

与图像内容无关的聚焦程度评价方法

夏晓华, 柴玉琳, 岳鹏举, 杨治, 秦绪芳

(长安大学道路施工技术与装备教育部重点实验室, 710000, 西安)

摘要: 针对现有的聚焦程度评价方法在图像场景内容改变时不能正确评价图像聚焦程度的问题, 基于“分类+拟合”的思想, 提出了一种与高斯模糊标准差完全等价的图像聚焦程度评价方法。首先, 建立了以有限高斯模糊标准差为标记的图像聚焦程度分类数据集; 然后, 构建了用于提取图像高斯模糊标准差分类分数的非对称核卷积神经网络(AKC-net); 最后, 采用三次样条插值函数拟合 AKC-net 全连接层输出的分类分数以及对应的高斯模糊标准差, 以最大分数对应的标准差作为图像的聚焦程度评价结果, 并在 Waterloo 数据集和实际拍摄图像上分别进行仿真实验和实拍实验。结果表明: 所提方法在不同聚焦图像上分类的平均准确率可达到 97.7%, 得到的评价结果与高斯模糊标准差真值的均方根误差和平均绝对误差均小于 0.07, 且实际拍摄图像的聚焦测度值与图像内容无关, 实现了图像聚焦程度的绝对评价。

关键词: 聚焦程度评价; 高斯模糊标准差; 非对称核卷积神经网络; 三次样条插值

中图分类号: TP399 **文献标志码:** A

DOI: 10.7652/xjtuxb202501011 **文章编号:** 0253-987X(2025)01-0116-09

Focus Evaluation Method with Independent of Image Content

XIA Xiaohua, CHAI Yulin, YUE Pengju, YANG Zhi, QIN Xufang

(Key Laboratory of Road Construction Technology and Equipment of MOE, Chang'an University, Xi'an 710000, China)

Abstract: In response to the challenge where current focus evaluation methods struggle to accurately assess image focus amidst fluctuations in image scene content, a focus evaluation method, fully equivalent to Gaussian blur standard deviation, is introduced based on the “classification + fitting” concept. Firstly, a dataset for image focus classification, annotated with finite Gaussian blur standard deviations, is established. Subsequently, an asymmetric kernel convolution neural network (AKC-net) is developed to derive image Gaussian blur standard deviation classification scores. Finally, a cubic spline interpolation function is applied to fit the classification score of the AKC-net fully connected layer output and the corresponding Gaussian blur standard deviation, using the standard deviation linked to the maximum score as the focus evaluation for the images. Simulation experiments and real shooting experiment are carried out on the Waterloo dataset and real shooting images respectively. The experimental findings indicate that the proposed method achieves an average classification accuracy of up to 97.7% across various focus images. The root mean square error and mean absolute error between the obtained evaluation results and the true Gaussian blur standard deviation values are less than 0.07.

收稿日期: 2024-05-12。 作者简介: 夏晓华(1987—), 男, 副教授, 博士生导师。 基金项目: 国家自然科学基金资助项目(61901056); 陕西省重点研发计划资助项目(2024GX-YBXM-197)。

网络出版时间: 2024-09-10

网络出版地址: <https://link.cnki.net/urlid/61.1069.t.20240910.1410.004>

Furthermore, the focus measurement values derived from real shooting images are independent of the image content, enabling the absolute evaluation of the image focus.

Keywords: degree of focus evaluation; Gaussian blur standard deviation; asymmetric kernel convolution neural network; cubic spline interpolation

聚焦程度评价方法在自动对焦^[1-2]、多聚焦图像融合^[3-4]、聚焦形貌恢复^[5]等领域中应用广泛,其性能决定了应用效果。一个理想的聚焦程度评价方法应满足独立于图像内容、具有良好的噪声鲁棒性、计算效率高、对离焦区域敏感等要求^[6]。近些年,国内外大量文献报道了提高抗噪性^[7]、计算效率^[4]和对离焦区域敏感^[8]的聚焦程度评价方法。但是,上述方法对图像内容有很强的依赖性,当图像内容发生改变时,不能正确评价其聚焦程度。

目前,尽管已有一些针对图像内容无关的聚焦程度评价方法研究,但相对较少。Bahrami 等^[9]利用广义高斯分布对图像内容感知总变化进行参数化,提取标准差和内容参数两个特征,对这两个特征进行权重分配得到与图像内容无关的聚焦程度评价方法。Zhan 等^[10]通过对最大梯度和梯度的可变性这两个梯度特征参数进行加权几何平均来评价不同图像内容的聚焦程度,其中最大梯度表示图像的聚焦位置,梯度的可变性表示图像内容的变化,然而该方法只能计算图像的聚焦位置,且需要对图像进行分块处理。Liu 等^[11]提出利用边缘方向检测算法得到图像的边缘宽度,建立宽度直方图,通过与距离因子加权求和得到不同图像内容的聚焦程度。上述方法在评价图像聚焦程度上受人为提取特征的影响^[12],评价准确性不高。Li 等^[13]提出了一种基于语义特征聚合的聚焦程度评价方法,利用预训练的分类卷积神经网络获取图像固有的内容感知特性,提取深度语义特征,通过线性回归模型进行拟合评价图像的聚焦程度,对不同内容的图像有较好的评价效果。然而,该方法是基于主观评价价值来判断图像的聚焦程度,对于无法实时获得主观评价值的实际拍摄图像,只通过聚焦测度值并不能判断图像的聚焦程度,无法实现对不同图像内容聚焦程度的绝对评价,且该方法使用的卷积神经网络中卷积结构为普通的方形卷积核,提取图像特征的能力较弱,线性回归模型对于离散型变量如分类网络的类别拟合精度低。高斯模糊虽然能够很好地模拟图像离焦效果,但是对于聚焦程度未知的图像,现有的聚焦程度评价方法无法直接测量其高斯模糊值。

针对现有聚焦程度评价方法存在的上述问题,

考虑到卷积神经网络可识别图像的聚焦区域和离焦区域^[4],本文提出了一种非对称核卷积神经网络(AKC-net)与数据拟合相结合的聚焦程度评价方法。该方法以高斯模糊标准差作为聚焦程度的评价标准,可以实现不同场景内容图像聚焦程度的绝对评价,且 AKC-net 中的非对称卷积核^[14]增强了模型对图像的特征提取能力,在不增加计算量的前提下提高网络的分类准确率,之后利用提取的网络中相关参数,通过数据拟合计算极值,评价图像的聚焦程度,实验验证了本文方法的有效性和优越性。

1 提出方法

考虑到高斯模糊能模拟图像离焦^[6],本研究以高斯模糊标准差作为图像聚焦程度的评价标准,首先建立以有限高斯模糊标准差为标记的图像聚焦程度分类数据集,然后构建用于提取图像高斯模糊标准差分类分数的 AKC-net,最后采用三次样条插值函数拟合 AKC-net 全连接层输出的分类分数以及对应的高斯模糊标准差,以最大分数对应的标准差作为图像的聚焦程度评价结果。聚焦程度评价方法流程如图 1 所示。

1.1 图像聚焦程度分类数据集

假定待检测图像离焦程度所对应的高斯模糊标准差范围为 $[0, \sigma_{\max}]$, 图像聚焦程度分类数据集将其划分为 n 个类别, 则每个类别对应的高斯模糊标准差依次为 $0, \frac{\sigma_{\max}}{n-1}, \frac{2\sigma_{\max}}{n-1}, \dots, \frac{(n-2)\sigma_{\max}}{n-1}, \sigma_{\max}$ 。将拍摄的聚焦图像或公共数据集中的聚焦图像依次进行高斯模糊标准差为 $\frac{\sigma_{\max}}{n-1}, \frac{2\sigma_{\max}}{n-1}, \dots, \frac{(n-2)\sigma_{\max}}{n-1}, \sigma_{\max}$ 的离焦模糊, 以高斯模糊标准差作为分类标签对图像数据进行标记, 聚焦图像的高斯模糊标准差记为 0, 建立图像聚焦程度分类数据集。

数据集的高斯模糊标准差范围 $[0, \sigma_{\max}]$ 和分类数量 n 关系到最终图像聚焦程度的检测精度, 分类高斯模糊标准差的间隔和平均检测误差 $\bar{E}$ 应满足以下经验公式

$$\frac{\sigma_{\max}}{n-1} = \mu \bar{E} \quad (1)$$

式中: μ 取经验值 5~10。

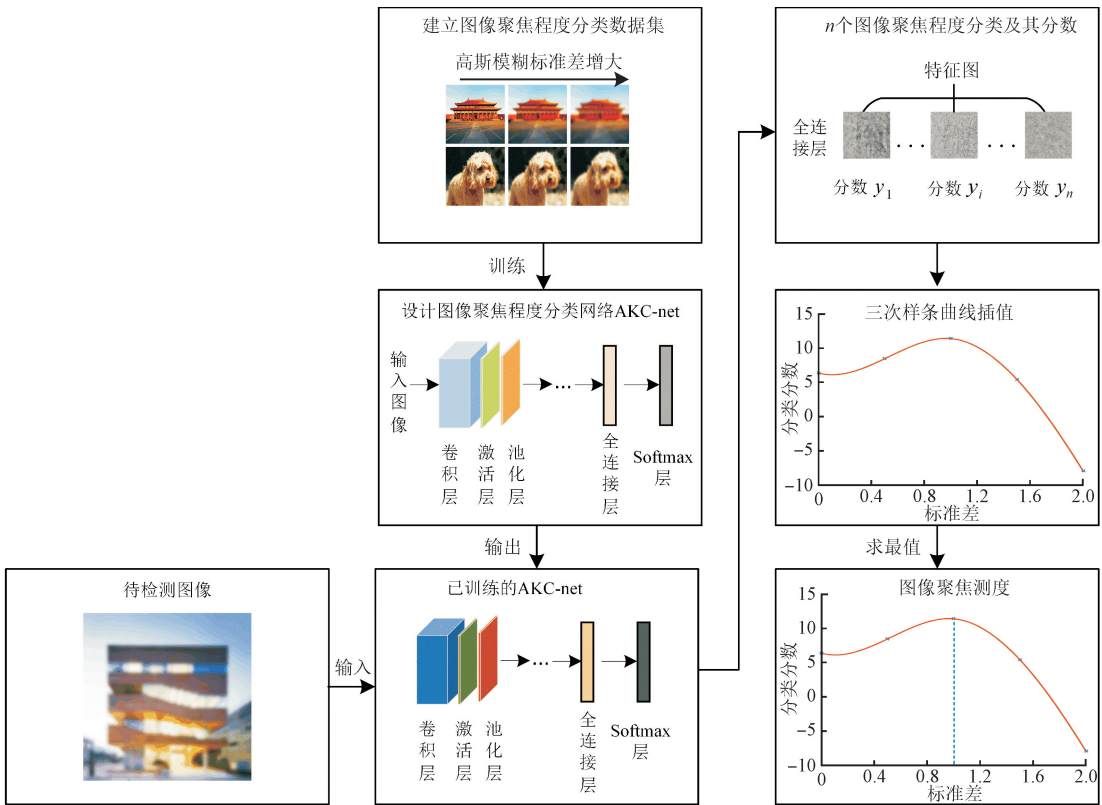


图 1 聚焦程度评价方法流程图

Fig. 1 Flow chart of focus degree evaluation method

1.2 聚焦程度分类网络

不同的卷积神经网络结构对图像分类影响很大。为了提高分类准确率,本文设计了 AKC-net 模型,其结构如图 2 所示。该模型结构包含 6 个卷积单元,每一个卷积单元包括非对称卷积块、Rule 激活层和最大池化层 3 部分。非对称卷积块是用 3 个并行的尺寸 3×3 、 1×3 和 3×1 卷积核代替 3×3 方

形卷积核,其示意图如图 3 所示。新网络结构卷积核数是旧网络的 3 倍,极大地提高了网络对图像特征的提取能力,新网络通过对并行的 3 个卷积核结果进行融合,获得非对称卷积块的最终输出。卷积核过滤器尺寸为 3×3 ,第 1 个卷积单元中卷积核数为 32,卷积后获得 32 个尺寸为 224×224 的特征图。

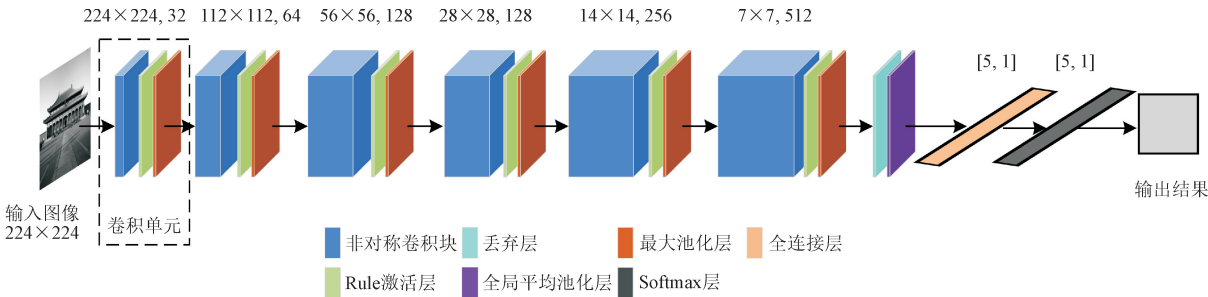


图 2 AKC-net 模型结构图

Fig. 2 Structure of AKC-net model

在非对称卷积块之后连接 Rule 激活层,卷积单元中最后一层为最大池化层,对非对称卷积块输出的特征图计算最大池化,设置滤波尺寸为 2×2 ,步长为 2。最大池化层采用下采样的方式,减小特征

图的大小并删除冗余的空间信息,池化后得到 32 个尺寸 112×112 的特征图。输入图像经过最后一个卷积单元后得到 512 个尺寸为 7×7 的特征图,将 512 个特征图进行全连接运算得到一维特征向量

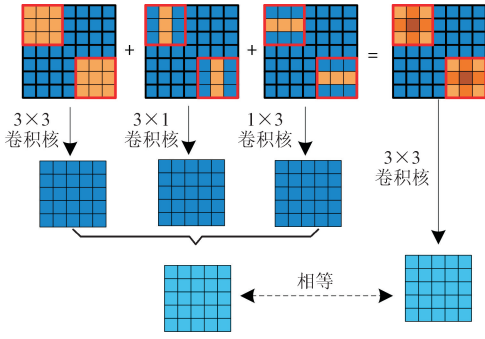


图 3 非对称卷积块结构图

Fig. 3 Structure of asymmetrical convolution block

F_1 , 使用 Softmax 函数对 F_1 进行运算, 得到大小为 $[5, 1]$ 的特征向量 F_2 , F_2 表示图像分属于每个标签的概率。

在 AKC-net 模型训练过程中, 采用前向传播和反向传播两个阶段表示卷积参数的传递和优化。在前向传播中读取输入图像的特征信息, 根据卷积神经网络的稀疏特性, 只有输入层的局部神经元与下一层的卷积神经元相连接。同时, 引入非线性函数作为激活函数, 使一部分神经元输出为 0, 减少层参数之间的依存关系。卷积层和最大池化层的前向传播表达式分别为

$$X_j = f\left(\sum_{i=1}^c X_i * K_{ij} + b_j\right) \quad (2)$$

$$W_j = f(\beta_j \text{down}(X_j) + \alpha_j) \quad (3)$$

式中: X_i 表示第 i 个输入特征图; K 表示对应的卷积核; b_j 表示第 j 个输出特征图的偏置值; c 为输入特征图的个数; $f()$ 为 Rule 激活函数; β_j 和 α_j 分别表示每个输出特征图的乘性偏置和加性偏置; $\text{down}()$ 为下采样函数。

反向传播主要是对网络参数进行更新优化, 通过链式法则求出每个参数的梯度。网络参数的更新量 $\Delta\alpha_i$ 可以描述为

$$\Delta\alpha_i = \eta \frac{\partial L}{\partial \alpha_i} = \eta \frac{\partial L}{\partial z_i} \frac{\partial z_i}{\partial \alpha_i} \quad (4)$$

式中: η 表示学习率; L 表示损失函数; z_i 表示层对应的输出; α_i 为层的参数。

1.3 聚焦程度拟合计算

分类网络仅能得到图像有限的离散聚焦程度。为了实现图像连续聚焦程度的检测, 本研究在上述分类结果的基础上采用三次样条插值法, 通过图像在有限聚焦程度处的分类分数, 估算出图像的实际聚焦程度。设在区间 $[0, \sigma_{\max}]$ 上给定一组节点, $0 =$

$$x_0 < x_1 = \frac{\sigma_{\max}}{n-1} < x_2 = \frac{2\sigma_{\max}}{n-1} < \dots < x_n = \sigma_{\max},$$

令 $f(x_i) = y_i$, y_i 为全连接层输出的分类分数, 若函数 $F(x)$ 在子区间 $[x_i, x_{i+1}]$ 上为不超过 3 次的多项式, $F(x)$ 的二阶导数在 $[0, \sigma_{\max}]$ 连续, $F(x_i) = y_i$, 则利用 $F(x)$ 对 y_i 和高斯模糊标准差进行拟合计算, 得到分数最大值对应的高斯模糊标准差, 以此作为图像的聚焦程度评价结果。

2 实验与性能测试

2.1 实验设置

本研究在处理器为 Intel(R)Core(TM)i7-9750H CPU、内存为 8 GB、显卡为 NVIDIA GeForce GTX1650 的计算机上进行实验。考虑到仿真实验能够获取图像高斯模糊标准差的真值, 因此采用仿真实验验证所提方法的有效性, 最终通过实拍实验进一步检验该方法对实际拍摄图像的聚焦程度评价效果。

仿真实验采用全为聚焦图像的 Waterloo 数据集^[15]作为原始数据集, 对比实验选择常用的 LIVE 数据集^[16]和 TID2008 数据集^[17]中的聚焦图像。对 3 个数据集中所有图像进行灰度化处理, 图像像素设置为 224×224 。从 Waterloo 数据集中随机选择 1 000 张图像进行高斯模糊标准差为 0.5、1.0、1.5、2.0 的离焦处理, 得到的图像构成网络的训练集和验证集, 再从灰度化处理后的 Waterloo 数据集中另选 500 张图像进行高斯模糊标准差分别为 0.1、0.2、...、1.9 和 2.0 的离焦处理构成测试集。图 4 中第 1 列表示 Waterloo 数据集中部分灰度化原图, 其余列

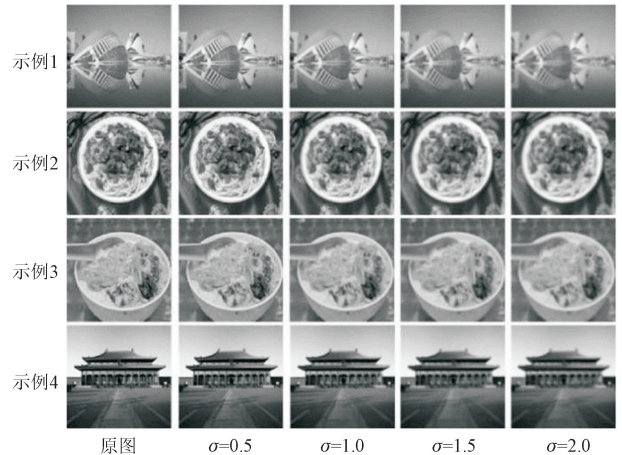


图 4 Waterloo 数据集中部分灰度化原图与相应离焦处理后的图像

Fig. 4 Partial images and corresponding defocus images in Waterloo dataset

表示对原图进行高斯模糊标准差分别为0.5、1.0、1.5、2.0的离焦处理后的图像。后续实验分别对构建的分类网络模型优越性、采用的数据拟合方法有效性、整个聚焦程度评价方法的优越性以及实际离焦图像的有效性进行验证。

2.2 分类网络优越性验证实验

为了验证本文方法构建的 AKC-net 模型在模糊图像分类性能上的优越性,将 AKC-net 模型与广泛应用的 AlexNet^[18]、VGGNet-16^[19]、GoogLeNet^[20]、ResNet50^[21]、SqueezeNet^[22]、MobileNetV2^[23] 等网络模型和先进的 ACNet^[14]、Swin-T^[24]、ConvNeXt^[25] 等网络模型进行对比。其中,各网络的动量初始化设置为 0.9,求解器选择带动量的随机梯度下降优化器,初始学习率设置为 0.005,每次训练迭代的最小批次设置为 32,训练轮数 40 轮。表 1 展示了各网络在测试集上的平均准确率和损失值,可以看到其他网络在测试集中的平均准确率最高为 92.4%,而 AKC-net 的平均准确率可达 97.7%,比其他网络最高准确率高出 5.3 个百分点,且损失值最小,分类效果得到显著提升。因此,AKC-net 在模糊图像分类性能方面优于其他网络。

表 1 各网络模型在测试集上的平均准确率和损失值
Table 1 The average accuracy and loss of each network on test dataset

模型	平均准确率/%	损失值
AlexNet ^[18]	70.2	0.57
VGGNet-16 ^[19]	91.4	0.21
GoogLeNet ^[20]	90.2	0.29
ResNet50 ^[21]	87.3	0.41
SqueezeNet ^[22]	89.5	0.34
MobileNetV2 ^[23]	90.6	0.26
ACNet ^[14]	92.4	0.18
Swin-T ^[24]	90.5	0.27
ConvNeXt ^[25]	87.9	0.36
AKC-net	97.7	0.13

注:加粗表示最优结果。

2.3 数据拟合有效性验证实验

为验证三次样条插值法拟合效果的有效性,将三次样条插值法与三次多项式、四次多项式等拟合方法进行对比,采用最大绝对误差($E_{\max}$)和平均绝对误差(E_{mean})作为数据拟合的评价标准。为了避免随机选择图像所引起的实验不确定性,每次实验任意选取测试集中 100 张图像,共进行了 3 次实验,结果如表 2 所示。从表 2 可以看出,在 3 次实验

中,三次样条插值法的 $E_{\max}$ 和 E_{mean} 均是最小的,表明三次样条插值法与全连接层输出的分类分数拟合效果较好,在评价图像聚焦程度上具有较高的准确性。

表 2 不同拟合方法的最大绝对误差和平均绝对误差
Table 2 Maximum absolute error and mean absolute error for different fitting methods

拟合方法	第 1 次实验		第 2 次实验		第 3 次实验	
	$E_{\max}$	E_{mean}	$E_{\max}$	E_{mean}	$E_{\max}$	E_{mean}
三次多项式	0.34	0.12	0.34	0.11	0.34	0.12
四次多项式	0.22	0.08	0.23	0.09	0.22	0.11
三次样条插值	0.20	0.06	0.20	0.07	0.19	0.09

注:加粗表示最优结果。

2.4 与图像内容无关的聚焦程度评价方法优越性验证实验

为验证本文提出的“分类+数据拟合”思想与聚焦程度评价方法的优越性,将本文方法与单纯 AKC-net 分类的聚焦程度评价方法(F-AKCnet)、基于 AKC-net 回归的聚焦程度评价方法(R-AKCnet)以及文献[10-11]、文献[13]和文献[26]的与图像内容无关的聚焦程度评价方法进行对比。

考虑到各方法聚焦测度值的量纲不一致,对所有数值进行归一化处理。采用均方根误差(E_{RMS})、平均绝对误差(E_{mean})、Pearson 相关系数(P)、Spearman 等级相关系数(S)作为聚焦程度评价标准。 E_{RMS} 和 E_{mean} 越小, P 越大,表示图像聚焦程度评价准确率越高。 E_{RMS} 和 P 表达式如下

$$E_{\text{RMS}} = \sqrt{\frac{1}{m} \sum_{k=1}^m (f_k - f_k^*)^2}$$
 (5)

$$P = \frac{\sum_{k=1}^m (f_k - \bar{f}_k)(f_k^* - \bar{f}_k^*)}{\sqrt{\sum_{k=1}^m (f_k - \bar{f}_k)^2 \sum_{k=1}^m (f_k^* - \bar{f}_k^*)^2}}$$
 (6)

式中: m 为数据集的图像数量; f_k 表示第 k 张图像归一化后的真实聚焦测度值; f_k^* 表示第 k 张图像归一化后的预测聚焦测度值; $\bar{f}_k$ 和 $\bar{f}_k^*$ 分别表示所有图像的 f_k 和 f_k^* 的平均值。

S 表示图像聚焦程度评价方法的预测单调性, S 越大表明 f_k^* 随着 f_k 的增大而增大的趋势越明显,其表达式为

$$S = 1 - \frac{6 \sum_{k=1}^m d_k^2}{m(m^2 - 1)}$$
 (7)

式中: d_k 表示第 k 张图像的 f_k 和 f_k^* 对应的等级排

序位置的差。

表 3 展示了各方法在测试集、LIVE 数据集和 TID2008 数据集上的评价结果,可以看出,在 3 个数据集中,Final^[10]、CEWD^[11]和 PSI^[26]方法的 E_{RMS} 和 E_{mean} 较大, P 和 S 较小,说明这 3 种方法在图像内容变化时对聚焦测度值影响较大。从表 3 可以看到,SFA^[13]方法在多个评价指标中取得了仅次于本文方法的效果,但存在网络模型参数量大的问题,其参数

量为 24.42 MB,远大于本文方法的参数量 14.11 MB,且本文在测试集中 E_{RMS} 和 E_{mean} 均小于 0.07,因此本文方法在评价不同内容图像的聚焦程度性能方面要优于其他评价方法。此外,本文方法在 3 个数据集上的表现优于基于分类的方法 F-AKCnet 和基于回归的方法 R-AKCnet,证明了“分类+数据拟合”思想在与图像内容无关的聚焦程度评价方面优于分类思想和回归思想。

表 3 不同方法在测试集、LIVE 和 TID2008 数据集上的评价结果
Table 3 Evaluation result for different methods on test set, LIVE and TID2008 dataset

数据集	方法	E_{RMS}	E_{mean}	P	S
测试集	Final ^[10]	0.181 2	0.160 7	0.774 8	0.751 2
	CEWD ^[11]	0.133 1	0.137 2	0.834 9	0.794 5
	SFA ^[13]	0.079 1	0.071 9	0.891 7	0.887 9
	PSI ^[26]	0.104 9	0.093 5	0.874 1	0.864 9
	F-AKCnet	0.126 2	0.118 6	0.901 6	0.893 7
	R-AKCnet	0.119 5	0.107 7	0.918 2	0.891 1
	本文方法	0.069 3	0.067 1	0.931 0	0.900 4
LIVE	Final ^[10]	0.149 0	0.135 5	0.841 7	0.839 5
	CEWD ^[11]	0.127 8	0.120 3	0.892 5	0.889 3
	SFA ^[13]	0.081 5	0.079 3	0.923 3	0.918 1
	PSI ^[26]	0.131 1	0.126 9	0.862 3	0.861 2
	F-AKCnet	0.108 3	0.101 1	0.916 6	0.902 0
	R-AKCnet	0.102 6	0.098 8	0.929 9	0.921 0
	本文方法	0.070 2	0.068 2	0.936 8	0.922 8
TID2008	Final ^[10]	0.180 2	0.185 5	0.831 9	0.840 8
	CEWD ^[11]	0.132 9	0.131 6	0.858 6	0.846 2
	SFA ^[13]	0.081 0	0.073 9	0.898 3	0.892 1
	PSI ^[26]	0.145 2	0.131 3	0.861 0	0.858 5
	F-AKCnet	0.102 5	0.093 6	0.901 4	0.891 9
	R-AKCnet	0.103 9	0.089 5	0.911 4	0.913 2
	本文方法	0.075 6	0.072 2	0.920 7	0.911 1

注:加粗表示最优结果。

2.5 统计显著性实验

为验证 2.4 节中各方法的比较结果是否具有显著性,对各方法的评价结果与高斯模糊标准差真值之间的残差进行统计学显著性检验。

利用不同方法对 3 个数据集上的图像进行聚焦测度值的计算,以 1% 的显著性水平进行双样本 t 检验。表 4 展示了本文方法对其他方法的统计学显著

性检验结果,其中,1 表示本文方法在统计上优于比较方法,0 表示本文方法在统计上与比较方法相当。从表 4 可以看出,在测试集中,本文方法优于其他方法;尽管在 LIVE 和 TID2008 数据集上,SFA 和 R-AKCnet 方法取得了与本文方法相当的性能,但是与图像内容无关的性能不应局限于某个特定的数据集,因此本文方法优于对比的 6 种方法。

表 4 不同方法在测试集、LIVE 和 TID2008 数据集上的 t 检验结果

Table 4 The t -test of different methods on test set, LIVE and TID2008 dataset

方法	t 检验结果		
	测试集	LIVE 数据集	TID2008 数据集
Final ^[10]	1	1	1
CEWD ^[11]	1	1	1
SFA ^[13]	1	0	1
PSI ^[26]	1	1	1
F-AKCnet	1	1	1
R-AKCnet	1	1	0

2.6 计算效率

为比较不同方法的计算效率,对测试集中的 500 张聚焦图像进行聚焦程度评价,表 5 展示了每张图像在不同方法上的平均计算时间。可以看出,本文方法的计算时间远低于所对比的方法,表明其计算效率要高于其他方法,计算速度更快,具有更好的实时性。

表 5 不同方法在测试集上的平均计算时间

Table 5 The average computation time of different methods on test set

方法	平均计算时间/s
Final ^[10]	0.013 3
CEWD ^[11]	0.011 6
SFA ^[13]	0.009 2
PSI ^[26]	0.020 5
F-AKCnet	0.009 7
R-AKCnet	0.010 2
本文方法	0.008 9

注:加粗表示最优结果。

2.7 实拍实验

为进一步检验本文方法对实拍图像聚焦程度的评价效果,利用移动镜头的对焦系统分别拍摄 5 个物体(依次为分辨率板、印刷盒、咖啡盒、测量仪盒、打印盒)。首先,调节系统使其准确对焦到第 1 个平面物体上,拍摄第 1 幅图像,然后等间距移动镜头并将拍摄目标依次更换为第 2~5 个平面物体(物体位置保持不变),拍摄第 2~5 幅图像,实验拍摄的 5 幅图像如图 5 所示。

由于实验拍摄的第 1 幅图像处于对焦状态,后续图像都是通过移动等间距的镜头拍摄得到,由几



图 5 实际拍摄的图像
Fig. 5 Real shooting images

何光学可知,这 5 幅图像的弥散圆直径存在线性关系,由于高斯模糊标准差与弥散圆直径也存在线性关系^[27],因此这 5 幅图像的高斯模糊标准差也存在线性关系。利用本文方法计算 5 幅图像的聚焦测度值,结果如图 6 所示。由图 6 可知,本文方法计算 5 幅图像的聚焦测度值近似为线性关系,与理论结果一致,并且不随图像内容的改变而变化,说明本文提出的聚焦程度评价方法与图像内容无关,验证了方法的有效性。

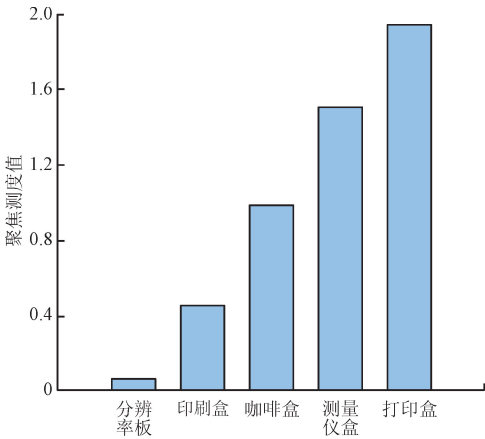


图 6 实际拍摄的图像聚焦测度值
Fig. 6 The focus measure value of real shooting images

3 结 论

本文针对现有聚焦程度评价方法对图像内容有较大依赖性的问题,研究了与图像内容无关的聚焦程度评价方法,主要结论如下。

(1)构建了用于提取图像高斯模糊标准差分类

分数的卷积神经网络模型 AKC-net,采用非对称卷积核增强模型对图像的特征提取能力,平均准确率可达 97.7%,对比比模型的最高准确率高 5.3 个百分点,验证了所构建模型在模糊图像分类性能上的优越性。

(2)基于“分类+拟合”思想,提出了一种非对称核卷积神经网络与数据拟合相结合的聚焦程度评价方法,采用三次样条插值函数拟合 AKC-net 全连接层输出的分类分数以及对应的高斯模糊标准差,以最大分数对应的标准差作为图像的聚焦程度评价结果。仿真与实拍实验结果表明,本文提出的方法对不同内容图像具有更好的聚焦程度评价性能且适用于对实时性要求更高的场合。

本文方法评价的是固定尺寸的图像,在未来的工作中,有必要对变尺寸图像的聚焦程度评价方法进行研究。

参考文献:

[1] HELMY I, CHOI W. Machine learning-based automatic focusing for high magnification systems [J]. Engineering Applications of Artificial Intelligence, 2023, 118: 105648.

[2] 杨梦雪,李祝莲,李荣旺,等. 基于 1.2 m 望远镜系统实时自动调焦的清晰度评价算法 [J]. 光学学报, 2023, 43(6): 0612006.

YANG Mengxue, LI Zhulian, LI Rongwang, et al. Sharpness evaluation algorithm based on real-time automatic focusing of 1.2 m telescope system [J]. Acta Optica Sinica, 2023, 43(6): 0612006.

[3] 蒋留兵,张点,潘波,等. 基于生成对抗网络的多聚焦图像融合 [J]. 计算机辅助设计与图形学学报, 2021, 33(11): 1715-1725.

JIANG Liubing, ZHANG Dian, PAN Bo, et al. Multi-focus image fusion based on generative adversarial network [J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(11): 1715-1725.

[4] TANG Han, XIAO Bin, LI Weisheng, et al. Pixel convolutional neural network for multi-focus image fusion [J]. Information Sciences, 2018, 433-434: 125-141.

[5] 韦号,崔海华,程筱胜,等. 一种用于评价聚焦形貌恢复算法的图像离焦仿真技术 [J]. 光学学报, 2019, 39(11): 1111001.

WEI Hao, CUI Haihua, CHENG Xiaosheng, et al. Image defocus simulation technology applied to evaluation of focused morphology recovery algorithm [J]. Acta Optica Sinica, 2019, 39(11): 1111001.

[6] XIA Xiaohua, YIN Lijuan, YAO Yunshi, et al. Combining two focus measures to improve performance [J]. Measurement Science and Technology, 2017, 28(10): 105401.

[7] LI Qiaoyue, LI Leida, LU Zhaolin, et al. No-reference sharpness index for scanning electron microscopy images based on dark channel prior [J]. KSII Transactions on Internet and Information Systems, 2019, 13(5): 2529-2543.

[8] 周汶,李旦,张建秋. 评价彩色图像自动聚焦清晰度的互补色小波测度 [J]. 自动化学报, 2020, 46(8): 1615-1627.

ZHOU Wen, LI Dan, ZHANG Jianqiu. A complementary color wavelet-based measure on color image sharpness assessment for autofocus [J]. Acta Automatica Sinica, 2020, 46(8): 1615-1627.

[9] BAHRAMI K, KOT A C. Efficient image sharpness assessment based on content aware total variation [J]. IEEE Transactions on Multimedia, 2016, 18(8): 1568-1578.

[10] ZHAN Yibing, ZHANG Rong. No-reference image sharpness assessment based on maximum gradient and variability of gradients [J]. IEEE Transactions on Multimedia, 2018, 20(7): 1796-1808.

[11] LIU Zhaoyang, HONG Huajie, GAN Zihao, et al. An improved method for evaluating image sharpness based on edge information [J]. Applied Sciences, 2022, 12(13): 6712.

[12] LIU Yu, CHEN Xun, PENG Hu, et al. Multi-focus image fusion with a deep convolutional neural network [J]. Information Fusion, 2017, 36: 191-207.

[13] LI Dingquan, JIANG Tingting, LIN Weisi, et al. Which has better visual quality: the clear blue sky or a blurry animal? [J]. IEEE Transactions on Multimedia, 2019, 21(5): 1221-1234.

[14] DING Xiaohan, GUO Yuchen, DING Guiguang, et al. ACNet: strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2019: 1911-1920.

[15] MA Kede, DUANMU Zhengfang, WU Qingbo, et al. Waterloo exploration database: new challenges for image quality assessment models [J]. IEEE Transactions on Image Processing, 2017, 26(2): 1004-1016.

[16] SHEIKH H R, SABIR M F, BOVIK A C. A statistical evaluation of recent full reference image quality assessment algorithms [J]. IEEE Transactions on Image

- Processing, 2006, 15(11): 3440-3451.
- [17] PONOMARENKO N. TID2008-a database for evaluation of full-reference visual quality assessment metrics [J]. Adv Modern Radioelectron, 2009, 10(4): 30-45.
- [18] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [19] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2015-04-10) [2022-02-04]. <https://arxiv.org/abs/1409.1556>.
- [20] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going deeper with convolutions [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2015: 1-9.
- [21] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 770-778.
- [22] IANDOLA F N, HAN Song, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size [EB/OL]. (2016-11-04) [2024-05-01]. <https://arxiv.org/abs/1602.07360>.
- [23] HOWARD A G, ZHU Menglong, CHEN Bo, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [EB/OL]. (2017-04-17) [2024-05-01]. <https://arxiv.org/abs/1704.04861>.
- [24] LIU Ze, LIN Yutong, CAO Yue, et al. Swin transformer: hierarchical vision transformer using shifted windows [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2021: 9992-10002.
- [25] LIU Zhuang, MAO Hanzi, WU Chaoyuan, et al. A ConvNet for the 2020s [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 11966-11976.
- [26] FEICHTENHOFER C, FASSOLD H, SCHALLAUER P. A perceptual image sharpness metric based on local edge gradient analysis [J]. IEEE Signal Processing Letters, 2013, 20(4): 379-382.
- [27] SUBBARAO M, SURYA G. Depth from defocus: a spatial domain approach [J]. International Journal of Computer Vision, 1994, 13(3): 271-294.

(编辑 亢列梅)