器，ResNet-50 网络的相关参数见表 2，将批处理参数 Batch_size 设置为 128，学习率设置为 0.001，通过将整幅图像作为 ResNet-50 的输入来计算图像的高级语义特征，提取中间层的输出特征 $d_1 \sim d_l$ ，见图 3。

每一幅图像提取出的特征可以表示为：

$$d_l = DCNN(l; \theta, \omega), \quad l = 1, 2, \dots, n \quad (1)$$

式中： l 为提取特征的所在卷积层； θ, ω 表示预训练 DCNN 的权重参数。

表2 ResNet-50 网络的参数
Tab.2 Parameters of ResNet-50 network

层	输出	ResNet-50
Conv1	112×112	7×7, 64, stride 2 3×3 max pool, stride 2
Conv2	56×56	1×1, 64 3×3, 64×3 1×1, 256 1×1, 128
Conv3	28×28	3×3, 128×4 1×1, 512 1×1, 256
Conv4	14×14	3×3, 256×6 1×1, 1024 1×1, 512
Conv5	7×7	3×3, 512×3 1×1, 2048
Pool	1×1	global average pool 1000-d fc, softmax

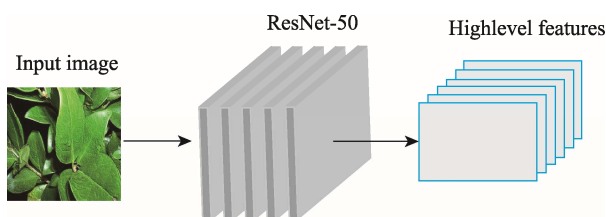


图3 基于 DCNN 的高级语义特征提取示意
Fig.3 Schematic diagram of high-level semantic feature extraction based on DCNN

2.4 二级高级语义特征聚合

提出 4 种聚合方式对提取的一级高级语义进行特征重构，见图 4。

2.4.1 均值聚合

考虑到每一维度高级语义特征的均值，计算剔除 (percent/2) % 的最大值和最小值后的样本均值，则第 1 种聚合特征 f_1 可以表示为：

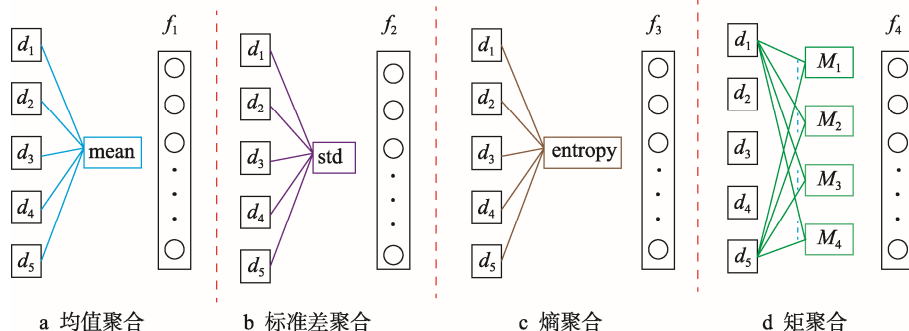


图4 高级语义特征聚合示意

Fig.4 Schematic diagram of high-level semantic feature aggregation

$$f_1 = f_{\text{mod_mean}} \{d_l\} \quad (2)$$

2.4.2 标准差聚合

标准差能反映一组数据的离散程度，是概率统计中最常用的用来统计数据分布程度的测量方法，则第 2 种聚合特征 f_2 可以表示为：

$$f_2 = f_{\text{std}} \{d_l\} \quad (3)$$

2.4.3 熵聚合

由信息论可知，如果信息的随机性越大，那么把它确定下来的价值也越大，信息论通过定义一种定量的学习方法“熵”来比较随机变量的信息价值，用熵值作为第 3 种聚合特征：

$$f_3 = E \{d_l\} \quad (4)$$

$$E = -\sum_{l=1}^n p(d_l) \log p(d_l) \quad (5)$$

2.4.4 矩聚合

矩能够描述数据分布的统计信息，为了平衡对更多信息的需求和特征空间的降维，用中心阶矩作为第 4 种聚合特征：

$$f_4 = M_1 \oplus M_2 \oplus M_3 \oplus M_4 \quad (6)$$

$$M_k = \left(\sqrt[k]{\frac{\sum_{i=1}^n (d_{i1} - m_1)^k}{n}}, \dots, \sqrt[k]{\frac{\sum_{i=1}^n (d_{in} - m_n)^k}{n}} \right)^T \quad (7)$$

式中： $k=1, 2, 3, 4$ 。

2.4 高容量特征回归器

由于图像的高级语义特征的维数较高，使用传统的回归算法容易引起因数据维度过高造成的运行成本过大的问题。全连接层的一个主要作用是能够实现维度的变换，尤其是可以把高维变到低维，同时能够保留下来有用的信息。全连接层的另一个作用是可以表达隐含的语义，把原始特征映射到各个隐层语义节点。对于图像质量评价模型的最后一层全连接而言，就是质量回归的显示表达。文中通过统计实验确定了由输入层、4 个完全连接层和输出层组成的一个 MFNN 作为高容量回归框架，每一层全连接层后面紧连着一个非线性激活函数，见图 5。

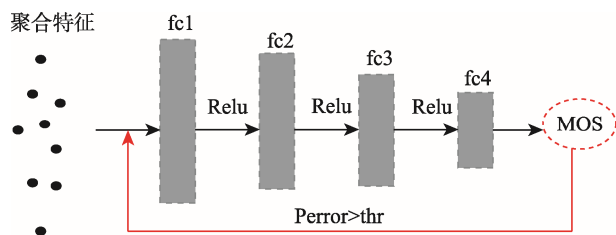


图 5 高容量回归器模型

Fig.5 High-capacity regression model

假设训练样本总数为 N ，训练周期 epochs 和迭代次数 iteration 存在着以下关系：

$$1\text{epoch} = \frac{N}{\text{Batch_size}} \times \text{iteration} \quad (8)$$

$$P_{\text{error}} = S_{\text{test}} - S_{\text{predict}} \quad (9)$$

$$\text{Accuracy} = \frac{\sum(|P_{\text{error}}| \leq \text{thr})}{N} \quad (10)$$

式中： P_{error} 表示预测出来的质量分数与实际 MOS 值的差值；Accuracy 表示测试集中误差绝对值在 thr 内的样本量占测试总样本量的比例；thr 表示回归预测的质量分数精度。

2.5 算法评价指标

采用以下 2 种性能指标对质量评价算法进行客观评价：Spearman 秩序相关系数 SROCC 和 Pearson 线性相关系数 PLCC，两者的值越接近于 1，说明主观一致性越强。

SROCC 定义如下：

$$\text{SROCC} = 1 - \frac{6 \sum_{i=1}^n e_i^2}{n(n^2 - 1)} \quad (11)$$

式中： e_i 为第 i 个图像的主观质量分数与客观质量分数之差。

PLCC 定义如下：

$$\text{PLCC} = \frac{\sum_{i=1}^n (s_i - \bar{s})(q_i - \bar{q})}{\sqrt{\sum_{i=1}^n (s_i - \bar{s})^2} \sqrt{\sum_{i=1}^n (q_i - \bar{q})^2}} \quad (12)$$

式中： s_i 和 q_i 分别为第 i 个图像主观和客观质量分数； $\bar{s}$ 和 $\bar{q}$ 分别为主观和客观质量分数的平均值。

3 实验结果与分析

所有对比实验都基于 Windows10 CPU Intel Core i7，内存 8 G，使用了深度学习框架 caffe 的 Matlab 接口。

3.1 测试数据库

采用 Hanhe Lin 等提出的 KonIQ-10k 真实失真图

像库，该图像库从大约 480 万幅图像中采样了 10 073 幅图像，采样算法使用了清晰度、颜色、亮度、对比度、噪声 5 项质量特征指标。对于每幅图像，均由 1467 名群众工作者通过主观统计实验获得 120 个可靠的质量评级。与现有的 IQA 数据库相比，KonIQ-10k 图像库包含大量的图像，具有更广泛的实际应用价值。

3.2 图像预分类类别数的确定

为确定最优的聚类中心的个数，即图像预分类的类别数，统计分析了类别总数 k 与测试集内质量分数的预测准确率的关系曲线，见图 6。

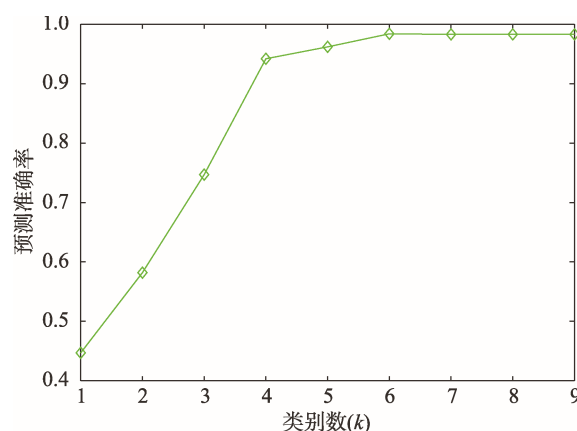


图 6 预测准确率与类别数的关系曲线

Fig.6 Relation curve between forecast accuracy and categories

从图 6 可以看出，KonIQ-10k 图像库未分类时，测试集内图像质量分数的预测准确率有 0.447，随着类别数 k 的不断上升，图像的测试准确率不断上升，当 k 取 6 时，对训练集内图像进行测试的准确率较高且趋于平缓，所以将 KonIQ-10k 图像库分为 6 类。

3.3 高容量回归器训练参数的确定

针对 KonIQ-10k 图像库的训练集图像和测试集图像，与预训练模型提取特征的学习参数保持一致，将批处理训练参数 Batch_size 设置为 128，学习率设置为 1×10^{-3} ，回归预测的质量分数精度 thr 设置为 0.25，然后设置了具体的训练时间参数，分别做了仿真实验，记录了均方根误差 (Root mean squared error, Rmse) 和损失函数值 (Loss function value, Loss) 的变化趋势，实验结果见图 7—8。

图 7a—b 分别为高容量回归器中，训练集的迭代次数与 Rmse 和 Loss 的关系曲线，由图 7 可知提出的高容量回归器能够快速而又准确地实现数据的稳定收敛。

训练周期与测试集的 Rmse 和 Accuracy 的关系曲线分别见图 8a—b，纵坐标分别为训练集的预测质量分数与真实质量分数的 Rmse 值，以及在预定误差

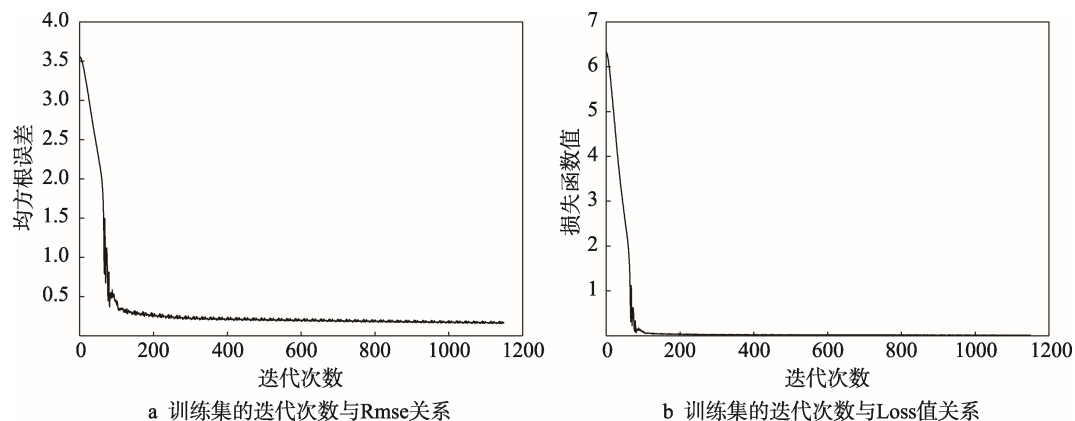


图7 训练集的迭代次数与Rmse和Loss的关系曲线
Fig.7 Iterations and Rmse & Loss relation curve in training set

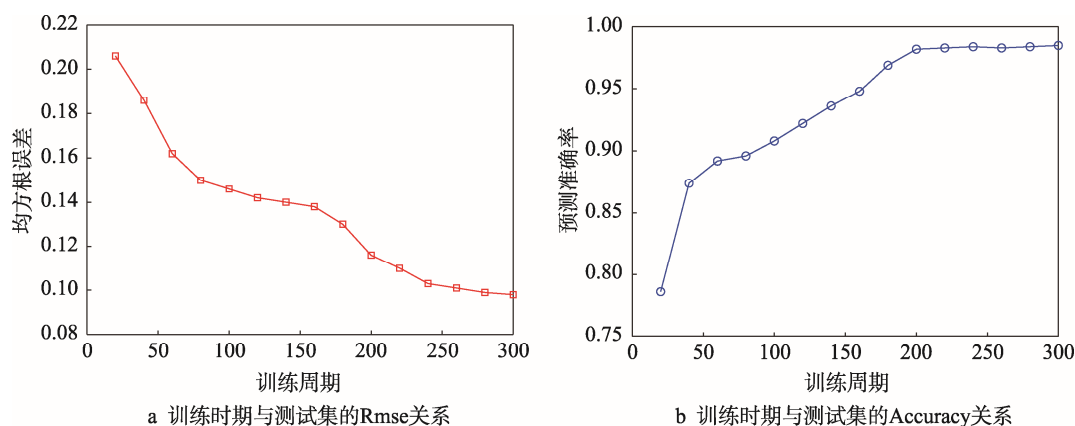


图8 训练时期与测试集的Rmse和Accuracy关系曲线
Fig.8 Rmse and Accuracy performance with different epochs

内质量分数预测正确的样本占总样本数的比例,即训练集的预测准确率。由图8可以看出,当训练 epochs 选择为 240 时,测试集的 Rmse 和 Accuracy 均能达到较理想的数值。

3.4 质量评价性能测试与分析

统计分析了不同质量评价算法在 KonIQ-10k 图像库上的性能,见表3。

表3 不同质量评价算法在 KonIQ-10k 上的性能对比
Tab.3 Performance comparison of different IQA algorithms on KonIQ-10k

Method	SROCC	PLCC
BRISQUE ^[3]	0.700	0.704
BLINDS-II ^[4]	0.575	0.584
SSEQ ^[24]	0.596	0.615
IQA_CNN ^[5]	0.634	0.671
BosICIP ^[6]	0.654	0.673
SFA ^[10]	0.685	0.764
DeepBIQ ^[25]	0.900	0.920
DeepRN ^[11]	0.920	0.950
Proposed	0.956	0.974

从表3可以看出,提出的质量评价模型在 KonIQ-10k 这个庞大真实失真图像库上的实验结果,明显优于传统手动特征提取和质量回归算法 BRISQUE, BLINDS-II, SSEQ, 同时与 BosICIP, SFA, DeepBIQ, DeepRN 相比,进一步反映了该模型对真实失真图像客观评分的有效性,即文中模型对图像质量的客观评分与人眼主观质量得分具有较好的一致性。

同时对利用深度学习技术进行图像质量评价的算法进行对比分析,以复现相关算法所用时间为统计指标,见表4。相比于同样针对 KonIQ-10k 真实失真图像库提出的质量评价算法 DeepRN 而言,提出的算法模型不仅在训练样本量上有了较明显的降低,而且降低了时间成本。与其他利用深度学习技术对图像质量作出评价的算法而言,在样本量和时间成本上得到了很好的折中效果,充分证明了文中评价模型具有较高的运行效率和预测精度。

3.5 消融研究

为验证提出的质量评价模型的有效性,在 KonIQ-10k 图像库上进行了消融验证实验,分析研究

了模型中的各个组成部分对 SROCC 和 PLCC 的影响,实验结果见图 9。

表 4 不同深度学习算法的性能对比
Tab.4 Performance comparison between different deep learning algorithms

算法	样本量	时间/h
IQA_CNN ^[5]	493577	96
BosICIP ^[6]	493577	96
SFA ^[10]	90657	84
DeepRN ^[11]	104949	336
DeepBIQ ^[25]	10073	6
Proposed	40292	8

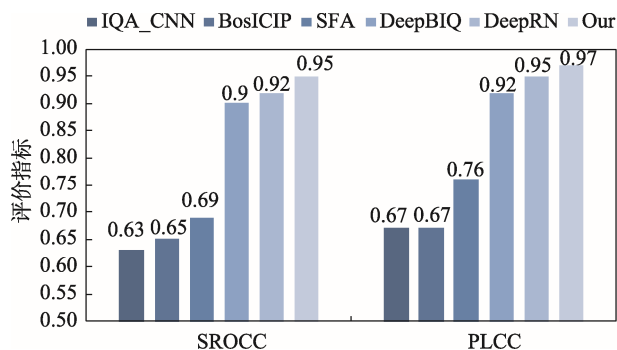


图 9 KonIQ-10k 图像库中不同算法的消融研究结果
Fig.9 Ablation results of different algorithms in the KonIQ-10k image library

1) 预分类机制。基于图像底层特征的预分类机制,相比于 DeepRN 中将其按照 MOS 值 1~5 进行预分类训练的模型而言,其 SROCC 和 PLCC 分别提高了 3.3%和 5.4%。

2) 预训练模型。相比于 KANG^[5]中 IQA_CNN 算法使用端对端的训练学习方法对图像质量作出评价,本算法模型运用深度卷积预训练模型提取图像特征然后进行质量回归的质量评价模型,评价指标 SROCC 和 PLCC 达到了 95%和 97%,比使用 IQA_CNN 算法分别提高了 50.8%和 44.8%。

3) 高级语义聚合。为了验证高级语义特征聚合对图像质量评价模型效率的提升,将本模型与 DeepBIQ 模型进行对比。DeepBIQ 运用了简单的特征池化和连接的方法对提取出来的高级语义特征进行整合,而本算法利用 4 种统计函数对高级语义聚合,使得 SROCC 和 PLCC 有了 3.3%和 4.5%的提升。

4) 高容量回归。提出由多层前馈神经网络构成的高容量回归模型,相对于 SFA 中运用最小二乘回归法将聚合后的高级语义特征回归到图像的质量分数上,本算法在真实失真图像库上的质量评价结果的 SROCC 和 PLCC 都有了明显的提高,分别能达到 0.95 和 0.97。

4 结语

以 KonIQ-10k 这个真实失真图像库为基础分析数据,针对人眼对真实失真图像底层特征的感知对图像进行预分类,然后结合深度学习预训练模型提取图像的高级语义特征,提出由全连接层和激活层组成的高容量回归器,在图像子类中展开对失真图像质量分数的回归预测。通过实验证明,与传统自然场景统计特性的质量评价算法和当前的一些基于机器学习的质量评价方法相比,文中的质量评价模型得出的图像质量分数与人类主观质量评价结果具有较好的一致性,而且运用预训练模型进行特征提取的方式大大降低了训练时间成本,运行效率得到了有效提升,为快速判断失真图像质量和后续图像增强奠定了基础。除此之外,如何针对 KonIQ-10k 图像库外的真实失真图像进行底层指标的预测,并根据其指标分布有针对性地进行图像增强,将是今后研究工作的重点。

参考文献:

- [1] 卢鹏,林根巧,邹国良.基于信息熵和深度学习的无参考图像质量评价方法研究[J].计算机应用研究,2018,35(11):3508—3512.
LU Peng, LIN Gen-qiao, ZOU Guo-liang. Study on Non-reference Image Quality Evaluation Method Based on Information Entropy and Deep Learning[J]. Application Research of Computer, 2018, 35(11): 3508—3512.
- [2] MITTAL A, MOORTHY A K, BOVIK A C. No-reference Image Quality Assessment in the Spatial Domain[J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2012, 21(12): 4695.
- [3] SAAD M A, BOVIK A C, CHARRIER C. Blind Image Quality Assessment: A Natural Scene Statistics Approach in the DCT Domain[J]. IEEE Transactions on Image Processing, 2012, 21(8): 3339—3340.
- [4] MOORTHY K, BOVIK, CONRAD A. Blind Image Quality Assessment: From Natural Scene Statistics to Perceptual Quality[J]. IEEE Transactions on Image Processing, 2011, 20(12): 3350—3364.
- [5] LE Kang, YE Peng, YI Li, et al. Convolutional Neural Networks for No-Reference Image Quality Assessment[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus: IEEE Computer Society, 2014: 1733—1740.
- [6] BOSSE S, MANIRY D, WIEGAND T, et al. A Deep Neural Network for Image Quality Assessment[C]// 23th IEEE International Conference on Image Processing (ICIP). Phoenix: IEEE, 2016: 3773—3775.
- [7] 杨蓝平.基于真实失真图像无参考质量评价的研究[D].北京:北京邮电大学,2017:9—20.

- YANG Lan-ping. Research on Non-reference Quality Evaluation Based on Real Distortion Image[D]. Beijing: Beijing University of Posts and Telecommunications, 2017: 9—20.
- [8] GHADIYARAM D, BOVIK A C. Feature Maps- driven No-reference Image Quality Prediction of Authentically Distorted Images[J]. Proceedings of SPIE-The International Society for Optical Engineering, 2015, 9394: 93940J-93940J-14.
- [9] 唐敏. 基于主观感知的真实图像质量评价方法的研究[D]. 北京: 北京邮电大学, 2018: 21—30.
TANG Min. Research on Real Image Quality Evaluation Method Based on Subjective Perception[D]. Beijing: Beijing University of Posts and Telecommunications, 2018: 21—30.
- [10] JAYARAMAN D, MITTAL A, MOORTHY A K et al. Objective Quality Assessment of Multiply Distorted Images[C]// ASILOMAR: 20132012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, IEEE Computer Society, 2013: 1693—1697.
- [11] VIRTANEN T, NUUTINEN M, VAAHTERANOKSA M, et al. CID2013: A Database for Evaluating No-Reference Image Quality Assessment Algorithms[J]. IEEE Trans Image Process, 2015, 24(1): 390—402.
- [12] LI Ding-quan, JIANG Ting-ting, JIANG Ming. Exploiting High-Level Semantics for No-Reference Image Quality Assessment of Realistic Blur Images[C]// ACM2017: Proceedings of the 2017 ACM on Multimedia Conference. New York: ACM, 2017: 378—385.
- [13] VARGA D, SAUPE D, SZIRANYI T. DeepPrn: A Content Preserving Deep Architecture for Blind Image Quality Assessment[C]// ICME 2018: 2018 IEEE International Conference on Multimedia and Expo (ICME), San Diego, CA: IEEE, 2018: 1—6.
- [14] LIN Han-he, HOSU V, SAUPE D. KonIQ-10k: Towards an Ecologically Valid and Large-scale IQA Database[C]// CVPR 2018: Proceeding of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2018: 1—5.
- [15] 何南南, 解凯, 李桐, 等. 图像质量评价综述[J]. 北京印刷学院学报, 2017, 25(2): 47—50.
- HE Nan-nan, XIE Kai, LI Tong, et al. A Review of Image Quality Evaluation[J]. Journal of Beijing Institute of Graphic Communication, 2017, 25(2): 47—50.
- [16] MANSOURIN A, MAHMOUDIA A. SSVD: Structural SVD-based Image Quality Assessment[J]. Signal Processing, 2019(74): 54—63.
- [17] LIU Xuan-xuan, WANG Xiao-hong, ZHANG Ting. An Image Quality Evaluation Method Based on Visual Underlying Features[J]. Optical, 2015, 41(3): 280—284.
- [18] SUN Jun-ding, LI Hai-hua, JIN Jiao-lin, et al. Automatic Image Annotation Based on Multi-feature Fusion and PLSA-GMM[J]. Measurement & Control Technology, 2017, 36(4): 31—35.
- [19] SIAHAAN E, HANJALIC A A, REDI J. Augmenting Blind Image Quality Assessment Using Image Semantics[C]// ISM 2016: 2016 IEEE International Symposium on Multimedia, San Jose, ISM, 2016: 307—312.
- [20] GAO Fei, WANG Yu, LI Pan-peng, et al. DeepSim: Deep Similarity for Image Quality Assessment[J]. Neuro-computing, 2017: S092523121—7301480.
- [21] YU Shao-de, WU Shi-bin, JIANG Fan, et al. A Shallow Convolutional Neural Network for Blind Image Sharpness Assessment[J]. PLOS ONE, 2017, 12(5): 2—5.
- [22] WANG Yu-xiong, HEBERT M. Learning to Learn: Model Regression Networks for Easy Small Sample Learning[C]// ECCV 2016: European Conference on Computer Vision. Amsterdam: Springer, 2016: 616—620.
- [23] JIA Yan, ZHANG Wei-xia, FENG Tian-peng. Blind Image Quality Assessment Based on Natural Redundancy Statistics[C]// ACCV 2016: Asian Conference on Computer Vision. Taiwan: Springer, 2016: 4—11.
- [24] LIU Li-xiong, LIU Bao, HUANG Hua, et al. No-reference Image Quality Assessment Based on Spatial and Spectral Entropies[J]. Signal Processing: Image Communication, 2014, 29(8): 856—863.
- [25] BIANCO S, CELONA L, NAPOLETANO P, et al. On the Use of Deep Learning for Blind Image Quality Assessment[J]. Signal, Image and Video Processing, 2018(12): 355—362.