

整体车牌图像超分辨率重建研究

倪申龙,曾接贤,周世健
(南昌航空大学,江西 南昌 330063)

摘要:为了增强低分辨率车牌(LP)图像的整体重建效果,将深度学习思想应用到车牌图像超分辨率重建任务中,提出了一种基于卷积神经网络(CNN)的单幅整体车牌图像超分辨率重建(SR)算法(LPSRCNN)。对初始高分辨率车牌图像进行预处理,利用双三次插值法 Bicubic 对原始车牌图像进行下采样后再上采样得到和初始图像尺寸大小一样的低分辨率车牌图像作为网络的输入图像,通过训练好的卷积神经网络直接学习低分辨率车牌图像和高分辨率车牌图像之间的映射关系,利用该映射关系输出重建后的高分辨率整体车牌图像。将得到的研究结果与双三次插值法和字典学习方法进行比较,结果表明计算得到的峰值信噪比(PSNR)都高于这两种方法。将该方法应用到整体车牌图像超分辨率重建问题中,可以获得更丰富的细节信息,得到更好的视觉效果,达到了整体车牌图像超分辨率的增强任务。

关键词:图像超分辨率;车牌图像;深度学习;卷积神经网络

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2019)04-0195-05

doi:10.3969/j.issn.1673-629X.2019.04.039

Study on Reconstruction of Overall Image of License Plate

NI Shen-long, ZENG Jie-xian, ZHOU Shi-jian
(Nanchang Aeronautical University, Nanchang 330063, China)

Abstract: In order to enhance the overall reconstruction effect of low-resolution license plate (LP) images, applying the idea of deep learning into the super-resolution reconstruction of license plate images, we propose a super-resolution algorithm based on convolutional neural network (CNN) for single license plate image, referred to as LPSRCNN. The original high-resolution license plate image is pre-processed, and the low-resolution license plate image with the same size as the original image is obtained after the Bicubic interpolation Bicubic sampling of the original license plate image. The trained convolutional neural network is used to directly learn the mapping relationship between low-resolution license plate image and high-resolution license plate image, and the reconstructed high-resolution overall license plate image is output based on the mapping relationship. Compared with the Bicubic interpolation method and the dictionary learning method, the results show that the PSNR is higher than that of the two methods. The method of deep learning is applied to the problem of super-resolution reconstruction of overall license plate image, which can obtain more details and better visual effect, achieving the task of enhancing the super resolution of the overall license plate image.

Key words: image super resolution; license plate image; deep learning; convolution neural network

0 引言

随着车辆保有量的增加,在给人们出行带来方便的同时也极大增加了交警执勤的任务。因此,车牌识别任务已经成为交管部门对车辆管理的重要工作之一,而准确、有效的识别车牌将为交管部门的交通执法和破案带来了极大的便利。准确、有效的车牌识别的前提是所获取到的车牌图像具有很高的分辨率。然而现实生活中,由于摄像设备、拍摄角度、光线、天气和车

辆运动等因素,往往得到的车牌图像是模糊的,大大降低了车牌的识别率。因此,针对车牌图像分辨率低的问题,文中通过超分辨率(super resolution, SR)重建^[1-2]的方法来增强整体车牌图像的空间分辨率。

图像超分辨率重建技术是一种采用信号处理和图像处理的方法,通过软件算法的方式从一幅或多幅低分辨率(low-resolution, LR)图像中重建出一幅高分辨率(high-resolution, HR)图像的图像处理技术。随着

收稿日期:2018-05-23

修回日期:2018-09-20

网络出版时间:2018-12-20

基金项目:国家自然科学基金(61763033)

作者简介:倪申龙(1992-),男,硕士研究生,CCF会员(93080G),研究方向为图像处理与模式识别、计算机视觉;曾接贤,硕士,教授,研究方向为图像处理与模式识别;周世健,博士,教授,研究方向为大数据分析。

网络出版地址: <http://cnki.net/kcms/detail/61.1450.TP.20181220.1049.058.html>

图像超分辨率重建理论和技术的日益成熟,超分辨率重建技术已经广泛应用于卫星遥感、军事侦察、医学影像、视频监控和数字电视等领域。

车牌识别^[3-6]是图像处理领域中一个经典并具挑战性的难题,而车牌图像的分辨率高低最能影响车牌识别的成功率。

传统的车牌识别方法采用分割字符的方式,即将车牌图像上的字符单个分割出来进行单独识别,这样不仅增加了识别难度,而且增加了车牌识别的时间。近年来,基于深度学习^[7-10]的卷积神经网络算法在行人检测^[11]、目标跟踪^[12-13]、人脸识别^[14]等图像处理领域应用广泛,取得了非常好的效果。Dong 等^[10]提出了 SRCNN (super resolution convolutional neural network),将深度学习应用在图像超分辨率重建领域,取得了显著效果。

鉴于深度学习在图像处理领域取得的显著效果,文中重点研究利用 SRCNN 网络来重建高分辨率整体车牌图像,提出一种基于卷积神经网络的整体车牌图像超分辨率重建算法(LPSRCNN)。该算法不需要通过单独分割车牌字符单独识别,而是直接学习网络映射重建出高分辨率车牌图像,从而达到识别整体车牌图像的效果。

1 算法描述

基于深度学习的卷积神经网络直接学习低分辨率车牌图像和对应的高分辨率车牌图像之间的映射关系 F 。这里所说的映射就表示为一个深层的卷积神经网络

络,该映射将低分辨率车牌图像 Y 作为网络输入,然后输出对应的高分辨率车牌图像 $F(Y)$ 。假设原始高分辨率车牌图像表示为 X ,目的是希望输出的高分辨率车牌图像 $F(Y)$ 尽可能地接近原始高分辨率车牌图像 X 。

在建立模型之前,首先要做的是获取低分辨率车牌图像作为模型输入图像 Y ,这是该模型中唯一的一个预处理过程。对原始的高分辨率车牌图像先利用双三次插值法 Bicubic 进行下采样,然后再利用相同的缩放因子上采样到跟原始图像一样的尺寸,得到的低分辨率图像即为网络模型的输入。针对输入图像是 RGB 图像的情况,先将图像从 RGB 颜色空间转换到 YCbCr 颜色空间,并且只对其中的 Y 通道(亮度通道)进行重建,因为人眼对亮度比对颜色更敏感。其余的 Cb、Cr 通道均采用双三次插值法 Bicubic 缩放放到目标尺寸大小。

1.1 网络模型

提出的 LPSRCNN 模型由三层卷积层构成:特征提取层、非线性映射层和重建层。网络模型如图 1 所示。

对网络的前两层,文中选用修正线性单元(rectified linear unit, ReLU)^[15-16]作为网络模型的激活函数,因为 ReLU 形式简单,计算简便,运算量小,能加快模型的收敛速度,并且可以有效避免在训练网络过程中产生的“梯度爆炸/消失”的问题^[17]。ReLU 模型如图 2 所示。

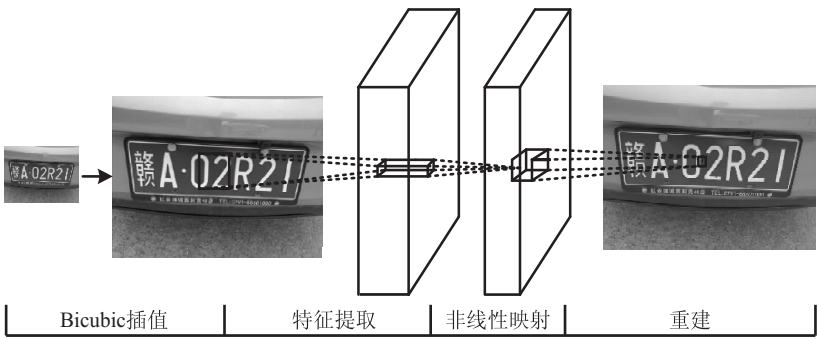


图 1 LPSRCNN 模型示意

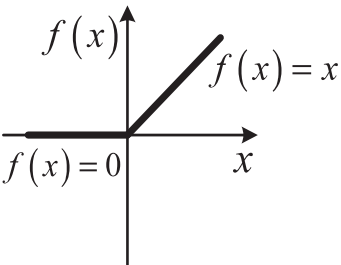


图 2 ReLU 模型

$$f(x) = \max(0, x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{1}$$

在卷积层中,前一层输出的特征图和一个卷积核进行卷积再经过激活函数后形成这一层的特征图。每一幅输出的特征图与不同的卷积核进行卷积后再经过激活函数,又形成下一层的特征图。

1.1.1 特征提取层

为了获得车牌图像的边缘和纹理特征,在输入的低分辨率车牌图像上密集地提取特征小块,并将每个

提取到的特征小块表示为一个个高维向量,然后将这些高维向量组成一组特征图,该特征图的数量等于向量的维度。该过程表示为 $F_1(Y)$:

输入:

$$Z_1 = W_1 * Y + B_1 \quad (2)$$

输出:

$$A_1 = f(Z_1) \quad (3)$$

$$F_1(Y) = A_1 = \max(0, W_1 * Y + B_1) \quad (4)$$

其中, W_1 为该卷积层的滤波器组,且每一个滤波器的大小是 $f_1 \times f_1$, 数量是 n_1 , 所以 W_1 的大小就是 $n_1 \times f_1 \times f_1 \times c$, c 表示输入图像的通道数(因为文中只对 Y 通道进行处理,所以 $c=1$); B_1 表示该卷积层的偏差(是 n_1 维的);“ $*$ ”表示卷积运算。每一个滤波器提取一种特征,所以经过卷积层卷积后输出的结果是由 n_1 个特征组成的特征图。

1.1.2 非线性映射层

如图1所示,为了获取高维特征向量,将第一层提取出的低维特征图非线性地映射到另一个高维矢量上,即在该层将第一层提取到的 n_1 维特征向量映射到 n_2 维上。这一过程表示为 $F_2(Y)$:

输入:

$$Z_2 = W_2 * F_1(Y) + B_2 \quad (5)$$

输出:

$$A_2 = f(Z_2) \quad (6)$$

$$F_2(Y) = A_2 = \max(0, W_2 * F_1(Y) + B_2) \quad (7)$$

其中, W_2 表示该层的滤波器组,每一个滤波器大小为 $n_1 \times f_2 \times f_2$, 数量为 n_2 , 所以 W_2 的大小为 $n_2 \times n_1 \times f_2 \times f_2$; B_2 表示 n_2 维的偏差。非线性映射层输出 n_2 维的特征图,且这些特征图都是高分辨率图像小块,这些生成的高分辨率图像特征小块将用于第三层的重建。

1.1.3 重建层

将第二层获得的特征图即高分辨率图像特征小块进行平均融合得到最终的高分辨率重建图像。该过程表示为 $F(Y)$:

$$F(Y) = W_3 * F_2(Y) + B_3 \quad (8)$$

其中, W_3 表示为 c 个大小为 $n_2 \times f_3 \times f_3$ 的滤波器组,网络的输出层没有采用激活函数,并且 W_3 可以看作是一个均值滤波器,整个重建过程是一个线性操作过程; B_3 表示一个 c 维的偏差。

1.2 损失函数

期望获得的重建车牌图像 $F(Y)$ 和原始高分辨率车牌图像 X 越接近越好,即它们之间的差值(损失函数)越小越好。由以上分析可知,学习到的端到端的映射函数 F 需要学习各层中的网络参数 W 和 B , 将各层网络参数组成一组参数向量 $\Theta = \{W_1, W_2, W_3, B_1, B_2, B_3\}$, 整个网络的训练就是对这组参数向量的估计

和优化,通过最小化 $F(Y_i; \Theta)$ 和 X_i 之间的差值得到参数最优解。

文中通过均方误差(MSE)来计算网络的损失函数:

$$L(\Theta) = \text{MSE} = \frac{1}{n} \sum_{i=1}^n \|F(Y_i; \Theta) - X_i\|^2 \quad (9)$$

其中, n 为训练样本数量; Y_i 为输入的低分辨率图像块; X_i 为对应的原始高分辨率图像块。

1.3 反向传播算法

为了使损失函数最小化,文中采用反向传播算法进行训练,主要利用随机梯度下降法去最优化每一层的网络参数。权重 W_i 和偏差 B_i 的更新公式如下:

$$W_i = W_i - \alpha \frac{\partial L(\Theta)}{\partial W_i} \quad (10)$$

$$B_i = B_i - \alpha \frac{\partial L(\Theta)}{\partial B_i} \quad (11)$$

其中, W_i 表示第 i 层的权重; B_i 表示第 i 层的偏差; α 表示学习率。

然后通过计算参数的偏导数来更新权重参数,计算公式如下:

$$\Delta W^{(l)} = \Delta W^{(l)} + \nabla_{W^{(l)}} L(\Theta) \quad (12)$$

$$\Delta B^{(l)} = \Delta B^{(l)} + \nabla_{B^{(l)}} L(\Theta) \quad (13)$$

得到更新后的权重参数之后,重复梯度下降法使目标损失函数达到最小,直至收敛,使得最终重建出来的图像更接近于原始的高分辨率图像。

2 实验与分析

2.1 图像质量评价指标

在超分辨率算法评价标准中,有主观评价和客观评价两类标准。其中主观标准主要是通过人眼判定,查看重建的图像是否有充分的细节信息、边缘是否模糊,或者重建的图像是否有严重的其他缺陷、图像的整体效果是否符合审美标准等。客观评价是计算重建后图像和原始图像之间的相似程度。将提出的深度卷积神经网络算法应用到车牌图像超分辨率重建中,旨在提高车牌图像的视觉效果,为了比较算法的公平性,将其与传统的双三次插值 Bicubic 算法以及基于字典学习的 K-SVD 算法^[18-19]进行比较。采用峰值信噪比(PSNR)作为车牌图像重建的质量评价指标。PSNR 表示信号的最大功率和噪声功率的比值,单位为 dB,是目前广泛用于评价图像质量的评价指标之一,公式如下:

$$\text{PSNR} = 10 \times \log(255^2 / \text{MSE}) \quad (14)$$

其中, MSE 是重建后图像和原始图像之间的均方误差。峰值信噪比 PSNR 值越大,重建效果越好。

2.2 实验设置

为了初始化不同的参数,选用均值为 0、标准差为

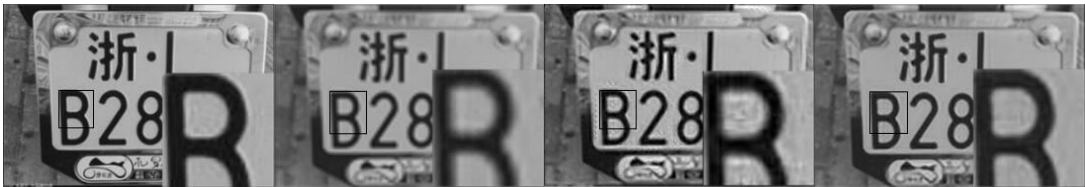
10^3 的高斯分布来初始化每一层的滤波器权重。然后使用随机梯度下降法求得目标损失函数的最优解。经过经验性的参数调整,最终设置参数如下: $f_1 = 9, f_2 = 5, f_3 = 5, n_1 = 64, n_2 = 32$,并设置放大因子为 3 倍,获得的图像视觉效果最佳。对于输入的彩色 RGB 图像,文中首先将图像从 RGB 颜色空间转换到 YCbCr 颜色空间,并且只对亮度通道 Y 进行超分辨率重建,剩下的两个 Cb、Cr 通道均采用双三次插值法 Bicubic 缩放到目标尺寸大小。



(a) 图像 1



(b) 图像 2



(c) 图像 3

图 3 放大因子为 3 的重建效果对比

从人眼视觉的美观角度分析,文中算法相比于上述两种方法能够获得更加清晰的重建效果。与 Bicubic 算法相比,文中算法得到的图像边缘更加清晰、平滑,整体图像视觉感较好,符合人眼的美学要求。而与 K-SVD 算法相比,文中算法能够有效抑制 K-SVD 算法产生的振铃现象。值得一提的是,虽然车牌图像经过文中算法得到的整体重建效果较好,但重建后的车牌图像中的中文文字部分不清晰,细节信息难以恢复,与原始真实的高清车牌图像中的文字存在一定的差距。

从客观角度分析,表 1 是不同算法对所选车牌图像重建后的 PSNR 值比较,PSNR 值越大,表示重建的效果越好。由表 1 可知,文中算法得到的 PSNR 值高于其他两种算法,说明文中算法重建质量有效。双三次插值法和 K-SVD 算法虽然提高了车牌图像的质量,但重建后的图像仍然存在边缘模糊的缺点。与传统的双三次插值法和基于字典学习的 K-SVD 算法相比,LPSRCNN 模型重建后的车牌图像边缘、纹理更清晰,并且得到的 PSNR 值均高于双三次插值和 K-SVD

2.3 实验结果与分析

为了验证 LPSRCNN 算法的有效性,选取了部分车牌图像进行实验。在放大因子为 3 的情况下,将文中算法分别与传统的双三次插值 Bicubic 算法和基于字典学习的 K-SVD 算法进行重建效果对比,如图 3 所示。其中,图像 1~3 中从左至右分别是原始高分辨率图像以及算法 Bicubic、K-SVD、LPSRCNN 重建后得到的图像。

算法得到的 PSNR 值,整体效果更接近原始的高分辨率图像。表明了卷积神经网络能够有效地直接学习高分辨率车牌图像和低分辨率车牌图像之间端到端的映射关系,重建后的车牌图像视觉效果更加明显。

表 1 不同算法对车牌图像重建后的 PSNR 值 dB

图像	倍数	Bicubic	K-SVD	LPSRCNN
图像 1	3	26.89	27.46	29.01
图像 2	3	24.32	25.11	27.33
图像 3	3	25.81	26.35	27.98

3 结束语

针对整体车牌图像超分辨率重建这一问题,将卷积神经网络的方法应用到整体车牌图像超分辨率重建中,直接学习高分辨率车牌图像和低分辨率车牌图像之间端到端的映射关系。实验结果表明,LPSRCNN 算法处理后的图像边缘、纹理等细节信息更加丰富,视觉效果得到了提升,有效提高了整体车牌图像的质量,达到了预期实验效果。但该算法中输入的不是原始的

低分辨率车牌图像,而是将原始的高分辨率车牌图像进行 Bicubic 插值得到的低分辨率车牌图像,这增加了算法的计算复杂度,并且得到的图像有人工操作的痕迹。因此,下一步将重点研究当输入图像是原始的低分辨率车牌图像时,能否改进算法的优越性,以及网络深度、滤波器大小、滤波器数量等因素对车牌图像重建效果的影响。

参考文献:

[1] 苏衡,周杰,张志浩. 超分辨率图像重建方法综述[J]. 自动化学报,2013,39(8):1202-1213.

[2] 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.

[3] 李志强,李永斌. 车牌识别技术的发展及研究现状[J]. 科技信息,2012(5):110.

[4] 李波,曾致远,周建中,等. 车牌识别系统研究与实现[J]. 计算机技术与发展,2006,16(6):10-11.

[5] 汤玉垚. 复杂背景下车牌识别算法的研究[D]. 合肥:中国科学技术大学,2016.

[6] 顾弘,赵光宙,齐东莲,等. 车牌识别中先验知识的嵌入及字符分割方法[J]. 中国图象图形学报,2010,15(5):749-756.

[7] LIANG Yudong, WANG Jinjun, ZHOU Sanping, et al. Incorporating image priors with deep convolutional neural networks for image super-resolution [J]. Neurocomputing, 2016, 194:340-347.

[8] 李素梅,雷国庆,范如. 基于卷积神经网络的深度图超分辨率重建[J]. 光学学报,2017,37(12):1210002.

[9] 黄凤,王晓明. 增强的单幅图像自学习超分辨方法[J]. 计算机应用,2017,37(9):2636-2642.

[10] DONG Chao, LOY C C, HE Kaiming, et al. Learning a deep convolutional network for image super-resolution [C]//European conference on computer vision. Berlin: Springer,

2014:184-199.

[11] OUYANG Wanli, WANG Xiaogang. Joint deep learning for pedestrian detection [C]//IEEE international conference on computer vision. Sydney, NSW, Australia: IEEE, 2014:2056-2063.

[12] OUYANG Wanli, ZENG Xingyu, WANG Xiaogang, et al. DeepID-net: deformable deep convolutional neural networks for object detection [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(7):1320-1334.

[13] ZHANG Ning, DONAHUE J, GIRSHICK R, et al. Part-based R-CNNs for fine-grained category detection [M]//13th European conference on computer vision. [s. l.]: Springer International Publishing, 2014:834-849.

[14] SUN Yi, CHEN Yuheng, WANG Xiaogang, et al. Deep learning face representation by joint identification-verification [C]//Proceedings of the 27th international conference on neural information processing systems. Montreal, Canada: MIT Press, 2014:1988-1996.

[15] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]//Proceedings of the 25th international conference on neural information processing systems. Lake Tahoe, Nevada: Curran Associates Inc., 2012:1097-1105.

[16] NAIR V, HINTON G E. Rectified linear units improve restricted Boltzmann machines [C]//International conference on machine learning. Omnipress: [s. n.], 2010:807-814.

[17] HOCHREITER S, BENGIO Y, FRASCONI P, et al. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies [M]. [s. l.]: IEEE Press, 2001.

[18] 刘志州. 基于字典学习的超分辨率图像重构 [D]. 西安:西安电子科技大学, 2011.

[19] 史国川, 龚连友. 基于系数复用和字典训练的图像超分辨率算法 [J]. 计算机技术与发展, 2018, 28(3):114-117.