

基于 CNN 的中文字体图像超分辨率重构算法

李 昕^{1,2}, 雷雨晴^{1,2}, 闫 宇^{1,2}

(1. 大连民族大学 计算机科学与工程学院, 辽宁 大连 116650;

2. 大连市计算机字库设计技术创新中心, 辽宁 大连 116605)

摘 要:提出一种针对中文字体生成图像的超分辨率字体重构方法,以提高字体图像矢量化化的准确度。结合字体行业领域知识和图像处理技术将网络生成的 9 169 个字体图像进行超分辨率重构,该方法联合优化了所有网络层,使用双三次插值将单个低分辨率字体图像放大到所需的大小,通过卷积神经网络(CNN)拟合非线性映射输出较高分辨率的字体图像。实验结果表明:提出的方法生成的字体分辨率更高,能更好辅助字体设计师进行字体设计。

关键词:超分辨率重构;深度卷积网络;字体设计

中图分类号:TP391

文献标志码:A

Super-resolution Reconstruction Algorithm for Chinese Font Image Based on CNN

LI Xin^{1,2}, LEI Yu-qing^{1,2}, YAN yu^{1,2}

(1. School of Computer Science and Engineering, Dalian Minzu University, Dalian Liaoning 116650, China;

2. Dalian Chinese Font Design Technology Innovation Center, Dalian Liaoning 116605, China)

Abstract: A super-resolution font reconstruction method for Chinese font generated images is proposed to improve the accuracy of font image vectorization. In this paper, 9169 font images generated by the network are reconstructed by combining the domain knowledge of font industry and image processing technology. This method optimizes all network layers, uses bicubic interpolation to enlarge a single low-resolution font image to the desired size, and outputs a higher-resolution font image by fitting nonlinear mapping with convolution neural network (CNN). The experimental results show that the font generated by this method has higher resolution and can assist font designers in font design better.

Key words: super-resolution reconstruction; deep convolution neural network; font design

汉字字体不仅量大,而且复杂。一套字体设计周期长、反复工作量大,想要实现独特性表达几乎不能完成。设计一款新的字体,首先是对原始字稿进行设计。原始字稿中不仅包括设计师直接在计算机上创作的字稿还包括设计师手写的字稿。方正倩体的设计草图如图1。接着再针对设计师的手写字稿进行扫描和数字化拟合操作,将字稿扫描后的点阵图形制作成尽可能接近原稿的曲线轮廓信息。最后需要字体设计师们对字形进

行精细设计、反复修改,才能产生一款新的字体。



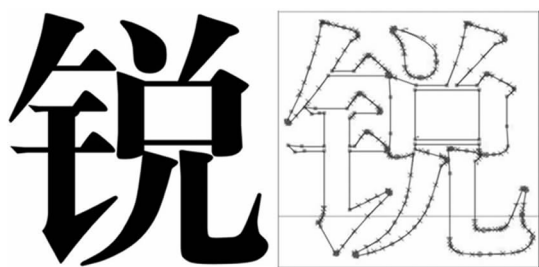
图1 方正倩体的设计草图

收稿日期:2021-06-13;最后修回日期:2021-08-05

基金项目:辽宁省自然科学基金项目(2020-MZLH-19);贵州省科技支撑计划项目(2021-534)。

作者简介:李昕(1997-),女,山东临沂人,大连民族大学计算机科学与工程学院硕士研究生,主要从事图像处理研究。

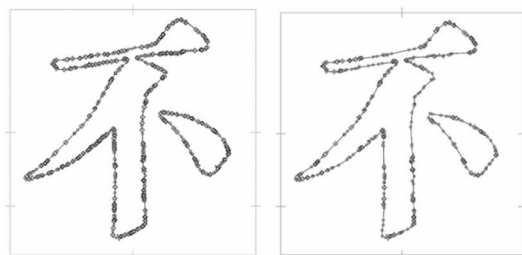
由于汉字字库是由矢量化曲线进行描述和表达,对于书法字体通常由字体图像进行矢量化转换,其中字体图像转为矢量化曲线是其中重要的环节之一。一个汉字的笔画周围的轮廓线可以看成曲线,曲线可以用一组首尾相接的矢量去近似地表示,字体矢量化如图2。每个矢量的起点称为曲线上的一个结点,于是,曲线就可以近似地用一组有次序结点的坐标来表示。用这种方法去描述汉字的字形,其信息量比直接用点阵描述一般要小得多。特别对于笔画比较平直的汉字,因为用到的结点很少,信息压缩的效果更明显。用这种方法做成的汉字字模叫做矢量字模,用矢量字模构成的字库就叫做矢量字库。矢量字库保存的是对每一个汉字的描述信息,比如一个笔划的起始、终止坐标,半径、弧度等等。在显示、打印这一类字库时,要经过一系列的数学运算才能输出结果,但是这一类字库保存的汉字理论上可以被无限地放大,笔划轮廓仍然能保持圆滑。



a) 标准字体图像 b) 写入字库后字体轮廓控制点

图2 字体图像矢量化

随着人工智能在字体设计领域应用的不断深入,深度学习的各种方法被广泛的用于字体设计中。尤其是生成式对抗网络方法,如 DCGAN、zi2zi 等方法可以利用小样本生成风格相近的中文字体图像再进行矢量化。在研究过程中发现字体图像分辨率的高低对于字体图像矢量化后的效果影响较大,如果字体图像分辨率低的话,点与点之间的距离会加大,矢量字体就会出现毛刺、冗余点过多等问题如图3。低分辨率的字体图像矢量化后字体边缘的锚点较多导致字体数据量更大,而高分辨率字体图像则可以生成锚点较少、边缘更为光滑曲线的矢量化字体。但用于字体生成的 GAN 网络,如果提高字体训练图像分辨率,会使得训练周期和时间复杂度急剧增加,因此构建一个能够提高字体图像分辨率的方法可以一定程度解决这一问题。



a) 低分辨率字体图像 矢量化后效果 b) 高分辨率字体图像 矢量化后效果

图3 不同分辨率字体图像矢量化后效果图

1 相关工作

图像超分辨率(image super-resolution, 简称 SR)是指由一幅低分辨率图像或图像序列恢复出高分辨率图像,是一种图像复原技术^[1]。图像超分辨率研究可分为3个主要范畴:基于插值、基于重建和基于学习的方法^[2]。基于插值的方法,如双三次插值法,是利用邻近像素的值来逼近损失的图像信息,从而实现图像分辨率的放大^[3]。虽然该方法能在一定程度上增加图像分辨率,但是在中文字体图像中容易丢失边缘细节,使得中文字体图像边缘模糊。基于重建的方法是重建退化图像丢失的高频信息^[4],虽然该方法计算量小,但是无法处理结构复杂的中文字体图像。由于上述两个方法存在自身局限性,一般把这两个方法用于基于学习方法的辅助处理过程。

随着机器学习的发展,基于学习的方法被广泛应用到图像超分辨率中。如 Yang 等人提出的传统基于稀疏编码算法是基于学习的图像超分辨率代表方法之一^[5]。此方法是从图像中密集地提取重叠的斑块,进行预处理,然后低分辨率字典再对这些补丁进行编码。但是传统的机器学习算法生成的字体往往由于网络模型参数等的原因,使得图像分辨率较低,从而使字体图像矢量化效果较差。为了进一步提高字体图像的分辨率,基于超分辨率卷积神经网络(SRCNN)被提出。该方法是低分辨率和高分辨率图像之间的端到端映射^[6],整个高分辨率图像(high resolution, HR)的流水线完全是通过学习获得的,只需要进行很少的预处理。同时它结构简单,可以提供更高的精确度。

为了解决上述中文字体分辨低问题,本文以基于学习的方法为研究对象,提出一种针对生成字体图像的超分辨率卷积神经网络(SCRNN)的方法来提高字体图像的分辨率,进而优化矢量化

效果。同时本实验辅助字体设计师在字体生成时解决了周期长、设计复杂等流程。

2 数据集

中文字体数据集是相关研究中较为规范和数据较容易获得的,本文采用设计师设计的2477个中文字体数据集如图4。并通过Fontforge导出分辨率分别为100像素与256像素的不同大小的字体图像。

数据预处理在深度学习中起着重要作用,在

网络模型的学习训练中,反向传播时传递到输入层的梯度会增大。由于梯度增大,直接影响网络学习效率。在该情况下,学习率的选择需要参考输入层数值大小,而直接将字体图像数据归一化操作,能很方便的选择学习率。故本文在获取到字体数据集后,对字体数据进行归一化处理。由于字体图像的像素值只有0与255,将这些像素值除以255,使他们缩放为0与1。进而将每一个字体图像减去数据的统计平均值(逐样本计算),这种归一化可以移除字体图像的平均亮度值。



图4 设计师设计的中文字体数据集

对于网络模型输入的字体图像数据集分为两个部分,一部分为图像尺寸均为 100×100 像素的字体图像,另一部分为图像尺寸均为 256×256 像素的字体图像。超分辨率网络学习低像素到高像素的映射,以此将低分辨率字体图像生成成为高分辨率字体图像。

3 超分辨率卷积神经网络

本文提出的超分辨率卷积神经网络(SCRNN)一种针对单幅字体图像超分辨率的深度学习。该方法首先使用双三次插值将单个低分辨率字体图像放大到所需的大小,通过卷积

神经网络(CNN)拟合非线性映射输出为较高的分辨率的目标图片。深度学习在经典的超分辨率计算机视觉问题中的应用,能达到很好的质量和速度。

网络由三个操作组成:面片提取和表示、非线性映射、重建。面片提取和表示是从低分辨率图像中提取面片,并将每个面片表示为高维向量;非线性映射是将每个高维向量非线性映射到另一个高维向量上,每个映射的矢量在概念上都是高分辨率面片的表示,这些矢量构成了另一组特征图;重建是聚集上述高分辨率拼接表示,以生成最终的高分辨率图像。SCRNN网络的结构如图5。

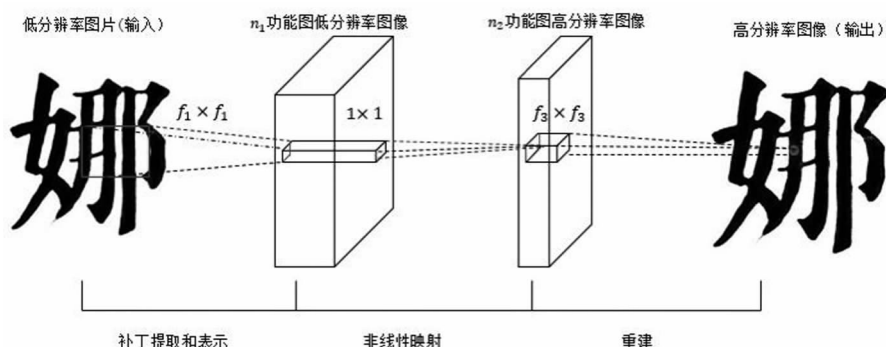


图5 SCRNN网络的结构示意图

首先,给定一张低分辨率字体图像,SCRNN 的第一卷积层提取一组特征映射。第二层将这些要素地图非线性映射到高分辨率面片表示。最后一层结合空间邻域内的预测,产生最终的高分辨率图像。

3.1 SCRNN 的原理

(1) 面片提取和表示。图像复原中一种流行的策略是密集的提取面片,然后通过用一组预先训练好的基表示,如 PCA、DCT、Haar 等。这相当于通过一组滤波器对图像进行卷积,每个滤波器都是一个基础。将这些基础的优化纳入到网络的优化中,形式上,把第一层表示为一个操作:

$$F_1(Y) = \max(0, W_1 * Y + B_1) \quad (1)$$

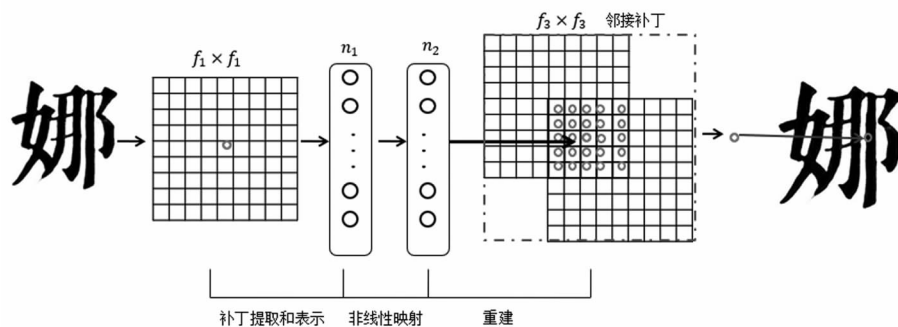


图6 从卷积神经网络的角度说明基于稀疏编码的方法

(3) 重建。在传统方法中,通常将预测的高分辨率重叠斑块平均生成最终的完整图像^[8]。平均可以被认为是—组特征图上的一个预定义的过滤器(其中每个位置是一个高分辨率补丁的“扁平”向量形式)。定义了一个卷积层来产生最终的高分辨率图像为

$$F(Y) = W_3 * F_2(Y) + B_3 \quad (3)$$

式中, W_3 的大小为 $n_2 \times f_3 \times f_3 \times c$; B_3 为 c 维向量。如果表示高分辨率的补丁位于字体图像域,则过滤器的作用类似于平均滤波器;如果位于某些其他域,则 W_3 的行为是将系数投影到图像域,然后进行平均。无论哪种方式, W_3 都是一组线性过滤器。把这三种操作放在一起,就形成了卷积神经网络如图4。该模型中所有的滤波权重和偏差都是最优化的。

3.2 损失函数

损失函数(Loss function)是将随机事件或其有关随机变量的取值映射为非负实数以表示该随机事件的“损失”的函数。学习端到端映射函数 F 需要估计参数 $\theta = \{W_1, W_2, W_3, B_1, B_2, B_3\}$ 。这是通过最小化重建图像 $F(Y; \theta)$ 与对应的地面真实高分辨率图像 X 之间的损失来实现的。给定—组

式中, W_1 和 B_1 分别表示滤波器和偏置。 W_1 的大小为 $c \times f_1 \times f_1 \times n_1$; c 是输入图像中的通道数; f_1 是滤波器的空间尺寸; n_1 是滤波器的数量。 W_1 对图像进行 n_1 卷积,每个卷积核大小为 $c \times f_1 \times f_1$ 。输出由 n_1 个特征地图组成。 B_1 是一个 n_1 维向量,它的每个元素都与一个过滤器相关联^[7]。

(2) 非线性映射。把第一层提取到每个面片的 n_1 维向量映射到 n_2 维向量。相当于应用具有 1×1 平凡空间支撑的 n_2 滤波器。第二层操作为

$$F_2(Y) = \max(0, W_2 * F_1(Y) + B_2) \quad (2)$$

式中, W_2 的大小为 $n_1 \times 1 \times 1 \times n_2$, B_2 是 n_2 维向量。每个输出的 n_2 维向量在概念上是将于重建的高分辨率面片的表示,如图6。

高分辨率图像 $\{X_i\}$ 及其对应的低分辨率图像 $\{Y_i\}$, 本文使用均方误差(MSE)作为损失函数:

$$L(\theta) = \frac{1}{n} \sum_{i=0}^n \|F(Y_i; \theta) - X_i\|^2 \quad (4)$$

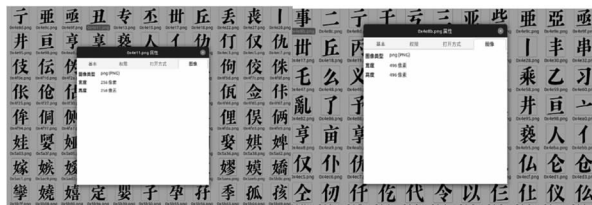
式中, n 是训练样本的数量。损失是最小的使用随机梯度下降与标准反向传播^[9]。

使用均方误差作为损失函数有利于高 PSNR。PSNR 是用于定量评估图像恢复质量的广泛使用的度量,并且至少部分地与感知质量相关^[10]。如果损失函数是可导的,卷积神经网络不排除使用其它类型的损失函数。如果在训练过程中给出了一个更好的感知激励度量,那么网络可以灵活地适应该度量。

4 实验

实验中,本文在字体数据集上迭代训练 100 个周期,每一个周期迭代训练 2 400 次,每一次交替训练判别器和生成器各一次,采用深卷积网络进行训练。将每一个 batch 的大小设置为 16,参数 $\lambda_1 = 10$,参数 λ_2 设置为 100。训练中,网络的学习速率设置为 0.000 1,采用 Adam 优化算法,其中参数 $\beta_{t1} = 0.5$, β_{t2} 设置为 0.9。每迭代 10 个周期保存一次训练模型。

在网络模型生成过程中,输入的字体图像像素为 256×256 ,经过超分辨率重构之后的字体图像为 496×496 像素,生成的字体图像细节更加明显、清晰。深卷积网络的结果看出其输入输出的字体图像在整体的内容表达上与原图是一致的,生成出的字体图像内容质量更高。输入输出端的字体图像如图 7。



a) 输入端字体图像

b) 输出端字体图像

图 7 网络模型生成的字体图像

在生成过程中,每一个生成字体图像比输入的图像更加清晰,结构更加完整。生成字体图像如图 6。图 8a 为低分辨率的字体图像,8b 为高分辨率字体图像字体图像。

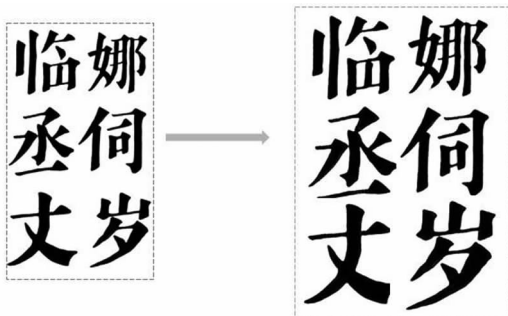
a) 输入端字体图像 256×256 像素b) 输出端字体图像 496×496 像素

图 8 单个生成的字体图像

为了测试文本提出的模型对设计师字库辅助设计的有效性,通过将生成的字体图像导入矢量字库,进而验证文本方法的价值所在,结果如图 9。图中从左至右分别为 256×256 像素以及生成的 496×496 像素的矢量图。通过本文方法生成的字体矢量字库锚点更少,字体结构更为清晰,进而达到字体设计师设置字体的要求,也证明了本文方法的有效性。

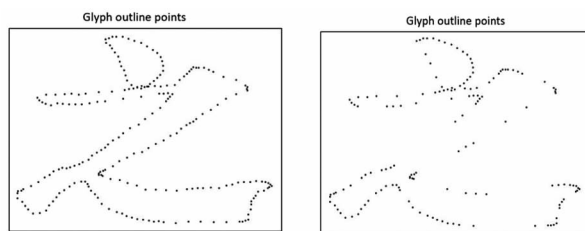
a) 256×256 像素字体矢量化效果b) 496×496 像素字体矢量化效果

图 9 生成矢量字库效果图

5 总 结

提出了一种新的关于字体的图像超分辨率深度学习。以往传统的基于稀疏编码的图像超分辨率方法可以转化为一种深度卷积的神经网络。本文所提出的方法 SRCNN 学习低分辨率和高分辨率字体图像之间的端到端映射,生成符合设计师要求的高分辨率字体图像。通过实验表明本文的方法更具有鲁棒性,对于字体设计行业,大大缩短了设计周期,可辅助字体设计师进行字体导入字库时的优化。

参考文献:

- [1] JIANG J, HU R, WANG Z. Face Super-Resolution via Multilayer Locality-Constrained Iterative Neighbor Embedding and Intermediate Dictionary Learning[J]. IEEE Transactions on Image Processing, 2014, 23(10):4220-31.
- [2] JIANG J, HU R, WANG Z. Noise Robust Face Hallucination via Locality-Constrained Representation[J]. IEEE Transactions on Multimedia, 2014, 16(5):1268-1281.
- [3] 谢海平, 谢凯利, 杨海涛. 图像超分辨率方法研究进展[J]. 计算机工程与应用, 2020, 56(19):34-41.
- [4] 彭亚丽, 张鲁, 张钰. 基于深度反卷积神经网络的图像超分辨率算法[J]. 软件学报, 2018, 29(4):926-934.
- [5] YANG J, WANG Z, LIN Z, HUANG T. Coupled dictionary training for image super-resolution[J]. IEEE Trans. on Image Processing, 2012, 21(8):3467-3478.
- [6] 王春霞, 苏红旗, 范郭亮. 图像超分辨率重建技术综述[J]. 计算机技术与发展, 2011, 21(5):124-127.
- [7] 樊国华. 基于卷积神经网络的小样本图像分辨率改进方法[J]. 电子技术与软件工程, 2020(14):165-167.
- [8] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [9] 徐磊, 张虹, 闫善文, 等. 图像处理过程中影响 PSNR 的因素分析[J]. 通化师范学院学报, 2021, 42(4):44-48.
- [10] DONG C, LOY C C, HE K. Learning a deep convolutional network for image super-resolution[C]//European conference on computer vision. Cham: Springer, 2014:184-199.

(责任编辑 王楠楠)