

# 基于人眼视觉特性的深度学习全参考图像质量评价方法

姚 旺<sup>1,2,3</sup>, 刘云鹏<sup>1,3</sup>, 朱昌波<sup>1,2,4</sup>

(1. 中国科学院沈阳自动化研究所, 辽宁 沈阳 110016; 2. 中国科学院大学, 北京 100049;  
3. 中国科学院光电信息处理重点实验室, 辽宁 沈阳 110016;  
4. 中国科学院沈阳自动化研究所 机器人学国家重点实验室, 辽宁 沈阳 110016)

**摘 要:** 针对现有的图像质量评价方法普遍为人工设计特征, 难以自动且有效提取到符合人类视觉系统的图像特征, 受人眼视觉特性的启发, 提出一种新的基于卷积神经网络的全参考图像质量评价方法(DeepFR)。该方法基于对数据集本身的学习设计了卷积神经网络 DeepFR 模型, 利用人眼视觉系统对梯度的敏感性进行加权优化, 提取了符合人眼视觉特性的视觉感知图。实验表明: 设计的 DeepFR 模型优于已有的全参考图像质量评价方法, 其预测结果与主观质量评价有很好的精确性与一致性。

**关键词:** 图像质量评价; 全参考; 深度学习; 卷积神经网络; 人眼视觉特性

**中图分类号:** TP391.4 **文献标志码:** A **DOI:** 10.3788/IRLA201847.0703004

## Deep learning of full-reference image quality assessment based on human visual properties

Yao Wang<sup>1,2,3</sup>, Liu Yunpeng<sup>1,3</sup>, Zhu Changbo<sup>1,2,4</sup>

(1. Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China;  
2. University of Chinese Academy of Sciences, Beijing 100049, China;  
3. Key Laboratory of Opto-Electronic Information Processing, Chinese Academy of Sciences, Shenyang 110016, China;  
4. State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China)

**Abstract:** Since the current image quality assessment methods are generally based on hand-crafted features, it is difficult to automatically and effectively extract image features that conform to the human visual system. Inspired by human visual characteristics, a new method of full-reference image quality assessment was proposed by this paper which was based on convolutional neural network (DeepFR). According to this method, the DeepFR model of convolutional neural network was designed which was based on the understanding of the dataset by itself using the human visual system to weight the sensitivity of the gradient, and the visual gradient perception map was extracted that was consistent with human visual characteristics. The experimental results show that the DeepFR model is superior to the current full-reference image quality assessment methods, its prediction score and subjective quality evaluation have good accuracy and consistency.

**Key words:** image quality assessment; full-reference; deep learning; convolutional neural network; human visual properties

收稿日期: 2018-04-10; 修订日期: 2018-05-20

基金项目: 装发部共用技术项目(Y6K4250401)

作者简介: 姚旺(1994-), 女, 硕士生, 主要从事图像质量评估方面的研究。Email: yaowang@sia.cn

导师简介: 刘云鹏(1980-), 男, 研究员, 博士, 主要图像处理、模式识别方面的研究。Email: ypliu@sia.cn

## 0 引言

图像在获取、传输、压缩等过程中易受到不同程度的失真,导致图像质量的下降,而图像质量的优劣将直接影响人与计算机获取信息的能力,因此开发一种对图像质量进行测量和量化的系统有重要应用价值。客观图像质量评价是通过模拟人眼视觉系统来衡量图像质量优劣的方法,一般根据在评价过程中对参考图像的依赖程度分为全参考、部分参考与无参考。全参考图像质量评价可以根据参考图像与失真图像之间的差异来量化视觉敏感度,从而使其能够与人类视觉系统有很好一致性,因此,文中采用全参考图像质量评价方法。

全参考图像质量评价一般采用两步进行设计:首先,通过利用数学模型对参考图和失真图的差异比较,如早期的均方误差(MSE)和峰值信噪比(PSNR)直接利用原始图像与扭曲图像之间的差,结构相似性指标(SSIM)<sup>[1]</sup>及其变体IW-SSIM<sup>[2]</sup>,主要利用了亮度、对比度与结构化信息3个因素的组合,视觉信息保真度(VIF)<sup>[3]</sup>通过计算感知的参考图像被保留在感知中的失真信息来表示差异。然后,通过汇聚池策略得到最终的质量分数。广泛使用的汇聚方式主要有平均池和加权池,例如MGSD<sup>[4]</sup>采用了平均池对GSD预测图像质量,FSIM<sup>[5]</sup>采用了基于相位一致性的加权池产生最终质量得分。GMSD<sup>[6]</sup>采用了标准偏差池的方法,提高了预测的精确性,降低了计算复杂度,但其适用的范围有限,可能使图像局部质量估计变得复杂。

近年来,深度学习已被广泛应用于计算机视觉领域,尤其是对目标识别和图像分类等困难问题产生了变革性的影响<sup>[7-8]</sup>。深度学习模型的主要优点之一是在对大规模有标签数据集进行训练时具有显著泛化能力,然而由于缺乏足够的被人类标定的图片数据库,将深度网络与图像质量评价结合起来的研究依然很少。参考文献[9]首次在图像质量评价领域中使用卷积神经网络,为了获取足够的数

据,将最后一层改为回归层,使学习到的特征映射到主观分数上。针对在训练时利用全局主观质量分数来标记局部图像会有一定误差,参考文献[11]提出了一个基于CNN的两阶段的无参考图像质量评价模型,该方法利用全参考图像质量评价产生的局部质量得分作为图像块得分标签。参考文献[12]中使用预先训练的AlexNet模型,提取特征图并比较特征相似性,将这些相似性分数汇集获得总体质量分数。参考文献[13]中作者将深度学习模型应用于全参考图像质量评价中,将参考图像、失真图像输入在ImageNet预训练的模型,通过计算输出特征图的局部相似性来获得全局图像质量分数,与传统的全参考质量模型相比,达到了较好的预测精度。

在真实情况下,人眼在观察一副失真图像时,会对边缘区域的失真感到明显,而上述基于深度学习的方法只是利用已有的数据集自动提取图像相关特征。针对这一问题,文中采用了卷积神经网络生成了视觉差异图映射,根据失真图像的局部空间特性对视觉差异图进行加权来进行图像质量评价。因此,文中的贡献如下:提出了一种新的基于深度学习的全参考图像质量评价模型,该模型考虑了人眼视觉系统对图像像素空间特性的感知,通过端到端的训练优化,自动学习到入眼视觉的敏感特征,实现了与人类主观评价的一致性。通过在不同数据集上对DeepFR模型进行实验验证,结果表明该模型性能优于已有的图像质量评价方法。

## 1 深度图像质量评价模型

### 1.1 网络结构

文中提出的DeepFR网络构架如图1所示,包含8个卷积层来提取特征图,3个最大池来缩减特征图的大小同时提取局部较强特征,以及两个全连接层组成的非线性回归函数。由参考文献[4]~[6]可知,失真图像不同局部区域遭遇不同程度退化,图像梯度对图像失真敏感,人眼对图像梯度具有敏感性。为了符合人眼的视觉特性,笔者采用梯度特征来定义梯度差异图来增强人眼视觉的敏感性。输入的原始图像要进行归一化的预处理,有助于网络学习到自然统计特性,具体预处理方法见下文。归一化后的梯度差异图与失真图像分别输入到卷积层

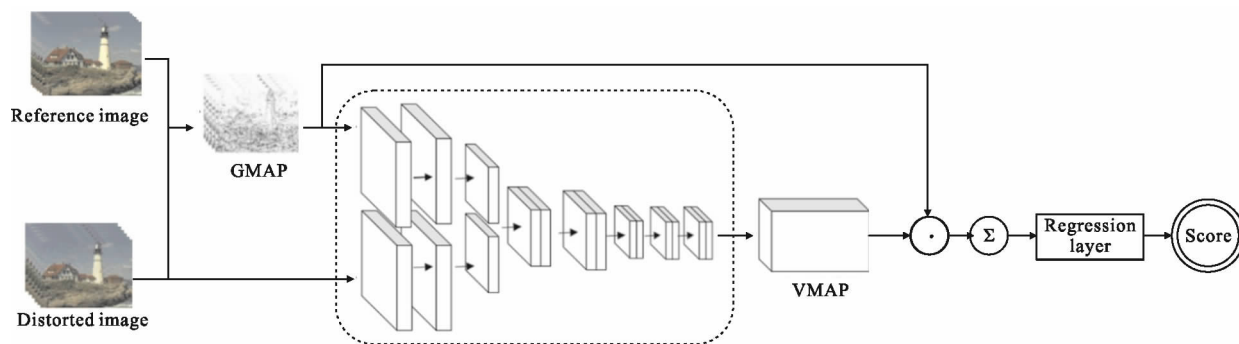


图 1 DeepFR 网络构架

Fig.1 Network architecture of DeepFR

CONV1-1 与 CONV1-2 中并在第 1 个 MaxPooling 后连接到一起。为了保持卷积过程中特征图的尺寸,卷积层采用零填充。同时,对前 7 个卷积层采用 LReLU 作为激活函数,最后 1 个采用 ReLU 作为激活函数。由于每张图片经过了两次池化操作,输出的图片尺寸为输入的 25%。模型的最后,使用 2 个全连接层进行回归,得到最终的主观得分。

## 1.2 图像预处理

在进行图像输入之前,对图像进行归一化预处理,有助于更快地达到最优值。参考文献[9]中提到了采用图像归一化进行预处理可以一定程度提高算法的性能。受参考文献[9]的启发,采用局部平均相减和除法归一化来计算局部归一化的亮度。首先,将参考图像  $I_r$  与失真图像  $I_d$  转化为灰度图。然后,分别对参考图像与失真图像进行归一化处理得到  $\hat{I}_r$  与  $\hat{I}_d$ 。通过对原始的输入图像做局部归一化,可以有效的去相关,不需要计算复杂的变换,有利于卷积神经网络学习到图像质量有关的特征。

## 1.3 基于人类视觉特性的梯度差异图定义

图像梯度由于能够有效捕获图像的局部结构,对人类视觉系统(HVS)高度敏感,而经常被应用于图像质量评估算法中。采用强化图像中的梯度差异来定义梯度差异图,使得 DeepFR 网络更好地学习到图像中的边缘信息。梯度特征利用线性滤波器对图像进行卷积来提取。为了简化计算,同时降低对噪声的敏感性,文中采取 Prewitt 算子来提取图像中的梯度特征。3×3 的 Prewitt 算子模板滤波器沿水平与垂直方向的定义为:

$$h_x = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \times \frac{1}{3}, h_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \times \frac{1}{3} \quad (1)$$

分别将参考图像  $\hat{I}_r$  和失真图像  $\hat{I}_d$  与 Prewitt 算子滤波器进行卷积,以便产生水平与垂直方向上的梯度图像,计算方式如下:

$$G_r^2 = (\hat{I}_r \otimes h_x)^2 + (\hat{I}_r \otimes h_y)^2 \quad (2)$$

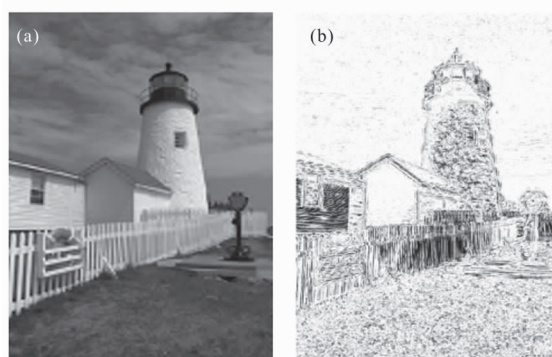
$$G_d^2 = (\hat{I}_d \otimes h_x)^2 + (\hat{I}_d \otimes h_y)^2 \quad (3)$$

定义的梯度差异图 (Gradient Map, GMAP) 方法如下:

$$\text{GMAP} = \frac{2G_r G_d + \varepsilon}{G_r^2 + G_d^2 + \varepsilon} \quad (4)$$

式中:  $\varepsilon$  为避免分母为零,保持系统稳定的正常数。由公式(4)可知,当输入的参考图像与失真图像相同时,即图像无失真,梯度差异图上每一点都达到最大值 1。

如图 2 所示,图 2(a)、(c)为 LIVE 数据集中 FF 类型产生失真的灰度图,其 DMOS 值分别为 16.50 与 17.90,图 2(b)与(d)分别相应的梯度差异图,可以看出对于失真明显的区域如图 2(d)中的边缘,梯度差异图对其做了加强。





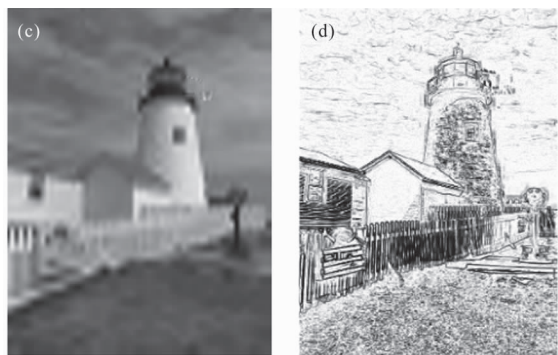


图 2 不同程度失真 GMAP 图

Fig.2 Examples of GMAP maps with various degrees of distortion

## 1.4 目标函数定义

将梯度差异图与失真图像一同输入到 DeepFR 的卷积神经网络模型中进行训练,得到视觉差异图 (Vision Map, VMAP):

$$\text{VMAP}=f(\hat{I}_d, \text{GMAP}; \theta) \quad (5)$$

式中:  $f(\cdot)$  表示参数为  $\theta$  的卷积神经网络模型。

为了模拟人类视觉系统,加强梯度在图像评价过程中的比重,利用梯度差异图对视觉差异图中每个像素点做权重得到视觉感知图 (Vision Gradient Map, VGMAP), 计算方式如下:

$$\text{VGMAP}=\text{VMAP} \odot \text{GMAP} \quad (6)$$

式中:  $\odot$  为哈达玛积运算。

为了减轻卷积过程中零充填造成的误差,对实验过程中的图像边界进行剪裁, 汇总得分为 VGMAP 剪裁后的图像区域 VGMAP' 的加权平均:

$$M=\frac{1}{h' \cdot w'} \sum_{(i,j) \in \text{VGMAP}'} \text{VGMAP}(i,j) \quad (7)$$

式中:  $h'$ 、 $w'$  为视觉感知图 (VGMAP) 剪裁后图像区域 VGMAP' 的长与宽。最后通过非线性回归得到最终得分, 目标函数定义为:

$$L=\|g(M)-S\|_F^2 \quad (8)$$

式中:  $g(\cdot)$  表示非线性回归函数;  $S$  为输入失真图像的真实主观得分。

## 1.5 训练方法

由于 LIVE 数据库图像尺寸不同, 要把它分成固定大小的块进行输入, 并且保证同一个失真图片的所有图像块应在同一个 batch 中, 注意输入失真图像的图像块与梯度差异图的图像块一一对应, 深

度卷积网络输出的同一幅图的所有特征图像块再按相对位置融合成视觉差异图。在此过程中, 需要注意避免特征图像块出现交叉重叠。因此, 在从失真图片获取失真图像块时, 需要设置适当的滑动步长, 文中设为 80。

DeepFR 模型网络架构具体参数如表 1 所示, 在初始输入, 分别将梯度差异图与失真图像输入到 Conv1-1 与 Conv2-1, 在各自进行 MaxPool 后, 使用 Fusion 直接将图像黏合到一起, 然后输入到 Conv3 中, 网络的具体变换过程见表 1。网络由端对端进行训练, 特征由卷积层提取, 通过两个完全连接层执行非线性回归, 得到最后的质量得分。

表 1 DeepFR 具体参数

Tab.1 DeepFR parameters

Fusion	Activation	Input channel	Output channel	Filter sizes	Stride
Conv1-1	LReLU	1	32	(3,3)	(1,1)
Conv2-1	LReLU	32	32	(3,3)	(1,1)
Maxpool1	—	32	32	(2,2)	(2,2)
Conv1-2	LReLU	1	32	(3,3)	(1,1)
Conv2-2	LReLU	32	32	(3,3)	(1,1)
Maxpool2	—	32	32	(2,2)	(2,2)

续表 1

Continued Tab.1

	Activation	Input channel	Output channel	Filter sizes	Stride
Conv3	LReLU	64	64	(3,3)	(1,1)
Conv4	LReLU	64	64	(3,3)	(1,1)
Maxpool	—	64	64	(2,2)	(2,2)
Conv5	LReLU	64	64	(3,3)	(1,1)
Conv6	ReLU	64	1	(3,3)	(1,1)
FC1	LReLU			(1,4)	
FC2	ReLU			(4,1)	

## 2 实验验证与性能分析

### 2.1 数据集及评价指标

实验过程中采用了 4 种标准的图像数据集 LIVE, CSIQ, TID2008<sup>[15]</sup>, TID2013 的全部图像, 每个数据库都包含了多种失真类型和相应 MOS 分数或

DMOS 分数。LIVE 数据集包含 29 幅参考图像及相应 982 幅失真图像,包含 JPEG 压缩, JPEG2000 压缩, 加性高斯白噪声 (Additive Gaussian White Noise, AWN), 高斯模糊(GB)和快速衰落(FF)5 种失真类型,为每幅图像提供了 0~100 的 DMOS 值。CSIQ 数据集包含 30 幅参考图像及相应 866 幅失真图像,包含了 JPEG 压缩, JP2K 压缩, 加性高斯白噪声(AWN),高斯模糊(GB),全局对比度(GCD)和加性高斯粉红噪声(PN)6 种失真类型,为每幅图像提供了 0~1 的 DMOS 值。TID2008 数据集有 17 种失真类型的 1 700 幅失真图像来源于 25 幅参考图像,提供了 0~9 的 MOS 值。TID2013 也是由 25 幅参考图像及 24 种失真类型所对应的 3 000 幅失真图像组成,同样为每幅图像提供了 MOS 值。

文中分别利用皮尔森线性相关系数(PCC)与斯皮尔曼秩相关系数(SRC)来评价模型的预测精确性与预测单调性,对图像质量评估算法进行性能比较。皮尔森线性相关系数的值越接近于 1 表示预测的精确性越高。斯皮尔曼秩相关系数计算后的值越接近于 1 证明预测的单调性越好。

## 2.2 参数设置

根据参考图像按照 4:1 的比例随机分为训练数据集与测试数据集,同时按照参考图像把相对应的失真图像也分为训练数据集与测试数据集。同时采用了水平翻转的方式增加训练样本的数量。由于不同数据集的 MOS 与 DMOS 值不同,需要对每个数据集的得分进行按比例进行缩放,调整至范围在 [0,1] 的 MOS 值。

训练过程中,采用了带有 Nesterov 动量项的 Adam 方法用于随机梯度下降,设置权重为  $10^3$ ,学习率初始化为  $5 \times 10^{-4}$ 。同时,利用 TV 正则化<sup>[14]</sup>作用于视觉差异图,得到总方差损失  $tv$ ,设置的权重为  $10^{-2}$ 。同时,使用了 L2 正则化的方式来防止过度拟合。为了实验结果具有说服力,在每个数据集随机进行了 10 次训练测试,最终取了结果的平均值。

## 2.3 人眼视觉一致性验证

为了验证 DeepFR 模型与人类视觉系统的一致性,图 3(a)~(d)依次是经过预处理的失真图像,梯度差异图,预测到的视觉差异图以及视觉感知图。图 3(a)为高斯模糊产生的失真图像的灰度图,它的边缘不清晰,在预测的图 3(c)、(d)边缘区域被明显加强。可以发现失

真明显区域,图像变暗被增强。可见模型可以根据失真程度进行合理的质量预测。

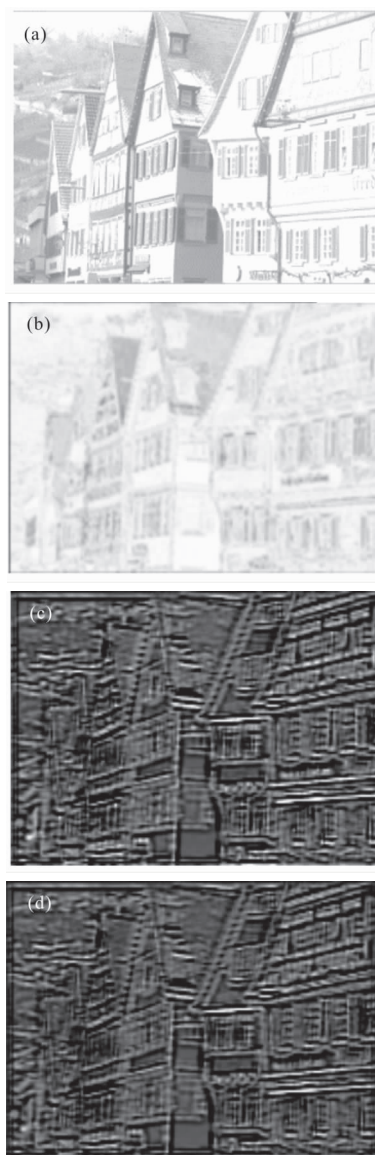


图 3 图像结果

Fig.3 Image results

为了进一步说明 DeepFR 模型与人类感知的一致性,文中分别绘出了在 LIVE, CSIQ, TID2013 数据集中 DeepFR 模型预测的图像质量分数与 MOS 值的散点图,如图 4 所示。图中每个点代表 1 幅图像,  $x$  轴标记了 MOS 值,  $y$  轴为预测分数。图 4(a)~(c) 分别为在 LIVE, CSIQ 与 TID2013 数据集中测试集上测试结果。图 4(d)~(f)分别为由 LIVE, CSIQ 与 TID2013 数据集上已训练好的模型在 TID2008 上的测试结果,由于 LIVE 与 CSIQ 包含的失真类

型小于 TID2008 数据集中的失真类型,因此,只选取其中与 TID2008 失真类型相同的图像测试。由图 4 中散点图可知,在所有图像数据库中,DeepFR 方法对图像质量的预测值与主观评价得分有很好的-致性,不仅说明了 DeepFR 模型效果不依赖于单一数据集,而且也说明了模型对于不同数据集进行质量预测的结果与人眼视觉感知相一致。

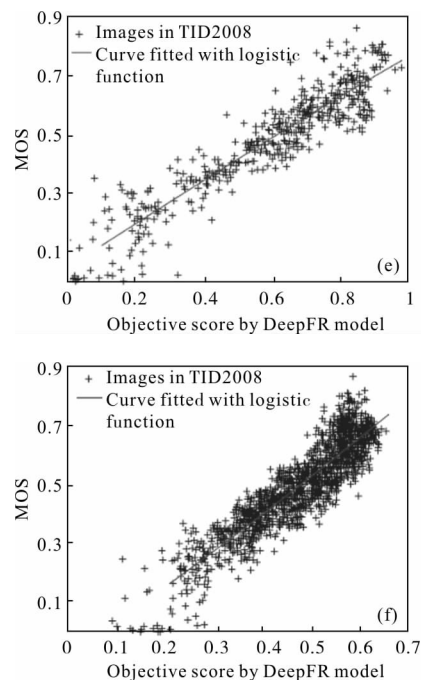
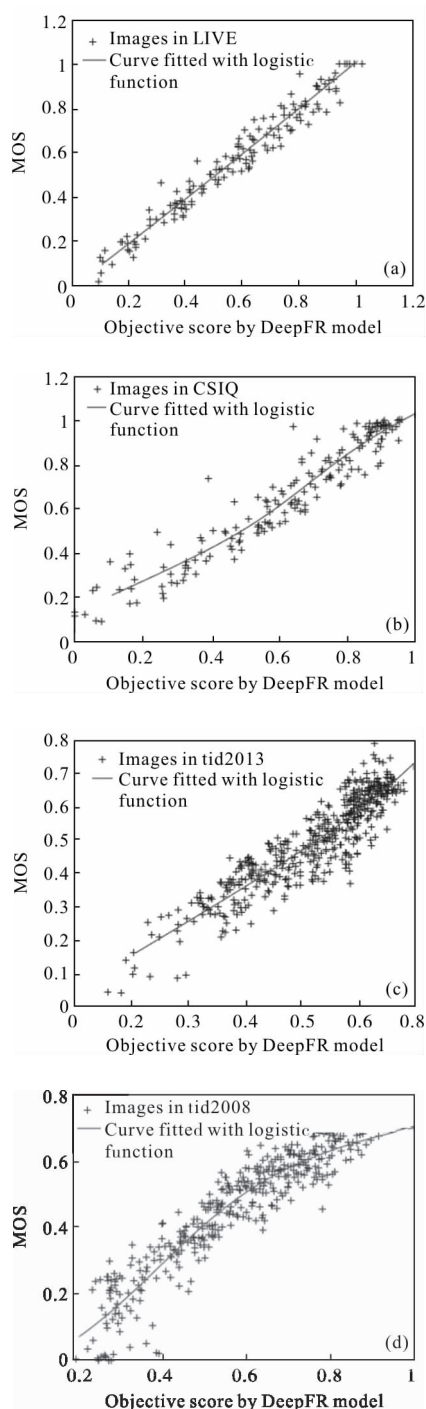


图 4 预测得分散点图

Fig.4 Scatter plots of the quality score

## 2.4 性能比较

为了验证 DeepFR 模型的有效性,文中对 LIVE、CSIQ、TID2013,3 个图像质量数据集进行评价,利用斯皮尔曼秩相关系数,皮尔森相关系数作为评价标准,对 DeepFR 模型方法进行评价后,与 PSNR, SSIM<sup>[1]</sup>及其变体 IW-SSIM<sup>[2]</sup>,VIF<sup>[3]</sup>,GMSD<sup>[6]</sup>,FSIM<sup>[5]</sup>,IQ(I<sub>T</sub>)<sup>[12]</sup>,DeepSim<sup>[13]</sup>等已有的最具代表性的全参考图像质量评价方法作比较,结果如表 2 中 3 个数据集上整体性能

表 2 3 个 IQA 数据集上整体性能比较

Tab.2 Performance comparison of three IQA databases

IQA method	LIVE(779)		CSIQ(866)		TID2013(3 000)		Weighted average	
	SRC	PCC	SRC	PCC	SRC	PCC	SRC	PCC
PSNR	0.876	0.872	0.806	0.800	0.636	0.706	0.708	0.751
SSIM	0.948	0.945	0.876	0.861	0.637	0.691	0.733	0.765
IW-SSIM	0.957	0.952	0.921	0.914	0.778	0.832	0.835	0.867
VIF	0.963	0.960	0.920	0.928	0.677	0.772	0.770	0.833
GMSD	0.960	0.960	0.957	0.954	0.804	0.859	0.859	0.894
FSIM	0.963	0.960	0.924	0.912	0.802	0.859	0.852	0.886
IQ(I <sub>T</sub> )	-	0.910	-	0.920	-	0.840	-	0.867
DeepSim	0.974	0.968	0.919	0.919	0.846	0.872	0.881	0.897
DeepFR	0.981	0.984	0.960	0.966	0.876	0.894	0.909	0.923



比较所示。其中除文中算法以外其他结果数据都是由相关论文提供。表2中结果可以看出 DeepFR 模型的评估方法在整个 LIVE,CSIQ 与 TID2013 数据集上的运行结果都明显优于其他的评价方法,而且其加权平均效果也有较高的预测准确性与一致性,说明了 DeepFR 模型对不同数据集都有较好的质量预测效果。

为了测试文中算法在不同失真类型的图像失真情况评价效果,对在 LIVE 数据集上已训练好的模型,采用了对 LIVE 数据集的每种失真类型进行测试,结果如表 3 所示。表 3 中结果显示了 DeepFR 模型不仅在整个 LIVE 数据集上显示了很好的一致性 with 精确性,而且在 JP2K,JPEG,WN,FF 失真类型上也比其他算法预测精确度高。

表 3 单一失真类型在 LIVE 上的比较

Tab.3 Performance comparison of individual distortion types on the LIVE

IQA method	JP2K		JPEG		WN		BLUR		FF	
	SRC	PCC	SRC	PCC	SRC	PCC	SRC	PCC	SRC	PCC
PSNR	0.895	0.876	0.881	0.903	0.985	0.917	0.782	0.780	0.891	0.880
SSIM	0.961	0.941	0.972	0.946	0.969	0.982	0.952	0.900	0.956	0.951
GMSD	0.968	0.963	0.973	0.976	0.974	0.977	0.957	0.954	0.942	0.939
FSIM	0.970	0.910	0.981	0.985	0.967	0.976	0.972	0.978	0.949	0.912
IQ(I <sub>r</sub> )	—	0.960	—	0.940	—	0.990	—	0.980	—	0.910
DeepFR	0.975	0.984	0.973	0.990	0.988	0.992	0.967	0.965	0.977	0.983

### 3 结 论

文中依据人眼视觉对梯度的敏感性,利用梯度差异图,设计了一种基于卷积神经网络模型的全参考图像质量评价方法。与传统的手工提取特征的全参考图像质量评估方法相比,文中采用了卷积神经网络方法模拟人类视觉系统,自动提取特征,利用梯度差异图增强梯度信息在评价过程的敏感性,得到的评价结果符合人眼视觉系统感知图像质量的过程。实验结果表明文中利用深度学习自动提取特征的方法能较准确的评价图像质量,整体性能优于已有的经典评价方法。进一步改善性能可以考虑采用常规方法进行预处理,再进行深度学习,以便减少训练参数,提高算法的运算速度。

#### 参考文献:

[1] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity [J]. *IEEE Trans Image Process*, 2004, 13(4): 600–612.  
[2] Wang Z, Li Q. Information content weighting for perceptual

image quality assessment [J]. *IEEE Transactions on Image Processing*, 2011, 20(5): 1185–1198.

[3] Sheikh H R, Bovik A C. Image information and visual quality [J]. *IEEE Transactions on Image Processing*, 2006, 15(2): 430–444.  
[4] Cheng G, Huang J C, Zhu C, et al. Perceptual image quality assessment using a geometric structural distortion model[C]// *IEEE International Conference on Image Processing*, 2010: 325–328.  
[5] Zhang D. FSIM: A feature similarity index for image quality assessment [J]. *IEEE Transactions on Image Processing*, 2011, 20(8): 2378–2386.  
[6] Xue W, Zhang L, Mou X, et al. Gradient magnitude similarity deviation: A highly efficient perceptual image quality index [J]. *IEEE Transactions on Image Processing*, 2014, 23(2): 684–695.  
[7] Luo Haibo, He Miao, Hui Bin, et al. Pedestrian detection algorithm based on dual-model fused fully convolutional networks [J]. *Infrared and Laser Engineering*, 2018, 47(2): 0203001. (in Chinese)  
罗海波, 何淼, 惠斌, 等. 基于双模全卷积网络的行人检测算法(特邀)[J]. *红外与激光工程*, 2018, 47(2): 0203001.  
[8] Luo Haibo, Xu Lingyun, Hui Bin, et al. Status and prospect

- of target tracking based on deep learning [J]. *Infrared and Laser Engineering*, 2017, 46(5): 0502002. (in Chinese)
- 罗海波, 许凌云, 惠斌, 等. 基于深度学习的目标跟踪方法研究现状与展望 [J]. 红外与激光工程, 2017, 46(5): 0502002.
- [9] Kang L, Ye P, Li Y, et al. Convolutional neural networks for No-reference image quality assessment [C]//Computer Vision and Pattern Recognition, IEEE, 2014: 1733–1740.
- [10] Li Y, Po L M, Feng L, et al. No-reference image quality assessment with deep convolutional neural networks [C]//IEEE International Conference on Digital Signal Processing, 2017: 685–689.
- [11] Kim J, Lee S. Fully deep blind image quality predictor[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2017, 11(1): 206–220.
- [12] Ali Amirshahi S, Pedersen M, Yu S X. Image quality assessment by comparing CNN features between images [J]. *Electronic Imaging*, 2016, 60(6): 6041010.
- [13] Gao F, Wang Y, Li P, et al. Deep Sim: Deep similarity for image quality assessment [J]. *Neurocomputing*, 2017 (1): 104–114.
- [14] Mahendran A, Vedaldi A. Visualizing deep convolutional neural Networks using natural pre-images[J]. *International Journal of Computer Vision*, 2016, 120(4): 1–23.
- [15] Ponomarenko N, Lukin V, Zelensky A, et al. TID2008 –a database for evaluation of full-reference visual quality assessment metrics [J]. *Adv Modern Radioelectron*, 2009, 10 (1): 30–45.