



面向人群计数的对偶卷积神经网络

郭华平, 王锐, 王敬, 孙艳歌, 李健, 李萌

引用本文:

郭华平, 王锐, 王敬, 孙艳歌, 李健, 李萌. 面向人群计数的对偶卷积神经网络[J]. 信阳师范学院学报自然科学版, 2021, 34(4): 650–654. doi: 10.3969/j.issn.1003–0972.2021.04.024

GUO Huaping, WANG Rui, WANG Jing, SUN Yange, LI Jian, LI Meng. Dual Convolutional Neural Networks for Crowd Counting[J]. Journal of Xinyang Normal University (Natural Science Edition), 2021, 34(4): 650–654. doi: 10.3969/j.issn.1003–0972.2021.04.024

在线阅读 View online: <https://doi.org/10.3969/j.issn.1003–0972.2021.04.024>

您可能感兴趣的其他文章

Articles you may be interested in

基于卷积神经网络进行电影院人群分布统计

The Distribution of Cinema Population Based on Convolutional Neural Network

信阳师范学院学报自然科学版, 2020, 33(4): 675–680. <https://doi.org/10.3969/j.issn.1003–0972.2020.04.028>

基于递归卷积神经网络的行人检测方法

Pedestrian Detection Method Based on Recurrent Convolutional Neural Networks

信阳师范学院学报自然科学版, 2021, 34(4): 655–660. <https://doi.org/10.3969/j.issn.1003–0972.2021.04.025>

一种加权词向量的混合网络文本情感分析方法

A Mixed Network Text Sentiment Analysis Method Based on Weighted Word Vectors

信阳师范学院学报自然科学版, 2021, 34(3): 472–477. <https://doi.org/10.3969/j.issn.1003–0972.2021.03.022>

基于卷积神经网络的英文篇章情感量化方法

Sentiment Quantization of English Texts Based on Convolutional Neural Networks

信阳师范学院学报自然科学版, 2021, 34(1): 130–137. <https://doi.org/10.3969/j.issn.1003–0972.2021.01.022>

基于随机森林算法的混凝土早期抗裂性预测研究

Prediction of Early Crack Resistance of Concrete Based on Random Forest Algorithm

信阳师范学院学报自然科学版, 2021, 34(1): 158–165. <https://doi.org/10.3969/j.issn.1003–0972.2021.01.026>

面向人群计数的对偶卷积神经网络

郭华平*, 王 锐, 王 敬, 孙艳歌, 李 健, 李 萌

(信阳师范学院 计算机与信息技术学院, 河南 信阳 464000)

摘 要:基于卷积神经网络(CNN)人群计数方法的核心是如何构建能将人群图有效映射为密度图的 CNN 模型,进而从密度图准确地估计出人群数量.传统构建 CNN 的方法只考虑了人群图到密度图的映射,并未考虑密度图到人群图的映射,以及该映射对模型性能的影响.为了解决以上问题,提出一种基于卷积神经网络的对偶模型(Dual Convolutional Neural Networks, DualCNN)以提高模型将人群图映射为密度图准确性. DualCNN 包含有两个映射子模型:1)将人群图映射到密度图的卷积神经模型,2)将密度图映射到人群图的对偶卷积神经模型.在学习过程中,通过两个子模型的相互影响,进而达到提高卷积神经模型在人群计数问题上的性能.在 UCF_CC_50 数据集和 ShangHaitech 数据集上的实验结果表明,该方法能有效提升 CNN 的计数性能,尤其在 UCF_CC_50 数据集中, DualCNN 将 MCNN 和 CSRNet 的平均绝对误差(MAE)分别降低 15.6% 和 15.8%, 最小均方误差(MSE)分别降低 18.1% 和 28.8%.

关键词:人群计数;人群图;密度图;对偶卷积神经网络

中图分类号:TP391

文献标识码:A

开放科学(资源服务)标识码(OSID):



Dual Convolutional Neural Networks for Crowd Counting

GUO Huaping*, WANG Rui, WANG Jing, SUN Yange, LI Jian, LI Meng

(College of Computer & Information Technology, Xinyang Normal University, Xinyang 464000, China)

Abstract: The core of crowd counting methods based on convolutional neural networks (CNN) is to construct a CNN model that can effectively mapping the crowd image to the density map, and then accurately estimate the number of crowd from the density map. The conventional CNN based methods only consider the mapping from crowd images to density maps without taking that from density maps to crowd maps and the impact of the mapping on model performance. A dual convolutional neural networks (DualCNN) is proposed to improve the accuracy of the mapping from crowd maps to density maps. DualCNN contains two mapping sub-models: 1) a convolutional neural model that mapping a crowd image to a density map; 2) a dual convolutional neural model that mapping a density map to a crowd image. In the learning process, through the mutual influence of the two sub-models, the performance of the convolutional neural model on crowd counting is improved. Experimental results on the datasets of UCF_CC_50 and ShangHaitech show that the presented method can effectively improve the counting performance of CNN. Especially in the UCF_CC_50 data set, DualCNN reduces the MAE of MCNN and CSRNet by 15.6% and 15.8%, and MSE by 18.1% and 28.8%, respectively.

Key words: crowd count; crowd image; density map; dual convolutional neural network

0 引言

城市化的快速发展使得人群越来越聚集,这也致使人群踩踏事件时有发生.例如,2015 年中国上海由大规模的人群集会导致的踩踏事件,造成 35

人死亡;不幸的是自此以后,世界各地的大规模踩踏事件层出不穷,造成了更多的人受害^[1].因此,如何从图像或视频中及时准确地估计人群数目以指导人群控制是一件非常有意义的工作.另外,人群数目相同的不同场景,人群分布可能完全不同,在

收稿日期:2020-12-09;修订日期:2021-01-27;*.通信联系人,E-mail:hpguo@xynu.edu.cn
基金项目:国家自然科学基金项目(31900710,62006205);河南省教师教育课程改革项目(2020-JSJYYB-034);信阳师范学院研究生科研创新基金(2020KYJJ15)
作者简介:郭华平(1982—),男,河南固始人,副教授,博士,硕士生导师,主要从事机器学习与数据挖掘研究.

准确获取人群数目的前提下,如何准确地获得人群分布情况对高风险环境中做出正确决策十分有益。

基于卷积神经网络(CNN)的人群计数方法^[2-5]能很好地估计人群数目和人群分布,其核心思想是:首先将人群图直接映射为人群分布密度图,然后根据密度图估计人群数目。如FU等^[2]用CNN模型映射图片的人群密度等级,按照等级估计人数;ZHANG等^[1]提出了多列卷积神经网络模型(MCNN),其使用尺寸不同的滤波器组在每一列上抽取不同尺寸的图像特征,以便每列CNN适应人/人头的尺寸变化。LI等^[3]结合数据驱动和空洞卷积的思想,提出了空洞卷积神经网络模型(CSRNet),生成高质量的人群密度图并估计出人群数目。

以上基于卷积神经网络的人群计数方法只考虑人群图到密度图的映射^[6-11],而忽略了映射后的密度图到人群图的重构(即 C 的对偶模型)及重构对原模型的影响,因此本文试图构建一个新模型,其既能将人群图映射到密度图,又能将密度图映射为人群图,同时考虑这种重构对模型的影响,进而提升模型在人群计数问题上的性能。

根据该思想,本文提出一种面向人群计数的对偶卷积神经网络(DualCNN)模型。该模型启发来自于双语言翻译对偶学习模型(Dual Learning Model)^[12],它通过最小化两个翻译模型的重构损失,同时训练两个相对偶的语言翻译模型。两个相对的翻译模型形成一个闭环,任意一个翻译模型的重构损失会产生信息反馈来训练两个翻译模型。与Dual Learning Model相似,DualCNN也包含两个互为对偶的卷积神经网络子模型,其中任意一个子模型的重构损失可以同时用于训练这两个子模型。

令DualCNN的两个子模型为 C_A 和 C_B ,其中 C_A 实现人群图到密度图的映射, C_B 实现密度图到人群图的映射。如果 C_A 映射得到密度图能够被 C_B 精确地映射(重构)为人群图,且 C_B 映射得到人群图能够被 C_A 精确地反映为密度图,那么就认为 C_A 模型生成的密度图质量较好,能够准确地估计出人群总数和反映出人群分布情况。

与Dual Learning Model的不同点主要表现在如下两个方面:1) Dual Learning Model的目标是实现语言间互译(两种语言一样重要),DualCNN目标是学习具有高精度的人群计数模型,即更关注于 C_A 的性能,这使得 C_A 和 C_B 重构损失更关注于对 C_A 的影响;2) CNN模型主要用来实现人群

图(密度图)到密度图(人群图)的映射,而不是序列翻译模型。相关实验结果表明,本文提出的方法能很好地提升模型在人群计数问题上的性能。

1 提出方法

传统的解决人群计数问题所提出来的卷积神经网络模型(CNN),都是将人群图有效映射为密度图,进而从密度图准确地估计出人群数量。本节提出一种对偶卷积神经网络模型(DualCNN),不仅实现了人群图到密度图的映射,还实现了密度图到人群图的映射。

1.1 对偶卷积神经网络

设 $x \in X$ 表示人群图像集 X 中的一个人群图像和 $y \in Y$ 是真实密度图集中相应于 x 的真实人群密度图,其使用几何自适应核方法^[1]获得。DualCNN的主要任务是学习一个卷积神经网络模型 $C_A: X \rightarrow Y$,实现人群图像 $x \in X$ 到相应密度图 $y \in Y$ 的映射;其对偶任务是学习 C_A 的对偶模型 $C_B: Y \rightarrow X$,实现密度图 $y \in Y$ 到真实人群图 $x \in X$ 的映射,将密度图 $y \in Y$ 映射为人群图像。DualCNN在学习 C_A 和 C_B 过程中考虑彼此间的影响,以提高模型在人群计数问题上的性能。图1给出了对偶卷积神经网络的总体结构图。

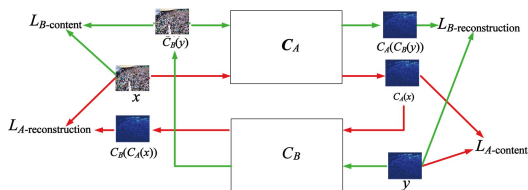


图1 对偶卷积神经网络结构图

Fig. 1 Schematic of dual convolutional neural network

如图1所示,使用 C_A 模型将人群图像 $x \in X$ 映射到密度图 Y ,即 $C_A(x)$ 。 $C_A(x)$ 的映射(内容)损失 $L_{A-content}$ 由与 x 相对应的真密度图 $y \in Y$ 监督,然后使用 C_B 将 $C_A(x)$ 重构回人群图像,即 $C_B(C_A(x))$,其重构损失 $L_{A-reconstruction}$ 由真实人群图像 x 监督。类似地,将真实密度图 y 映射为人群图像 $C_B(y)$ 并重构为 $C_A(C_B(y))$ 。 $C_B(y)$ 的内容损失 $L_{B-content}$ 和 $C_A(C_B(y))$ 的重构损失 $L_{B-reconstruction}$ 分别由真实图像 x 和真实密度图 y 监督。

1.2 学习目标

给定一组人群图和密度图(x, y),对偶卷积神经网络(DualCNN)单独学习每个子模型,即在学习 C_A 时,固定 C_B ,其中 $x \in X$ 为人群图像, $y \in$

Y 为 x 对应的人群密度图(相关符号参见 1.1 节). 类似地, DualCNN 在学习 C_B 时, 固定 C_A . 由图 1, DualCNN 的两个子神经网络模型 C_A 和 C_B 都同时受到内容损失和重构损失的影响, 因此定义用于学习 C_A 的损失函数为:

$$L_A(x, y) = \alpha L_{A-content} + (1 - \alpha) L_{A-reconstruction}, \tag{1}$$

其中 $\alpha \in [0, 1]$ 是用于调节内容损失和重构损失的权重, $L_{A-content}$ 和 $L_{A-reconstruction}$ 分别为 x 的内容损失和重构损失, 定义为:

$$L_{A-content} = \|y - C_A(x)\|, \tag{2}$$

$$L_{A-reconstruction} = \|x - C_B(C_A(x))\|. \tag{3}$$

在式(1)中, 若设置 $\alpha = 1$, 则式(1)退化为传统面向人群计数问题的卷积神经网络(CNN)模型的损失函数, 故可用于预训练 C_A . 在实验部分, 使用该方法预训练 C_A 模型. 类似地, 定义用于学习 C_B 的损失函数为:

$$L_B(x, y) = \alpha L_{B-content} + (1 - \alpha) L_{B-reconstruction}, \tag{4}$$

$L_{B-content}$ 和 $L_{B-reconstruction}$ 分别为 y 的内容损失和重构损失, 分别定义为:

$$L_{B-content} = \|x - C_B(y)\|, \tag{5}$$

$$L_{B-reconstruction} = \|y - C_A(C_B(y))\|. \tag{6}$$

类似于式(1), 若设置式(2)中 $\alpha = 1$, 重构损失 $L_{B-reconstruction}$ 无效, 故可用于预训练 C_B .

1.3 训练过程

使用梯度下降法学习对偶卷积神经网络(DualCNN)中子模型 C_A 和 C_B 的参数. DualCNN 迭代、交替学习 C_A 和 C_B : 每次迭代, 首先固定 C_B , 更新优化 C_A 中的参数; 然后固定 C_A , 更新优化 C_B 中的参数. 算法 1 给出了 DualCNN 具体细节.

算法 1 对偶卷积神经网络学习算法

输入: 人群图片集 X , 与 X 相对应的真实密度图片集 Y , 迭代次数 N , 内容损失和重构损失的权重协调因子 α

开始:

1. 随机初始化 C_A 和 C_B 的参数 θ_A 和 θ_B ;
2. do
3. for 每张人群图 $x \in X$ 及相应密度图 $y \in Y$
- do
4. 固定 C_B , 通过最小化式(1)更新参数 θ_A ;
5. 固定 C_A , 通过最小化式(4)更新参数 θ_B ;
6. end for
7. until 参数收敛
8. return 模型 C_A ;

DualCNN 的输入是人群图像集 X 和相应的

人群密度图集 Y . DualCNN 首先分别初始化 C_A 和 C_B 的参数 θ_A 和 θ_B (第 1 行), DualCNN 迭代学习模型参数直至模型参数收敛(第 2~7 行). 对于每次迭代, DualCNN 逐一以每张人群图 $x \in X$ 及相应密度图 $y \in Y$ 为训练数据更新 C_A 和 C_B 的参数(第 3~6 行). 对于当前人群图和相应密度图组 (x, y) , DualCNN 首先固定 C_B , 以式(1)为损失函数优化 C_A 参数(第 4 行), 然后 DualCNN 固定 C_A , 以式(4)为损失函数优化 C_B 参数(第 5 行).

2 实验

2.1 实验设置

分别选取 MCNN^[1]、Unet^[13] 和 CSRNet^[3] 作为对偶卷积神经网络(DualCNN)中的 C_A 模型, 并对相应的模型分别取名为 DualCNN-MCNN、DualCNN-Unet 和 DualCNN-CSRNet, 其中 C_A 是实现人群图像到相应密度图映射的卷积神经网络. 表 1 给出了这 3 种模型具体设置, 其中 C_B 表示 C_A 的对偶模型. DualCNN-MCNN 中 C_A 模型的每一列都有两次下采样, 映射的密度图尺寸减小为人群图的 1/16, 故 C_B 模型需要两个上采样层(与 C_A 下采样层相对应)将映射的密度图重构为人群图. 根据文献[1], MCNN 只是用了人群图的 1 个颜色通道, 故 DualCNN-MCNN 也只用了 1 个颜色通道. 类似与 DualCNN-MCNN 的采用做法, DualCNN-CSRNet 中的 C_B 模型也对特征图进行了上采样(与 C_A 下采样层对应). 注意: DualCNN-CSRNet 处理彩色图片, 所以设置 C_B 的最后一个卷积层包含 3 个卷积核, 以还原图片通道数. 由于 Unet 模型映射后的图片大小与映射前相同, 因此 DualCNN-Unet 的 C_B 模型不需要考虑图片采样问题, 但是需要考虑颜色通道问题, C_B 模型的最后一层也要有 3 个卷积核.

在实验中, 3 个 DualCNN 模型设置的优化器都是 Adam, 其中 DualCNN-MCNN 模型的迭代次数为 500, 学习率为 1e-4; DualCNN-Unet 的学习率为 1e-6 迭代次数为 200; DualCNN-CSRNet 模型学习率为 1e-5, 迭代次数为 500.

平均绝对误差(MAE)和最小均方误差(MSE)是常用于评估人群计数方法的指标, 其中 MAE 表示预测的准确性, MSE 表示预测的鲁棒性, 具体定义如下:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \text{pre}_i|, \tag{7}$$

$$\text{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \text{pre}_i)^2}, \tag{8}$$

其中: N 为测试图像的个数, y_i 和 pre_i 分别为第 i 组图像的实际计数和估计计数.

表 1 模型的具体设置

Tab.1 The specific setting of the model

模型	C_A	C_B
DualCNN-MCNN	包含 3 列子卷积神经网络,每一列包含 4 个卷积层,其中,第 1、2 卷积层后都跟着 1 个大小为 2×2 、步长为 2 的下采样层,第一列网络的卷积核尺寸依次为 9×9 、 7×7 、 7×7 、 7×7 ,对应的卷积核个数依次为 16、32、16、8,第二列卷积核尺寸为 7×7 、 5×5 、 5×5 、 5×5 ,个数为 20、40、20、10,第三列卷积核尺寸为 5×5 、 3×3 、 3×3 、 3×3 ,个数为 24、48、24、12	C_A 与 C_B 类似,不同点在于将 C_A 中的下采样层修改为 2×2 、步长为 2 的上采样层
DualCNN-Unet	包含 24 个卷积层,1~23 个卷积层的卷积核大小为 3×3 ,最后一个卷积层的卷积核大小为 1×1 ,第 1~24 个卷积层对应的卷积核个数依次为:32、32、64、64、128、128、256、256、512、512、256、256、256、128、128、128、64、64、64、32、32、32、2、1,其中第 2、4、6、8 个卷积层后都跟着 1 个大小为 2×2 、步长为 2 的下采样层,第 10、13、16、19 层后接一个尺寸 2×2 、步长为 2 的上采样层,在 11、14、17、20 层中卷积特征图分别与前面的 8、6、4、2 层输出特征图进行特征融合	C_A 与 C_B 类似,不同点在于 C_B 模型的第 24 个卷积层的卷积核个数是 3
DualCNN-CSRNet	包含 17 个卷积层,第 1~16 个卷积层的卷积核大小为 3×3 ,第 17 个卷积层卷积核大小为 1×1 ,前 10 层的结构与 VGG16 模型的前 10 层完全相同,11~16 层的卷积核空洞率 ^[3] 为 2,个数依次为 512、512、512、256、128、64	C_A 与 C_B 类似,不同点在于 C_B 模型的第 17 层的卷积核个数是 3

2.2 ShangHaitech 数据集

ShangHaitech 数据集是 ZHANG 等^[1] 在 2016 年发布的一个规模较大的数据集,其包含了 Part_A 和 Part_B 两部分,每一部分又各自包含训练集和测试集. Part_A 和 Part_B 总共有图片 1198 张,标记的人头数目为 330 165,其中 Part_A 是在网络上随机抽取的高密度人群图共有 482 张图片,平均每张图片 501 人,人数最多的图片中有 3139 人,人数最少的图片有 33 人. Part_B 是在上海街头抓拍的图片,相比于 Part_A 中的人数密度较稀疏,人数最少的图片仅有 9 人,最多的有 578 人.

表 2 给出了模型在 ShangHaitech 数据集上的性能. 从表 2 可以看出:1) DualCNN-MCNN 将 MCNN 在 Part_A 上 MAE 和 MSE 分别降低到 108.2 和 167.4,在 Part_B 上的分别降低到 21.0 和 36.0;2) DualCNN-UNet 大幅度提升 Unet 在 Part_A 上的性能,具体表现为 DualCNN-MCNN 将 Unet 在 Part_A 上的 MAE 和 MSE 分别降低了 100.4 和 124.7、在 Part_B 上的性能降低了 0.1 和 10.5;3) DualCNN-CSRNet 与 CSRNet 性能基本相当. 造成 DualCNN-CSRNet 与 CSRNet 性能相当的原因可能是 CSRNet 在 ShangHaitech 数据集上参数收敛到较为理想的值,难以进一步优化. 这些结果总体表明,提出的对偶神经网络能有效提升

卷积神经网络在人群计数问题上的性能.

表 2 ShangHaitech 数据集上的误差

Tab.2 Error in ShangHaitech dataset

方法	Part_A		Part_B	
	MAE	MSE	MAE	MSE
MCNN	110.2	173.2	26.4	41.3
DualCNN-MCNN	108.2	167.4	21.0	36.0
UNet	357.7	452.1	71.8	105.6
DualCNN-UNet	257.3	327.4	71.7	95.1
CSRNet	68.2	115.0	10.6	16.0
DualCNN-CSRNet	72.9	114.0	10.1	15.5

2.3 UCF_CC_50 数据集

UCF_CC_50 数据集^[14]是从公共可用的 Web 图像创建的. 作者收集了一些标签不同的图像,有体育场馆、音乐会、抗议和马拉松. 它总共包含 50 个分辨率不同的人群图像,每个图像平均有 1280 个人. 在整个数据集中共标记了 63 075 个人. 图片中的人数从 94 到 4543 不等,表明图像之间存在很大差异.

表 3 给出了模型在 UCF_CC_50 上的实验结果. 表 3 显示,提出对偶卷积神经网络(DualCNN)显著提升模型在 UCF_CC_50 上的性能. 具体表现为 DualCNN-MCNN 将 MCNN 的 MAE 和 MSE 分别降低 59.0 和 92.4, DualCNN-UNet 将 UNet 的 MAE 和 MSE 分别降低 0.3 和 4.1, DualCNN-

CSRNet 将 CSRNet 的 MAE 和 MSE 分别降低 42.3 和 114.5. 结合模型在 ShangHaitech 数据集上的结果(参考表 2),可以看出 DualCNN 在人数密集的数据集上效果更加明显.

表 3 UCF_CC_50 数据集上的误差

Tab. 3 Error in UCF_CC_50 dataset

方法	MAE	MSE
MCNN	377.6	509.1
DualCNN-MCNN	318.6	416.7
UNet	529.0	729.1
DualCNN-UNet	528.7	725.0
CSRNet	266.1	397.5
DualCNN-CSRNet	223.8	283.0

结合表 2 和表 3 的实验结果,可以得出:1)提出的对偶卷积神经网络能显著提升卷积神经网络

在人群计数上的性能;2)较之于在人数密集的数据集上的表现,该模型能更有效地提升卷积神经网络在人数密集的数据集上的性能.

3 结论

本文提出了一种对偶卷积神经网络模型用于提高卷积神经网络的性能,该模型包含两个子卷积神经网络模型,分别实现人群图到密度图的映射和密度图到人群图的映射.在学习过程中,利用两个子模型的相互影响,进而提高预测密度图的质量,获得高精度的人群分布密度图.实验结果表明,该方法能有效提升卷积神经网络在人群计数问题上预测准确率.

参考文献:

[1] ZHANG Y, ZHOU D, CHEN S, et al. Single-image crowd counting via multi-column convolutional neural network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. USA: Salt Lake City, UT, 2016:1091-1100.

[2] FU M, XU P, LI X, et al. Fast crowd density estimation with convolutional neural networks[J]. Engineering Applications of Artificial Intelligence, 2015, 43: 81-88.

[3] LI Y, ZHANG X, CHEN D. CSRNet: Dilated convolutional neural networks for understanding the highly congested scenes[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. USA: Salt Lake City, UT, 2018: 1091-1100.

[4] SAM D B, SURYA S, BABU R V. Switching convolutional neural network for crowd counting[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. USA: Honolulu, HI, 2017: 4031-4039.

[5] TIAN Y, LEI Y, ZHANG J, et al. PaDNet: Pan-density crowd counting[J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 2019, 29: 2714-2727.

[6] SHEN Z, XU Y, NI B, et al. Crowd counting via adversarial cross-scale consistency pursuit[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. USA : Salt Lake City, UT, 2018: 5245-5254.

[7] SINDAGI V A, PATEL V M. HA-CCN: Hierarchical attention-based crowd counting network [J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 2019, 29: 323-335.

[8] JIANG X, XIAO Z, ZHANG B, et al. Crowd counting and density estimation by trellis encoder-decoder networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. USA: Long Beach, CA, 2019: 6133-6142.

[9] LIU W, SALZMANN M, FUA P. Context-aware crowd counting[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. USA: Long Beach, CA, 2019: 5094-5103.

[10] CHEN J, SU W, WANG Z. Crowd counting with crowd attention convolutional neural network[J]. Neurocomputing, 2020, 382: 210-220.

[11] ZOU Z, CHENG Y, QU X, et al. Attend to count: Crowd counting with adaptive capacity multi-scale CNNs[J]. Neurocomputing, 2019, 367: 75-83.

[12] HE D, XIA Y, QIN T, et al. Dual learning for machine translation[C]//NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems. Spain: Barcelona, 2016: 820-828.

[13] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]// Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015, 9351: 234-241.

[14] IDREES H, SALEEMI I, SEIBERT C, et al. Multi-source multi-scale counting in extremely dense crowd images [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. USA: Portland, OR, 2013: 2547-2554.

责任编辑:郭红建