

# 基于深层卷积网络的单幅图像超分辨率重建模型\* Single Image Super-Resolution Restoration Model Using Deep Convolutional Networks

龙法宁<sup>1</sup>,朱晓姝<sup>1,2\*\*</sup>,胡春娇<sup>3</sup>

LONG Fanning<sup>1</sup>,ZHU Xiaoshu<sup>1,2</sup>,HU Chunjiao<sup>3</sup>

(1. 玉林师范学院计算机科学与工程学院,广西玉林 537000;2. 玉林师范学院广西高校复杂系统优化与大数据处理重点实验室,广西玉林 537000;3. 玉林师范学院教育技术中心,广西玉林 537000)

(1. Computer Science Department, Yulin Normal University, Yulin, Guangxi, 537000, China; 2. Guangxi Universities Key Lab of Complex System Optimization and Big Data Processing, Yulin Normal University, Yulin, Guangxi, 537000, China; 3. Educational Technology Center, Yulin Normal University, Yulin, Guangxi, 537000, China)

**摘要:**【目的】针对 Mean squared error(MSE)作为损失函数在人眼感知方面存在局限性,以及基于卷积神经网络的图像超分辨率(Super-resolution,SR)算法生成的图像存在参数较多、计算量较大、训练时间较长、纹理模糊等问题,设计基于深层卷积神经网络的单幅图像超分辨率重建模型。【方法】使用 ImageNet 预先训练的大型卷积神经网络 Visual geometry group(VGG)模型提取图像特征,利用该特征设计视觉感知损失函数进行训练学习,引入亚像素卷积层(Sub-pixel convolution)替换上采样层,缓解生成图像的棋盘效应。【结果】设计的模型对放大两倍的图像进行超分辨率修复,与其他4种超分辨率重建模型的 Peak signal to noise ratio(PSNR)值接近,且生成图像的视觉效果更加清晰逼真,细节更加细腻。【结论】该模型可以实现输入不同大小的低分辨率图像而不必多次训练学习不同比例的放大模型,可以实现对不同放大倍数图像的训练和预测,在保持一定 PSNR 正确率的前提下,放大后的超分辨率图像能够恢复更多纹理细节和更佳视觉效果。

**关键词:**超分辨率 深度学习 感知损失函数 卷积神经网络

**中图分类号:**TP391.41 **文献标识码:**A **文章编号:**1005-9164(2017)03-0231-05

**Abstract:**【Objective】On the limitations of human perception in MSE (Mean square error) as one of loss functions, and flaws remained in images generated by the super-resolution algorithm based on the convolution neural network, such as excessive parameters, large calculated amount, long training time and fuzzy texture and so on. This research aimed at designing a single image super-resolution reconstruction model based on a deep convolution of the neural network. 【Methods】The deep convolution neural network optimized perceptual loss functions based on high-level features extracted from the pre-trained networks Visual Geometry Group (VGG), and a sub-pixel convolution layer was used to replace the upscale layer and effectively

relieve to generate images' checkerboard artifacts. 【Results】Our proposed methods doubled the magnification of the image super-resolution repair, which was close to the other four SRCNN SR algorithms and set a better photorealistic image SR. 【Conclusion】Our proposed method performs a better visual improvement in our results which are easily noticeable. **Key words:** super-resolution, deep learning, perceptual losses, convolutional neural networks

收稿日期:2017-01-23

修回日期:2017-03-14

作者简介:龙法宁(1978-),男,硕士,讲师,主要从事图像处理、机器学习方面的研究。

\* 广西重点实验室科研课题项目(2016CSOBDP0302)和广西高校科研项目(2013YB202)资助。

\*\* 通信作者:朱晓姝(1973-),女,硕士,硕士生导师,教授,主要从事生物信息、大数据分类、分布式网络计算研究, E-mail: jgxyzs@126.com。

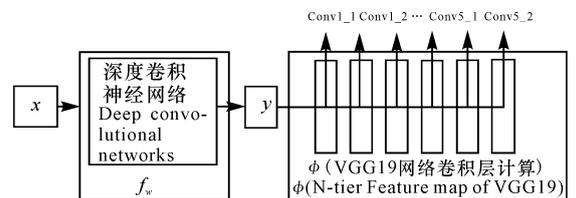
【研究意义】手机、平板等非专业照相设备的像素受到网络传输和拍摄距离的限制,导致大量手机图片或卫星影像等被人为压缩并降低画质。如今高分辨率显示屏在家庭和移动设备上已经普及,因此需要采用超分辨率技术对图像进行复原,提高清晰度,满足用户需求。【前人研究进展】图像超分辨率(Super-resolution, SR)处理技术是指根据一幅或者多幅低分辨率图像(Low-resolution, LR)来重建高分辨率高质量图像(High-resolution, HR)的技术。目前超分辨率技术主要分为两类:一是基于重建的超分辨率方法,利用多张图像的丰富信息,图像间的相同异或信息以及图像的先验知识进行重建,例如邻域像素重复技术和双三次图像插值技术(Bicubic)等,修复效果不佳。二是基于学习的超分辨率方法,目前用于单幅图像超分辨率的最先进方法大多是基于样本学习的。这些方法主要利用相同图像的内部相似性,或外部低和高分辨率的样本学习映射函数。基于样本的外部方法通常提供有大量丰富样本,但是受到有效紧凑建模数据困难的挑战<sup>[1]</sup>。随着深度学习在机器学习的多个领域取得显著的效果,许多基于深度卷积神经网络的超分辨率修复算法已经被广泛采用并取得了最佳效果。卷积神经网络被广泛应用于自然图像去噪, Dong 等<sup>[2]</sup>已经证明利用卷积神经网络可以用于学习 LR 到 HR 的映射函数,可以实现或多或少的消除噪声,其中比较成功的典型算法是 Super Resolution Convolutional Neural Network(SRCNN)算法。Kim 等<sup>[3]</sup>提出了一种利用 20 层的残差卷积网络进行学习预测的方法,采用比较大的学习率参数大幅提高训练速度,峰值信噪比(Peak Signal to Noise Ratio, PSNR)实验结果接近历史最好效果。Johnson 等<sup>[4]</sup>从人类视觉感知的角度思考,提出了一种基于感知损失函数的修复低分辨率视频图像方法,取得很好的视觉效果。改善图像光照补偿算法的效果,提高图像预处理的质,胡伟平、Shi 等<sup>[5-6]</sup>提出了亚像素卷积层(Sub-pixel convolution)的概念,通过在卷积神经网络的最后一层采用亚像素卷积层进行额外的超分辨率计算,获得很好的复原效果,并成功地用于超分辨率图像复原。【本研究切入点】PSNR 是评价图像的客观标准,是最普遍、最广泛使用于评鉴画质的客观量测法,特别是在超分辨率修复领域,常常针对 PSNR 进行损失函数(PSNR/MSE loss function)设计,大部分修复算法都在追求最高的 PSNR 数据。但其具有局限性,许多实验结果都显示,PSNR 的

数无法和人眼看到的视觉品质完全一致,有可能 PSNR 较高者看起来反而比 PSNR 较低者模糊。这是因为人眼的视觉对于误差的敏感度并不是绝对的,其感知结果会受到许多因素影响而产生变化,例如:人眼对空间频率较低的对比差异敏感度较高;人眼对亮度对比差异的敏感度较色度高;人眼对一个区域的感知结果会受到其周围邻近区域的影响。【拟解决的关键问题】参考 Johnson 等<sup>[4]</sup>的方法,考虑图像间的峰值、纹理、光照等人眼感知的差异,采用亚像素卷积层替换常用的反卷积层作为上采样层,缓解图像放大后的棋盘效应(Checkerboard artifacts),设计合理的深度卷积神经网络,采用基于视觉感知的损失函数进行训练学习,实现对放大两倍的图像进行超分辨率修复,并达到优秀视觉效果。

## 1 模型设计

### 1.1 深度学习模型

本研究设计的深度学习模型由两部分组成(图 1),一是学习 LR 到 HR 映射函数的深度卷积神经网络  $f_w$ ,其中  $W$  (weights)是  $f_w$  所有神经元的最优权值;二是采用由牛津视觉几何组(Visual Geometry Group, VGG)开发的卷积神经网络结构 VGG19 网络模型  $\phi$ ,用于计算损失函数。该模型通过卷积层提取不同层的卷积特征(Feature map),可以表达人类对图像的视觉感知情况(图 2)。利用卷积特征计算修复后预测图像和目标高分辨率图像的视觉感知对比损失情况,根据不同卷积层的特征图可以定义一系列的损失函数  $l_1, l_2, \dots, l_k$ 。假设一幅低分辨率图像  $x$  通过深层卷积神经网络  $f_w$  计算得到超分辨率输出图像  $y$ ,则定义  $y = f_w(x)$ ,每个损失函数计算高分辨率预测图像  $y$  和目标高分辨率图像  $\hat{y}$  的损失大小  $l_i(y_i, \hat{y}_i)$ ,损失函数可以从纹理、光照、直方图等多方面考虑。因此  $f_w$  的求解过程就转化为利用公式(1)进行梯度下降求最优解,其中  $\lambda_i$  为每个损失函数  $\ell_i$  的权重比例。



Conv1\_n 为 VGG19 网络模型第 n 层的卷积特征

Conv1\_n is N-tier Feature map of VGG19

图 1 网络模型

Fig. 1 Architecture of Network

$$W^* = \arg \min_w E_{x, \{y_i\}} \left[ \sum_{i=1}^k \lambda_i \ell_i(f_w(x), y_i) \right]. \quad (1)$$

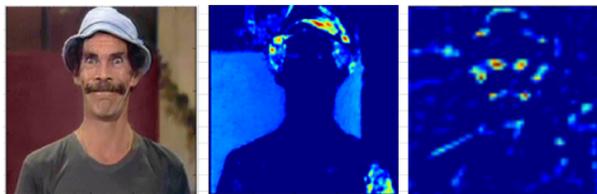
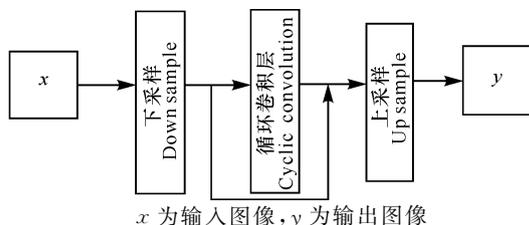


图2 VGG模型分层提取图像特征

Fig. 2 Feature maps of VGG Network

## 1.2 深度卷积神经网络

本研究参考 Kim 等<sup>[7]</sup>的模型设计思路,设计全卷积层的深度神经网络,深层卷积模型如图3所示。下采样模块采用  $4 \times 4$  大小滤波器(filter),步长为  $2 \times 2$  的卷积层,循环卷积模块包括8个相同的全卷积层,采用 PRelu 激活函数,卷积核的数量都为128个,滤波器大小为  $3 \times 3$ ,步长为  $1 \times 1$ ,外围边缘(padding)为1。上采样模块由一个卷积层和一个放大两倍图像尺寸的亚像素卷积层组成,其中卷积层的卷积核的数量为512,滤波器大小为  $3 \times 3$ ,步长为  $1 \times 1$ 。如果需要训练输出2倍超分辨率目标图像,则上采样模块数量  $\times 2$ ;需要训练4倍目标图像,则上采样模块数量  $\times 3$ ,以此类推。上采样模块连接一个滤波器大小为  $7 \times 7$ ,步长为  $3 \times 3$ ,卷积核的数量为3,padding为  $3 \times 3$  的卷积层,最终输出放大2倍或者4倍的目标预测图像。



$x$  为输入图像,  $y$  为输出图像

$x$  is the input image,  $y$  is the output image

图3 深层卷积神经网络

Fig. 3 Deep CNN Network

## 1.3 感知损失函数

与传统的基于像素级别损失函数的训练过程不同,针对人眼对图像高频信号更为敏感的特征,参考 Gatys 等<sup>[8]</sup>的方法,本研究定义了基于 VGG19 神经网络模型的 VGG 损失函数,VGG19 模型被设计用于预测图片的类别,采用官方提供的 VGG19 训练数据,把通过神经网络  $f_w$  获得的超分辨率图像  $I^{SR}$  和目标高清图像  $I^{HR}$  作为 VGG19 模型的输入图像,易于获得各卷积层的图像特征  $\phi$ ,定义  $\phi_{i,j}$  为 VGG19 的第  $j$  卷积层的特征图,计算均方误差 (Mean Squared Error, MSE),最终定义损失函数如公式(2)所示:

$$\sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(I^{SR})_{x,y})^2. \quad (2)$$

由于噪音信号(图像)相邻信号间不平滑,随机变化较大,通过最小化全变差正则化(Total variation regularization)可以去除噪音,平滑信号。为了缓解生成图像的棋盘效应,增加基于全变差正则化计算的损失函数,具体计算如公式(3)所示<sup>[9]</sup>:

$$l_{TV}^{SR} = \frac{1}{r^2 WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} \|\nabla \phi_{i,j}(I^{SR})_{x,y}\|. \quad (3)$$

## 1.4 训练过程

首先将高清图像随机裁剪为  $256 \times 256$  大小的 HR 子图像  $Y$ ,对 HR 子图像  $Y$  进行缩小 50% 处理,添加符合标准正态分布的噪点,进行高斯处理,得到“低分辨率”图像  $X$ ,输入图像  $X$  的大小为输出图像  $Y$  的  $1/2$ 。通过学习映射  $F$ ,能够输入低分辨率图像产生最终的高分辨率图像。

为了解决损失函数的最小化问题,训练过程采用 Adam 优化器,学习率为  $1e-4$ ,Mini-Batch 大小为 16。第一次初步训练采用 VGG 的 Conv1\_2 和 Conv2\_2 卷积层提取特征数据,用损失函数(2)计算均方误差然后进行梯度下降 75 个 epoch 计算最优解。第二次采用 VGG 的 Conv5\_2 卷积层提取特征数据,把损失函数与一个全变差正则损失函数相加进行辅助计算,训练过程适当增大全变差正则损失函数的权重比例,可以有效地缓解生成图像的棋盘效应,然后进行梯度下降 500 个 epoch 进行权值优化。当学习率不断下降以后,梯度下降越来越缓慢,采用梯度剪裁(Clip gradient)可以把全局的梯度二范数限制在一定的范围内,最终得到比较好的训练效果。

## 2 结果与分析

本研究使用谷歌公司最新发布的 Open Image 图片数据集训练一个  $\times 2$  的模型,该数据集包含约 900 万张图片,这些图片通过标签注释被分为 6 000 多类,都具有很高的分辨率。和其他一般图像数据集训练对比,实验证明采用高分辨率图片数据集相比较分辨率较低的图片数据集,能够更有效地避免神经网络预测生成图像的棋盘效应,有更强的鲁棒性。本研究随机从 Open Image 数据集中裁剪  $256 \times 256$  大小的 HR 子图像 200 000 张,对训练图像进行高斯模糊和增加噪点,然后缩小 50% 作为低分辨率输入图像。

本研究评估数据集使用超分辨率研究领域标准评测数据集 Set5<sup>[10]</sup>,Set14<sup>[11]</sup>,BSD100<sup>[12]</sup>,评测标准为 PSNR,单位是 dB,如公式(4)所示:

$$PSNR = 10 \times \log_{10} \left( \frac{(2^n - 1)^2}{MSE} \right), \quad (4)$$

其中, MSE 是原图像与处理图像之间均方误差;  $(2^n - 1)^2$  的是信号最大值的平方;  $n$  是每个采样值的比特数。

对比 Bicubic 算法及基于卷积神经网络进行超分辨率复原的 SRCNN、ESRCNN (Expanded Super Resolution Convolutional Neural Network) 和 DDSRCNN (Deep Denoising Super Resolution Convolutional Neural Network) 模型 (其中, ESRCNN 为 SRCNN 的改进模型, 增加了更多卷积层), 本模型较好保持了和该 3 个模型相近的 PSNR 数值 (表 1), 本研究的超分辨率模型的修复效果清晰度更高, 纹理细节更加明显 (图 4~5)。

表 1 5 种超分辨率重建模型的 PSNR 比较

Table 1 PSNR comparison of five super-resolution reconstruction models

数据集 Data set	Bicubic	SRCNN <sup>[1]</sup>	ESRCNN	DDSRCNN <sup>[13]</sup>	本模型 This model
Set5	33.95	36.66	33.19	36.55	33.58
Set14	32.49	32.42	31.83	32.47	32.18
BSD100	32.47	31.36	31.14	32.35	31.92

图 4~5 从左到右依次为原始高清图片, Bicubic 算法、SRCNN 算法和 DDSRCNN 算法处理后的图片, 本研究优化算法生成的超分辨率图像。从图中可以看出 Bicubic 算法处理后的图片是最模糊的, 本研究算法的处理效果纹理细节比较清晰。

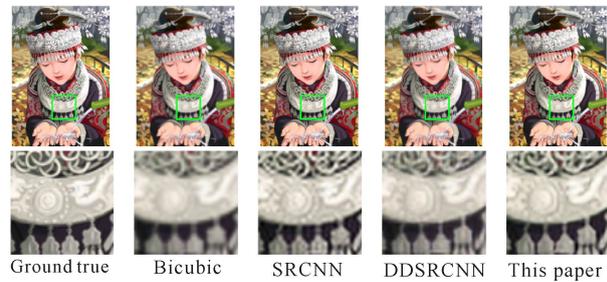


图 4 人物图像超分辨率复原结果

Fig. 4 Super-resolution results of figure image

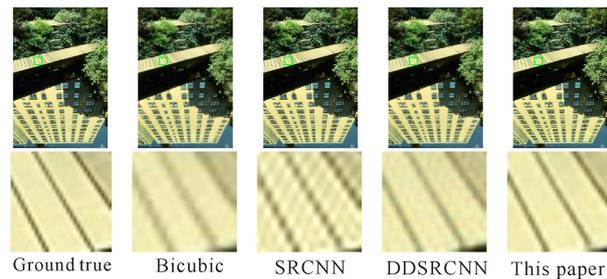


图 5 景物图像超分辨率复原结果

Fig. 5 Super-resolution results of scene image

### 3 结论

本研究设计了一种全部采用卷积层 (无池化层) 的神经网络模型, 结合亚像素卷积层对图像进行上采

样, 并采用基于视觉感知的损失函数进行训练学习。经验证, 该模型可以实现输入不同大小的低分辨率图像而不必多次训练学习不同比例的放大模型, 实现对不同放大倍数图像的训练和预测。在保持一定 PSNR 正确率的前提下, 该模型对放大后的超分辨率图像能够恢复更多纹理细节和更佳视觉效果, 同时发现输入图像分辨率大小、是否进行模糊预处理等操作, 对生成图像的棋盘效应影响很大。下一步工作考虑不仅仅通过 VGG 模型获得图像特征信息计算损失函数, 尝试结合图片的光照、颜色、形状等特征信息共同参与模型的训练和学习, 并进一步研究减少训练时间的方法。

### 参考文献:

- [1] DONG C, LOY C C, HE K, et al. Learning a deep convolutional network for image super-resolution [M]// FLEET D, PAJDLA T, SCHIELE B, et al (eds.). Computer Vision-ECCV 2014. ECCV 2014. Cham: Springer, 2014.
- [2] DONG C, LOY C C, HE K, et al. Image super-resolution using deep convolutional networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295-307.
- [3] KIM J, LEE J K, LEE K M, et al. Accurate image super-resolution using very deep convolutional networks [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: NV, United States, 2016.
- [4] JOHNSON J, ALAHI A, LI F F. Perceptual losses for real-time style transfer and super-resolution [C]// LEIBE B, MATAS J, SEBE N, et al (eds.). Computer Vision-ECCV 2016. Cham: Springer, 2016.
- [5] 胡伟平. 基于对称区块的人脸图像光照补偿算法 [J]. 广西科学院学报, 2015, 31(1): 64-68.
- [6] HU W P. Face illumination compensation algorithm based on symmetrical blocks [J]. Journal Guangxi Academy of Sciences, 2015, 31(1): 64-68.
- [7] SHI W, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016.
- [8] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C]// Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA: IEEE, 2016.
- [9] GATYS L A, ECKER A S, BETHGE M. Texture syn-

thesis using convolutional neural networks [C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge, MA, USA: MIT Press, 2015: 262-270.

[9] ALY H A, DUBOIS E. Image up-sampling using total-variation regularization with a new observation model [J]. IEEE Transactions on Image Processing, 2005, 14 (10): 1647-1659.

[10] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on nonnegative neighborhood embedding [C]//Proceedings of 2012 British Machine Vision Conference (BMVC). Surrey: BMVC, 2012.

[11] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse representations [M]//BOISSON-

NAT J D (ed.), Curves and Surfaces. Berlin: Springer, 2010.

[12] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars [C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA: IEEE, 2015.

[13] MAO X J, SHEN C, YANG Y B. Image restoration using convolutional auto-encoders with symmetric skip connections [EB/OL]. [2016-12-20]. <https://arxiv.org/pdf/1606.08921v3>.

(责任编辑: 陆 雁)

## 投稿指南

### 1 来稿要求

#### 1.1 稿件要素

稿件内容必须包括题目、作者姓名、作者所在单位、作者所在省份和城市、邮政编码、中文摘要、关键词、英文题目、作者英文姓名、作者英文单位、英文摘要、英文关键词、正文、致谢(非必选)、参考文献等内容。

#### 1.2 题目

应以简明、确切的语言反映稿件的重要思想和内容,一般不超过 20 字。

#### 1.3 作者与单位

多位作者姓名用逗号隔开。所有作者均须注明所在单位全称、省份城市及邮编。

#### 1.4 汉语姓名译法

姓在前名在后,姓用大写字母,名首字母大写(如:欧阳奋发, OUYANG Fenfa)。

#### 1.5 中、英文摘要

用第三人称撰写,应完整准确概括论文的实质性内容,试验研究论文摘要须标注【目的】……【方法】……【结果】……【结论】……4 个要素。英文摘要与中文摘要内容相对应(Abstract:【Objective】……【Methods】……【Results】……【Conclusion】……)。

#### 1.6 首页脚注标识要素

资助项目:项目名称(项目编号)。作者简介包括姓名(出生年-),性别,职称或职务,主要研究方向。如有通信作者,请注明×××为通信作者,包括姓名(出生年-),性别,职称或职务,主要研究方向,E-mail。

#### 1.7 稿件正文

试验研究论文应包括引言、材料与方法、结果与分析、讨论、结论等要素。引言须标注【研究意义】……【前人研究进展】……【本研究切入点】……【拟解决的关键问题】……等基本内容,“讨论”与“结论”部分须分开阐述。各层次标题用阿拉伯数字连续编号,如 0;1,1.1,1.1.1,1.1.1……;2,2.1,2.1.1……层次划分一般不超过 3 级。

#### 1.8 参考文献

所有类型的中文文献须对应翻译成英文。编排格式如下:

[1] 陈宝玲,宋希强,余文刚,等.濒危兰科植物再引入技术及其应用[J].生态学报,2010,30(24):7055-7063.

CHEN B L, SONG X Q, YU W G, et al. Re-introduction technology and its application in the conservation of endangered orchid[J]. Acta Ecologica Sinica, 2010, 30(24): 7055-7063.

#### 1.9 图和表

稿件可附必要的图和表,表用三线表表示,忌与文字表述重复,表的主题标目要明确。图表名、图表注及图表中所有的中文须有英文对照。图要大小适中,清晰,标注完整;照片尽量选用黑白照片。

#### 1.10 量和单位

量名称及其符号须符合国家标准,采用法定计量单位(用国际通用符号,如面积单位“亩”换算成“公顷 hm<sup>2</sup>”)。书写要规范化,并注明外文字母的大小写、正斜体及上下角标。容易混淆的字母、符号,请特别注明。

### 2 注意事项

2.1 本刊已开通网络投稿系统,投稿请登陆 <http://gxkk.cbpt.cnki.net/>,使用网上投稿和查稿系统。我刊审稿周期为 1 个月,1 个月后未收到审稿结果可另投他刊。

2.2 稿件一经采用,酌收版面费;刊登后,付稿酬含(《中国学术期刊(光盘版)》、中国期刊网、万方数据网及台湾华艺 CEPS 中文电子期刊服务网等)网络发行的稿酬,同时赠送样刊 2 本。

2.3 本刊入编《中国学术期刊(光盘版)》、中国期刊网、万方数据网及台湾华艺 CEPS 中文电子期刊数据库并已签订 CNKI 优先数字出版合作协议。

2.4 囿于人力、物力有限,本刊只通过期刊采编系统发送“稿件处理意见”,如需纸质意见,请向编辑部索取。