

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 8 月 27 日 (27.08.2020)



(10) 国际公布号
WO 2020/169043 A1

- (51) 国际专利分类号:
G06K 9/00 (2006.01) **G06N 3/04** (2006.01)
- (21) 国际申请号: PCT/CN2020/075795
- (22) 国际申请日: 2020 年 2 月 19 日 (19.02.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910129612.3 2019年2月21日 (21.02.2019) CN
- (71) 申请人: 苏州大学(SOOCHOW UNIVERSITY) [CN/CN]; 中国江苏省苏州市工业园区仁爱路 199 号, Jiangsu 215123 (CN)。
- (72) 发明人: 张莉(ZHANG, Li); 中国江苏省苏州市工业园区仁爱路 199 号, Jiangsu 215123 (CN)。 陆金刚(LU, Jingang); 中国江苏省苏州市工业园区仁爱路 199 号, Jiangsu 215123 (CN)。 周伟达(ZHOU, Weida); 中国江苏省苏州市工业园区仁爱路 199 号, Jiangsu 215123 (CN)。 王邦军(WANG, Bangjun); 中国江苏省苏州市工业园区仁爱路 199 号, Jiangsu

215123 (CN)。 章晓芳(ZHANG, Xiaofang); 中国江苏省苏州市工业园区仁爱路 199 号, Jiangsu 215123 (CN)。 屈蕴茜(QU, Yunqian); 中国江苏省苏州市工业园区仁爱路 199 号, Jiangsu 215123 (CN)。 赵雷(ZHAO, Lei); 中国江苏省苏州市工业园区仁爱路 199 号, Jiangsu 215123 (CN)。

(74) 代理人: 北京集佳知识产权代理有限公司 (UNITALEN ATTORNEYS AT LAW); 中国北京市朝阳区建国门外大街 22 号赛特广场 7 层, Beijing 100004 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: DENSE CROWD COUNTING METHOD, APPARATUS AND DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 一种密集人群计数的方法、装置、设备以及存储介质

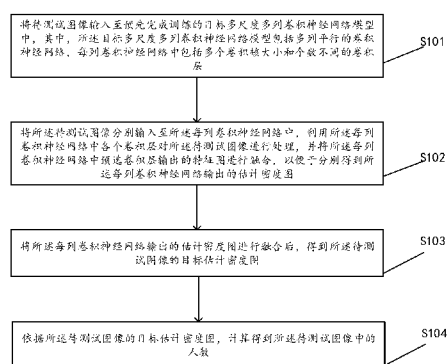


图 1

- S101 Input an image to be tested into a target multi-scale multi-column convolutional neural network model that has completed training in advance, wherein the target multi-scale multi-column convolutional neural network model comprises multiple columns of parallel convolutional neural networks, and each column of convolutional neural networks comprises multiple convolutional layers with different convolutional kernel sizes and quantities
- S102 Respectively input the image to be tested into each column of convolutional neural networks, process the image to be tested by using each convolutional layer in each column of convolutional neural networks, and fuse feature maps output by pre-selected convolutional layers in each column of convolutional neural networks, so as to respectively obtain estimated density maps output by each column of convolutional neural networks
- S103 Fuse the estimated density maps output by each column of convolutional neural networks to obtain a target estimated density map of the image to be tested
- S104 Calculate the number of people in the image to be tested according to the target estimated density map of the image to be tested

(57) Abstract: Provided are a dense crowd counting method, apparatus and device, and a computer-readable storage medium. The method comprises: inputting an image to be tested into a target multi-scale multi-column convolutional neural network model comprising multiple columns of parallel convolutional neural networks, wherein each column of convolutional neural networks comprises multiple convolutional layers with different convolutional kernel sizes and quantities; processing the image to be tested by using each convolutional layer in each column of convolutional neural networks, and fusing feature maps output by pre-selected convolutional layers in each column of convolutional neural networks, so as to obtain estimated density maps output by each column of convolutional neural networks; fusing the estimated density maps output by each column of convolutional neural networks to obtain a target estimated density map of the image to be tested; and calculating the number of people in the image to be tested according to the target estimated density map. By means of the provided method, apparatus and device and computer-readable storage medium, the accuracy of a dense crowd image prediction result is improved.

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

(57) 摘要: 一种密集人群计数的方法、装置、设备及计算机可读存储介质, 包括: 将待测试图像输入包括多列平行的卷积神经网络的目标多尺度多列卷积神经网络模型中; 每列卷积神经网络中包括多个卷积核大小和个数不同的卷积层。利用所述每列卷积神经网络中各个卷积层对待测试图像进行处理, 并将所述每列卷积神经网络中预选卷积层输出的特征图进行融合, 以便得到所述每列卷积神经网络输出的估计密度图; 将所述每列卷积神经网络输出的估计密度图进行融合后, 得到待测试图像的目标估计密度图; 依据所述目标估计密度图, 计算得到待测试图像中的人数。所提供的方法、装置、设备及计算机可读存储介质, 提高了密集人群图像预测结果的准确性。

一种密集人群计数的方法、装置、设备以及存储介质

本申请要求于 2019 年 2 月 21 日提交中国专利局、申请号为 201910129612.3、发明名称为“一种密集人群计数的方法、装置、设备以及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本发明涉及计算机视觉技术领域，特别是涉及一种密集人群计数的方法、装置、设备以及计算机可读存储介质。

10 背景技术

为了人群控制和公众安全，准确地估计来自图像或视频的人群已经成为计算机视觉技术越来越重要的应用。计算机视觉中的人群计数任务是自动计算图像或视频中的人数。为了在诸如公众集会和体育赛事等许多场景中帮助控制人群和公共安全，需要准确的人群计数。

15 传统的密集人群计数方法包括两种：基于检测的方法和基于回归的方法。基于检测的方法将人群视为一组检测到的个体实体。但是，行人经常被密集的人群遮挡，这在静止图像中估计人群时尤其具有挑战性。基于回归的方法对从人群图像中提取的各种特征的标量值（如人数）或密度图进行回归。它们基本上有两个步骤：首先，从人群图像中提取有效特征；第二，利用各种回归函数来估计人群数量。但是，通过回归进行的人群计数容易受到视角和尺度的急剧变化的影响，而这种变化通常存在于人群图像中。

25 与此同时，深度学习已经被成功地应用在密集人群图像的估计中。主流的估计方法采用密度图的思想，即设计一个神经网络，网络的输入为原始图像，而输出为人群的密度图。这类方法对密集人群图像处理的第一步，就是要通过一个高斯滤波器，根据图像的真实值 ground-truth 得到图像对应的密度图。Zhang 等人在“Single-Image Crowd Counting via Multi-Column Convolutional Neural Network”中提出了一个多列卷积神经网络。该网络由

-2-

三列平行的卷积神经网络组成，每一列使用感受野大小不同的卷积核，分别对应尺度大小不一样的人头；每一列除了卷积核大小和数量，其他构成都相同；采用大小为 的最大池化和 ReLU 激活函数；最后将三列的特征图在通道数上串联起来，用一个 的卷积核将其映射到估计的密度图输出。然而多列卷积神经网络结构偏简单，层数较少，前面的卷积层提取到的一些特征在后续过程中可能被丢弃且提取到的特征不够而影响到最终的结果。

综上所述可以看出，如何提高密集人群图像预测结果的准确性是目前有待解决的问题。

10 发明内容

本发明的目的是提供一种密集人群计数的方法、装置、设备以及计算机可读存储介质，以解决现有技术中提供的密集人群计数的神经网络性能较差的问题。

为解决上述技术问题，本发明提供一种密集人群计数的方法，包括：

15 将待测试图像输入至预先完成训练的目标多尺度多列卷积神经网络模型中；其中，所述目标多尺度多列卷积神经网络模型包括多列平行的卷积神经网络，每列卷积神经网络中包括多个卷积核大小和个数不同的卷积层；将所述待测试图像分别输入至所述每列卷积神经网络中，利用所述每列卷积神经网络中各个卷积层对所述待测试图像进行处理，并将所述每列卷积神经网络中预选卷积层输出的特征图进行融合，以便于分别得到所述每列卷积神经网络输出的估计密度图；将所述每列卷积神经网络输出的估计密度图进行融合后，得到所述待测试图像的目标估计密度图；依据所述待测试图像的目标估计密度图，计算得到所述待测试图像中的人数。

25 优选地，所述将待测试图像输入至预先完成训练的目标多尺度多列卷积神经网络模型中前包括：

利用高斯滤波器对预先创建的人群图像数据集进行滤波处理后，获取所述人群图像数据集中每幅图像的密度图，从而构建目标训练集；

采用所述目标训练集对多尺度多列卷积神经网络模型进行训练，得到完成训练后的目标多尺度多列卷积神经网络模型。

-3-

优选地, 所述利用高斯滤波器对预先创建的人群图像数据集进行滤波处理后, 获取所述人群图像数据集中每幅图像的密度图, 从而构建目标训练集包括:

获取预先采集的人群图像数据集 $D = \{(X_i, Y_i)\}_{i=1}^N$, 其中, X_i 为所述人

- 5 群图像数据集第 i 张图像, 大小为 $m*n$; Y_i 为所述第 i 张图像对应的人头坐标点图, 大小为 $m*n$, N 为所述人群图像数据集中图像总数;

利用高斯滤波器对所述人群图像数据集 $D = \{(X_i, Y_i)\}_{i=1}^N$ 中的每幅图像 X_i 进行滤波处理后, 得到所述每幅图像 X_i 的密度图 M_i , 利用所述每幅图像 X_i 的密度图 M_i 构建目标训练集 $D' = \{(X_i, M_i)\}_{i=1}^N$ 。

- 10 优选地, 所述采用所述目标训练集对多尺度多列卷积神经网络模型进行训练包括:

将所述目标训练集中的当前人群图像分别输入至所述多尺度多列卷积神经网络模型的每列卷积神经网络中;

- 15 其中, 所述多尺度多列卷积神经网络模型中的每列卷积神经网络相互平行, 所述每列卷积神经网络除卷积核大小和个数外, 其他网络结构相同;

将所述每列卷积神经网络输出的所述当前人群图像的估计密度图在通道数上串联后, 经过一个卷积核大小为 $1*1$ 的总卷积层, 并将所述总卷积层输出的特征图映射为所述当前人群图像的目标估计密度图, 以便于将所述当前人群图像的目标估计密度图作为所述多尺度多列卷积神经网络模型的网络输出。

- 20

优选地, 所述多尺度多列卷积神经网络模型的每列卷积神经网络包括:

第一卷积层、第二卷积层、第三卷积层、第四卷积层、第五卷积层、反卷积层、第六卷积层和第七卷积层;

- 25 其中, 所述第一卷积层和其他卷积层的卷积核大小不同, 所述第二卷积层、所述第三卷积层、所述第四卷积层、所述第五卷积层和所述第六卷积层的卷积核大小相同, 所述第三卷积层、所述第四卷积层、所述第五卷积层和所述第六卷积层的卷积核的个数相同;

所述第一卷积层、所述第二卷积层、所述第三卷积层和所述第四个卷

—4—

积层之间的池化层选用区域 2×2 ，步长为2的最大池化；

所述第四卷积层和所述第五卷积层之间的池化层选用 3×3 区域，步长为1的最大池化，以便于保持所述第四卷积层输出特征图和对所述第四卷积层输出特征池化后的特征图大小不变；

5 所述各个卷积层的激活函数采用 ReLU 函数；

所述第四卷积层输出的特征图和所述第五卷积层输出的特征图在通道数上串联后输入所述反卷积层，所述反卷积层输出的特征图和所述第三卷积层输出的特征图在通道数上串联后输入所述第六卷积层，所述第八卷积层输出所述待测试图像的估计密度图作为所述每列卷积神经网络模型的输出结果。

优选地，所述依据所述待测试图像的目标估计密度图，计算得到所述待测试图像中的人数包括：

将所述待测试图像 T 输入至所述目标多尺度多列卷积神经网络模型，得到所述待测试图像 T 的估计密度图 $\hat{M}$ 后，计算所述估计密度图 $\hat{M}$ 中所有像素值的和，得到所述待测试图像中的人数 $et\_count = \text{sum}(\hat{M})$ 。

本发明还提供了一种密集人群计数的装置，包括：

输入模块，用于将待测试图像输入至预先完成训练的目标多尺度多列卷积神经网络模型中；其中，所述目标多尺度多列卷积神经网络模型包括多列平行的卷积神经网络，每列卷积神经网络中包括多个卷积核大小和个数不同的卷积层；

处理模块，用于将所述待测试图像分别输入至所述每列卷积神经网络中，利用所述每列卷积神经网络中各个卷积层对所述待测试图像进行处理，并将所述每列卷积神经网络中预选卷积层输出的特征图进行融合，以便于分别得到所述每列卷积神经网络输出的估计密度图；

25 输出模块，用于将所述每列卷积神经网络输出的估计密度图进行融合后，得到所述待测试图像的目标估计密度图；

计算模块，用于依据所述待测试图像的目标估计密度图，计算得到所述待测试图像中的人数。

优选地，所述输出模块前包括：

—5—

训练模块，用于利用高斯滤波器对预先创建的人群图像数据集进行滤波处理后，获取所述人群图像数据集中每幅图像的密度图，从而构建目标训练集；

采用所述目标训练集对多尺度多列卷积神经网络模型进行训练，得到
5 完成训练后的目标多尺度多列卷积神经网络模型。

本发明还提供了一种密集人群计数的设备，包括：

存储器，用于存储计算机程序；处理器，用于执行所述计算机程序时实现上述一种密集人群计数的方法的步骤。

本发明还提供了一种计算机可读存储介质，所述计算机可读存储介质
10 上存储有计算机程序，所述计算机程序被处理器执行时实现上述一种密集人群计数的方法的步骤。

本发明所提供的密集人群计数的方法，利用预先完成训练的目标多尺度多列卷积神经网络模型对待测试图像进行预测。所述目标多尺度多列卷积神经网络模型包括多列平行的卷积神经网络。将所述待测试图像输入所述目标多尺度多列卷积神经网络模型后，将所述待测试图像分别输入所述
15 每列卷积神经网络中。所述每列卷积神经网络中包括多个卷积核大小和个数不同的卷积层，分别利用所述每列卷积神经网络中的不同卷积层对所述待测试图像进行计算，并将所述每列卷积神经网络中预选的卷积层输出的特征图进行融合，提取到所述待测试图像的不同尺度的特征；解决了现有技术中的卷积神经网络中前面卷积层提取到的一些特征在后续过程中可能被丢弃导致提取到的特征不够从而影响了对待测试图像预测结果的准确性的问题。本发明所提供的方法，引入了多尺度的思想，可以将前面卷积层提取到的特征和后面卷积层提取到的特征结合起来，即将详细程度不同的特征结合起来进而提取特征，这就弥补了传统神经网络前面的卷积层得到
20 的特征图经过池化可能被丢弃的一些特征，提高了密集人群计数的神经网络的性能以及密集人群图像预测结果的准确性。

附图说明

为了更清楚的说明本发明实施例或现有技术的技术方案，下面将对实

—6—

施例或现有技术描述中所需要使用的附图作简单的介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图 1 为本发明所提供的密集人群计数的方法的第一种具体实施例的流程图；

图 2 为本发明所提供的多尺度多列卷积神经网络结构图；

图 3 为本发明所提供的密集人群计数的方法的第二种具体实施例的流程图；

图 4 为本发明实施例提供的一种密集人群计数的装置的结构框图。

10

具体实施方式

本发明的核心是提供一种密集人群计数的方法、装置、设备以及计算机可读存储介质，提高了密集人群计数的神经网络的性能以及密集人群图像预测结果的准确性。

15 为了使本技术领域的人员更好地理解本发明方案，下面结合附图和具体实施方式对本发明作进一步的详细说明。显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

20 请参考图 1，图 1 为本发明所提供的密集人群计数的方法的第一种具体实施例的流程图；具体操作步骤如下：

步骤 S101：将待测试图像输入至预先完成训练的目标多尺度多列卷积神经网络模型中，其中，所述目标多尺度多列卷积神经网络模型包括多列平行的卷积神经网络，每列卷积神经网络中包括多个卷积核大小和个数不同的卷积层；

25

将待测试图像输入至预先完成训练的目标多尺度多列卷积神经网络模型中前需要对多尺度多列卷积神经网络（SaMCNN）进行训练。

对所述多尺度多列卷积神经网络进行训练时，首先利用高斯滤波

-7-

器对预先创建的人群图像数据集 $D = \{(X_i, Y_i)\}_{i=1}^N$ 进行滤波处理后, 获取所述人群图像数据集中每幅图像 X_i 的密度图 M_i , 从而构建目标训练集 $D' = \{(X_i, M_i)\}_{i=1}^N$; 其中, X_i 为所述人群图像数据集第 i 张图像, 大小为 $m*n$; Y_i 为所述第 i 张图像对应的人头坐标点图, 大小为 $m*n$, N 为所述人群图像数据集中图像总数。采用所述目标训练集 $D' = \{(X_i, M_i)\}_{i=1}^N$ 对多尺度多列卷积神经网络模型进行训练, 得到完成训练后的目标多尺度多列卷积神经网络模型。

如图 2 所示, 所述多尺度多列卷积神经网络中可以包括多列卷积神经网络, 在本实施例中以三列平行的卷积神经网络为例。所述每列卷积神经网络包括第一卷积层、第二卷积层、第三卷积层、第四卷积层、第五卷积层、反卷积层、第六卷积层和第七卷积层。其中, 所述第一卷积层和其他卷积层的卷积核大小不同, 所述第二卷积层、所述第三卷积层、所述第四卷积层、所述第五卷积层和所述第六卷积层的卷积核大小相同, 所述第三卷积层、所述第四卷积层、所述第五卷积层和所述第六卷积层的卷积核的个数相同。所述各个卷积层的激活函数采用 ReLU 函数。

所述第一卷积层、所述第二卷积层、所述第三卷积层和所述第四个卷积层之间的池化层选用区域 $2*2$, 步长为 2 的最大池化; 所述第四卷积层和所述第五卷积层之间的池化层选用 $3*3$ 区域, 步长为 1 的最大池化, 以便于保持所述第四卷积层输出特征图和对所述第四卷积层输出特征池化后的特征图大小不变。

所述第四卷积层输出的特征图和所述第五卷积层输出的特征图在通道数上串联后输入所述反卷积层, 所述反卷积层输出的特征图和所述第三卷积层输出的特征图在通道数上串联后输入所述第六卷积层, 所述第八卷积层输出所述待测试图像的估计密度图作为所述每列卷积神经网络模型的输出结果。

将所述每列卷积神经网络输出的所述当前人群图像的估计密度图在通道数上串联后, 经过一个卷积核大小为 $1*1$ 的总卷积层, 并将所述总卷积层输出的特征图映射为所述当前人群图像的目标估计密度图, 以便于将所述

—8—

当前人群图像的目标估计密度图作为所述多尺度多列卷积神经网络模型的网络输出。

步骤 S102：将所述待测试图像分别输入至所述每列卷积神经网络中，利用所述每列卷积神经网络中各个卷积层对所述待测试图像进行处理，并
5 将所述每列卷积神经网络中预选卷积层输出的特征图进行融合，以便于分别得到所述每列卷积神经网络输出的估计密度图；

将所述待测试图像输入所述目标多尺度多列卷积神经网络模型中，并分别将所述待测试图像分别输入所述目标多尺度多列卷积神经网络模型的每列卷积神经网络中。所述每列卷积神经网络中的卷积层对所述待测试数据进行处理。利用所述每列卷积神经网络中的各个卷积层和池化层
10 进行处理，将所述每列卷积神经网络的第四卷积层和第五卷积层之间选用 3×3 区域，步长为 1 的最大池化，来保持池化前后的特征图大小不变，便于将两次卷积之后的特征图在通道数上串联起来。在所述第五卷积层之后，使用反卷积层对之前的特征图进行上采样，进而和第三个卷积层得到的特征图在通道数上串联起来。
15

步骤 S103：将所述每列卷积神经网络输出的估计密度图进行融合后，得到所述待测试图像的目标估计密度图；

步骤 S104：依据所述待测试图像的目标估计密度图，计算得到所述待测试图像中的人数。

20 本实施例所提供的方法，利用多尺度多列卷积神经网络对待测试图像进行测试。所述多尺度多列卷积神经网络相对于多列卷积神经网络，增加了每列卷积神经网络的层数，并且引入了多尺度的思想，将前面卷积层提取到的特征图和后面卷积层提取到的特征图相结合；从而提高了密集人群计数的神经网络的性能以及密集人群图像预测结果的准确性。

25

基于上述实施例，在本实施例中，可以选择 Shanghai tech 数据集的第二部分作为人群图像数据集，利用所述人群图像数据集中第二部分图像的密级图对所述多尺度多列卷积神经网络模型进行训练。请参考图 3，图 3 为本发明所提供的密集人群计数的方法的第二种具体实施例的流程图；具

体操作步骤如下：

步骤 301：利用高斯滤波器对 Shanghai tech 数据集的第二部分的人群图像进行滤波处理后，获取所述第二部分的人群图像的度图，构建目标训练集；

5 在本实施例中可以选择 Shanghai tech 数据集的第二部分作为人群图像数据集 $D = \{(\mathbf{X}_i, \mathbf{Y}_i)\}_{i=1}^N$ ， $\mathbf{X}_i$ 为所述人群图像数据集第 i 张图像，大小为 768×1024 ； $\mathbf{Y}_i$ 为所述第 i 张图像对应的人头坐标点图，大小为 768×1024 ， N 为所述人群图像数据集中图像总数。

10 所述 Shanghai tech 数据集包含 1198 张标注的图像和 330165 个人头中心标注；所述 Shanghai tech 数据集被划分为两个部分，其中，第一部分包括 482 张随机从网上爬取的图像，其中 300 张用于训练，182 张用于测试；第二部分包括 716 张在上海街头拍取的图像，其中 400 张用于训练，316 张用于测试。

15 步骤 302：采用所述目标训练集对多尺度多列卷积神经网络模型进行训练，得到完成训练后的目标多尺度多列卷积神经网络模型；

 步骤 303：将待测试图像 $\mathbf{T}$ 输入至所述目标多尺度多列卷积神经网络模型中，其中，所述目标多尺度多列卷积神经网络模型包括多列平行的卷积神经网络，每列卷积神经网络中包括多个卷积核大小和个数不同的卷积层；

20 步骤 304：将所述待测试图像 $\mathbf{T}$ 输入至所述目标多尺度多列卷积神经网络模型后，输出所述待测试图像 $\mathbf{T}$ 的估计密度图 $\hat{\mathbf{M}}$ ；

 步骤 S305：计算所述估计密度图 $\hat{\mathbf{M}}$ 中所有像素值的和，得到所述待测试图像中的人数 $et\_count = \text{sum}(\hat{\mathbf{M}})$ 。

25 将本实施例所提供的多尺度多列卷积神经网络模型与多列卷积神经网络模型在相同的数据集上做人群计数比较。从表 1 可以得到，本实施例提出的网络模型的计数结果的平均完全误差（MAE）和均方误差（MSE）都比现有技术中网络模型的计数结果小，获得了更好的性能。

表-1 人群计数结果的对比

网络模型	Part_A	Part_B
------	--------	--------

	MAE	MSE	MAE	MSE
ScMCNN	89.08	142.83	20.30	37.23
MCNN	110.92	172.51	26.28	49.06

请参考图 4，图 4 为本发明实施例提供的一种密集人群计数的装置的结构框图。具体装置可以包括：

输入模块 100，用于将待测试图像输入至预先完成训练的目标多尺度多列卷积神经网络模型中；其中，所述目标多尺度多列卷积神经网络模型包括多列平行的卷积神经网络，每列卷积神经网络中包括多个卷积核大小和个数不同的卷积层；

处理模块 200，用于将所述待测试图像分别输入至所述每列卷积神经网络中，利用所述每列卷积神经网络中各个卷积层对所述待测试图像进行处理，并将所述每列卷积神经网络中预选卷积层输出的特征图进行融合，以便于分别得到所述每列卷积神经网络输出的估计密度图；

输出模块 300，用于将所述每列卷积神经网络输出的估计密度图进行融合后，得到所述待测试图像的目标估计密度图；

计算模块 400，用于依据所述待测试图像的目标估计密度图，计算得到所述待测试图像中的人数。

本实施例的密集人群计数的装置用于实现前述的密集人群计数的方法，因此密集人群计数的装置中的具体实施方式可见前文中的密集人群计数的方法的实施例部分，例如，输入模块 100，处理模块 200，输出模块 300，计算模块 400，分别用于实现上述密集人群计数的方法中步骤 S101，S102，S103 和 S104，所以，其具体实施方式可以参照相应的各个部分实施例的描述，在此不再赘述。

本发明具体实施例还提供了一种密集人群计数的设备，包括：存储器，用于存储计算机程序；处理器，用于执行所述计算机程序时实现上述一种密集人群计数的方法的步骤。

本发明具体实施例还提供了一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机程序，所述计算机程序被处理器执行时实现上

- 11 -

述一种密集人群计数的方法的步骤。

本说明书中各个实施例采用递进的方式描述，每个实施例重点说明的都是与其它实施例的不同之处，各个实施例之间相同或相似部分互相参见即可。对于实施例公开的装置而言，由于其与实施例公开的方法相对应，
5 所以描述的比较简单，相关之处参见方法部分说明即可。

专业人员还可以进一步意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本发明的范围。
10

结合本文中所公开的实施例描述的方法或算法的步骤可以直接用硬件、处理器执行的软件模块，或者二者的结合来实施。软件模块可以置于
15 随机存储器（RAM）、内存、只读存储器（ROM）、电可编程 ROM、电可擦除可编程 ROM、寄存器、硬盘、可移动磁盘、CD-ROM、或技术领域内所公知的任意其它形式的存储介质中。

以上对本发明所提供的密集人群计数的方法、装置、设备 以及计算机可读存储介质进行了详细介绍。本文中应用了具体个例对本发明的原理及
20 实施方式进行了阐述，以上实施例的说明只是用于帮助理解本发明的方法及其核心思想。应当指出，对于本技术领域的普通技术人员来说，在不脱离本发明原理的前提下，还可以对本发明进行若干改进和修饰，这些改进和修饰也落入本发明权利要求要求的保护范围内。

权 利 要 求

1、一种密集人群计数的方法，其特征在于，包括：

将待测试图像输入至预先完成训练的目标多尺度多列卷积神经网络模型中；其中，所述目标多尺度多列卷积神经网络模型包括多列平行的卷积神经网络，每列卷积神经网络中包括多个卷积核大小和个数不同的卷积层；

将所述待测试图像分别输入至所述每列卷积神经网络中，利用所述每列卷积神经网络中各个卷积层对所述待测试图像进行处理，并将所述每列卷积神经网络中预选卷积层输出的特征图进行融合，以便于分别得到所述每列卷积神经网络输出的估计密度图；

10 将所述每列卷积神经网络输出的估计密度图进行融合后，得到所述待测试图像的目标估计密度图；

依据所述待测试图像的目标估计密度图，计算得到所述待测试图像中的人数。

2、如权利要求 1 所述的方法，其特征在于，所述将待测试图像输入至预先完成训练的目标多尺度多列卷积神经网络模型中前包括：

利用高斯滤波器对预先创建的人群图像数据集进行滤波处理后，获取所述人群图像数据集中每幅图像的密度图，从而构建目标训练集；

采用所述目标训练集对多尺度多列卷积神经网络模型进行训练，得到完成训练后的目标多尺度多列卷积神经网络模型。

20 3、如权利要求 2 所述的方法，其特征在于，所述利用高斯滤波器对预先创建的人群图像数据集进行滤波处理后，获取所述人群图像数据集中每幅图像的密度图，从而构建目标训练集包括：

获取预先采集的人群图像数据集 $D = \{(\mathbf{X}_i, \mathbf{Y}_i)\}_{i=1}^N$ ，其中， $\mathbf{X}_i$ 为所述人群图像数据集第 i 张图像，大小为 $m*n$ ； $\mathbf{Y}_i$ 为所述第 i 张图像对应的人头坐
25 标点图，大小为 $m*n$ ， N 为所述人群图像数据集中图像总数；

利用高斯滤波器对所述人群图像数据集 $D = \{(\mathbf{X}_i, \mathbf{Y}_i)\}_{i=1}^N$ 中的每幅图像 $\mathbf{X}_i$ 进行滤波处理后，得到所述每幅图像 $\mathbf{X}_i$ 的密度图 $\mathbf{M}_i$ ，利用所述每幅图像 $\mathbf{X}_i$ 的密度图 $\mathbf{M}_i$ 构建目标训练集 $D' = \{(\mathbf{X}_i, \mathbf{M}_i)\}_{i=1}^N$ 。

-13-

4、如权利要求2所述的方法，其特征在于，所述采用所述目标训练集对多尺度多列卷积神经网络模型进行训练包括：

将所述目标训练集中的当前人群图像分别输入至所述多尺度多列卷积神经网络模型的每列卷积神经网络中；

5 其中，所述多尺度多列卷积神经网络模型中的每列卷积神经网络相互平行，所述每列卷积神经网络除卷积核大小和个数外，其他网络结构相同；

将所述每列卷积神经网络输出的所述当前人群图像的估计密度图在通道数上串联后，经过一个卷积核大小为1*1的总卷积层，并将所述总卷积层输出的特征图映射为所述当前人群图像的目标估计密度图，以便于将所述
10 当前人群图像的目标估计密度图作为所述多尺度多列卷积神经网络模型的网络输出。

5、如权利要求4所述的方法，其特征在于，所述多尺度多列卷积神经网络模型的每列卷积神经网络包括：

第一卷积层、第二卷积层、第三卷积层、第四卷积层、第五卷积层、
15 反卷积层、第六卷积层和第七卷积层；

其中，所述第一卷积层和其他卷积层的卷积核大小不同，所述第二卷积层、所述第三卷积层、所述第四卷积层、所述第五卷积层和所述第六卷积层的卷积核大小相同，所述第三卷积层、所述第四卷积层、所述第五卷积层和所述第六卷积层的卷积核的个数相同；

20 所述第一卷积层、所述第二卷积层、所述第三卷积层和所述第四个卷积层之间的池化层选用区域2*2，步长为2的最大池化；

所述第四卷积层和所述第五卷积层之间的池化层选用3*3区域，步长为1的最大池化，以便于保持所述第四卷积层输出特征图和对所述第四卷积层输出特征池化后的特征图大小不变；

25 所述各个卷积层的激活函数采用ReLU函数；

所述第四卷积层输出的特征图和所述第五卷积层输出的特征图在通道数上串联后输入所述反卷积层，所述反卷积层输出的特征图和所述第三卷积层输出的特征图在通道数上串联后输入所述第六卷积层，所述第八卷积层输出所述待测试图像的估计密度图作为所述每列卷积神经网络模型的输

出结果。

6、如权利要求1至5任一项所述的方法，其特征在于，所述依据所述待测试图像的目标估计密度图，计算得到所述待测试图像中的人数包括：

5 将所述待测试图像 T 输入至所述目标多尺度多列卷积神经网络模型，得到所述待测试图像 T 的估计密度图 $\hat{M}$ 后，计算所述估计密度图 $\hat{M}$ 中所有像素值的和，得到所述待测试图像中的人数 $et\_count = \text{sum}(\hat{M})$ 。

7、一种密集人群计数的装置，其特征在于，包括：

10 输入模块，用于将待测试图像输入至预先完成训练的目标多尺度多列卷积神经网络模型中；其中，所述目标多尺度多列卷积神经网络模型包括多列平行的卷积神经网络，每列卷积神经网络中包括多个卷积核大小和个数不同的卷积层；

15 处理模块，用于将所述待测试图像分别输入至所述每列卷积神经网络中，利用所述每列卷积神经网络中各个卷积层对所述待测试图像进行处理，并将所述每列卷积神经网络中预选卷积层输出的特征图进行融合，以便于分别得到所述每列卷积神经网络输出的估计密度图；

输出模块，用于将所述每列卷积神经网络输出的估计密度图进行融合后，得到所述待测试图像的目标估计密度图；

计算模块，用于依据所述待测试图像的目标估计密度图，计算得到所述待测试图像中的人数。

20 8、如权利要求7所述的装置，其特征在于，所述输出模块前包括：

训练模块，用于利用高斯滤波器对预先创建的人群图像数据集进行滤波处理后，获取所述人群图像数据集中每幅图像的密度图，从而构建目标训练集；

25 采用所述目标训练集对多尺度多列卷积神经网络模型进行训练，得到完成训练后的目标多尺度多列卷积神经网络模型。

9、一种密集人群计数的设备，其特征在于，包括：

存储器，用于存储计算机程序；

处理器，用于执行所述计算机程序时实现如权利要求1至7任一项所述一种密集人群计数的方法的步骤。

—15—

10、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有计算机程序，所述计算机程序被处理器执行时实现如权利要求1至7 任一项所述一种密集人群计数的方法的步骤。

— 1/3 —

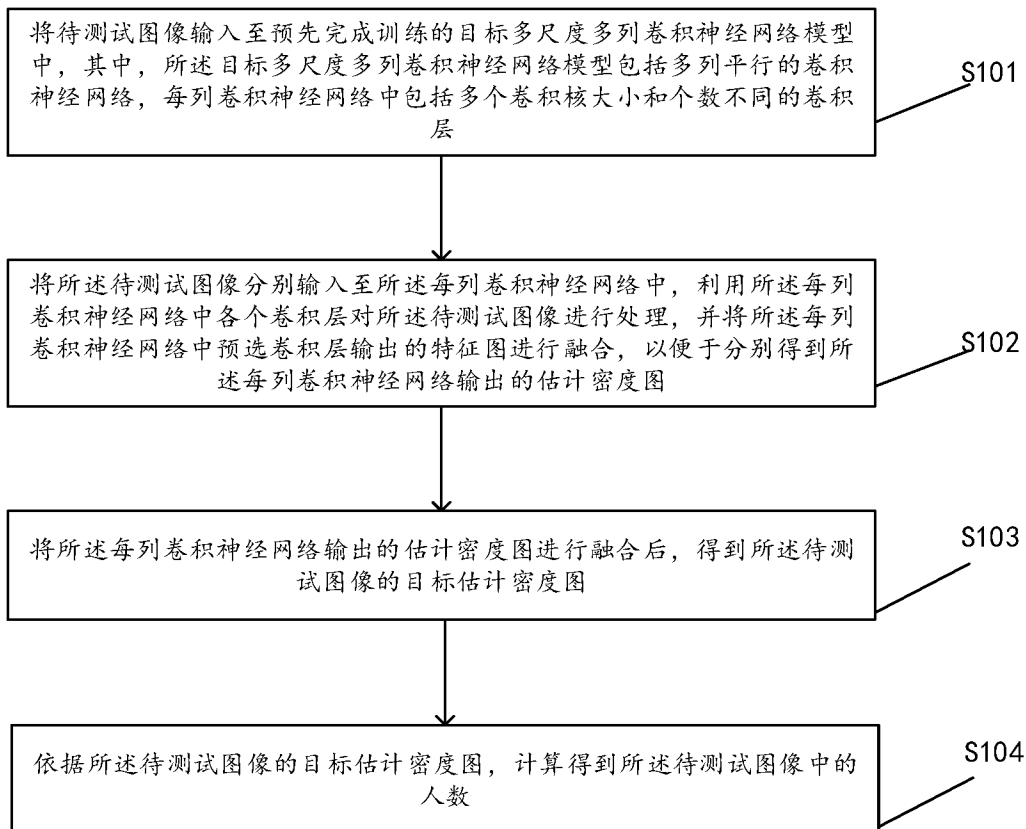


图 1

— 2/3 —

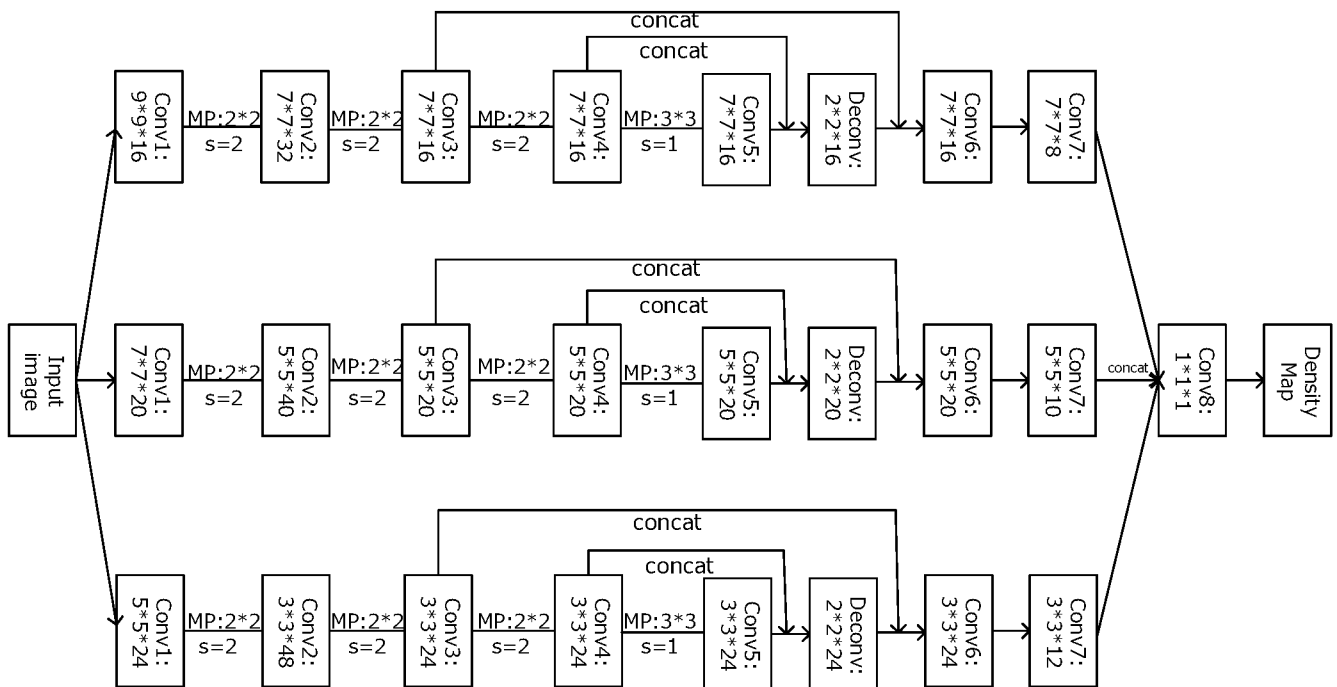


图 2

— 3/3 —

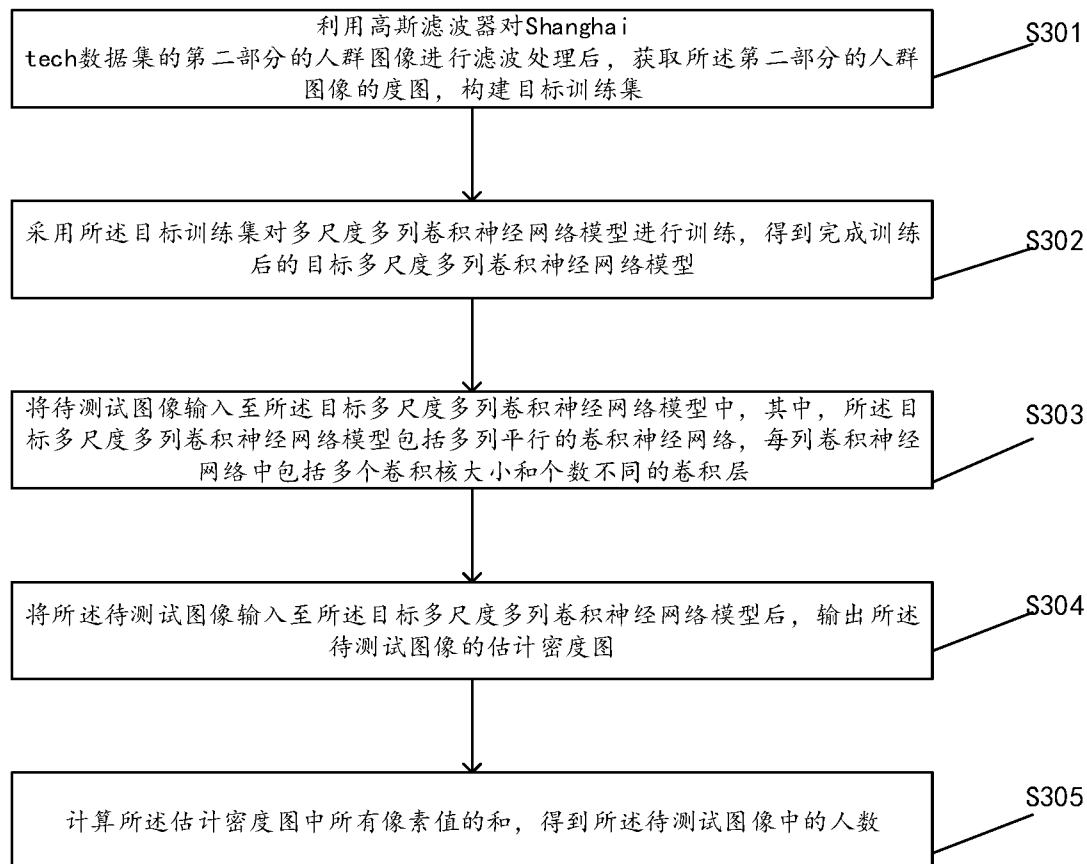


图 3

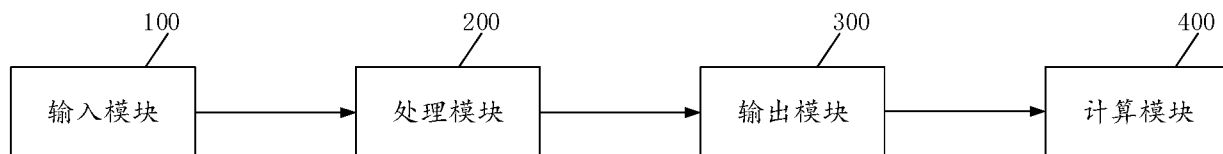


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/075795

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i; G06N 3/04(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K; G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI: 卷积, 神经网络, 训练, 人群, 密集, 密度, 多路, 多列, 多尺度, 密度图, 融合, CNN, dense, crowd, multi column, multi scale, merge

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109858461 A (SOOCHOW UNIVERSITY) 07 June 2019 (2019-06-07) description, paragraphs [0007]-[0037], and claims 1-10	1-10
X	唐斯琪 等 (TANG, Siqu et al.). "一种多列特征图融合的深度人群计数算法 (A Deep Crowd Counting Algorithm Based on Multi-Column Feature Map Fusion)" 郑州大学学报 (理学版) (Journal of Zhengzhou University (Natural Science Edition)), Vol. 50, No. (2), 30 June 2018 (2018-06-30), section 2	1, 6-7
Y	唐斯琪 等 (TANG, Siqu et al.). "一种多列特征图融合的深度人群计数算法 (A Deep Crowd Counting Algorithm Based on Multi-Column Feature Map Fusion)" 郑州大学学报 (理学版) (Journal of Zhengzhou University (Natural Science Edition)), Vol. 50, No. (2), 30 June 2018 (2018-06-30), section 2	2-4, 8-10
Y	CN 109214337 A (SOOCHOW UNIVERSITY) 15 January 2019 (2019-01-15) description, paragraphs [0053]-[0098]	2-4, 8-10
A	CN 109271960 A (YANSHAN UNIVERSITY) 25 January 2019 (2019-01-25) entire document	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

07 May 2020

Date of mailing of the international search report

19 May 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/075795**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2018052586 A1 (KONICA MINOLTA LABORATORY U.S.A., INC.) 22 March 2018 (2018-03-22) entire document	1-10
A	CN 109101930 A (HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY) 28 December 2018 (2018-12-28) entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/075795

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109858461	A	07 June 2019	None			
CN	109214337	A	15 January 2019	None			
CN	109271960	A	25 January 2019	None			
WO	2018052586	A1	22 March 2018	US	2019236411	A1	01 August 2019
CN	109101930	A	28 December 2018	None			

国际检索报告

国际申请号

PCT/CN2020/075795

A. 主题的分类 G06K 9/00(2006.01)i; G06N 3/04(2006.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类																							
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G06K; G06N 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) WPI, EPDOC, CNPAT, CNKI: 卷积, 神经网络, 训练, 人群, 密集, 密度, 多路, 多列, 多尺度, 密度图, 融合, CNN, dense, crowd, multi column, multi scale, merge																							
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 109858461 A (苏州大学) 2019年 6月 7日 (2019 - 06 - 07) 说明书第[0007]-[0037]段, 权利要求1-10</td> <td>1-10</td> </tr> <tr> <td>X</td> <td>唐斯琪 等. "一种多列特征图融合的深度人群计数算法" 《郑州大学学报(理学版)》, 第50卷, 第2期, 2018年 6月 30日 (2018 - 06 - 30), 第2部分</td> <td>1, 6-7</td> </tr> <tr> <td>Y</td> <td>唐斯琪 等. "一种多列特征图融合的深度人群计数算法" 《郑州大学学报(理学版)》, 第50卷, 第2期, 2018年 6月 30日 (2018 - 06 - 30), 第2部分</td> <td>2-4, 8-10</td> </tr> <tr> <td>Y</td> <td>CN 109214337 A (苏州大学) 2019年 1月 15日 (2019 - 01 - 15) 说明书第[0053]-[0098]段</td> <td>2-4, 8-10</td> </tr> <tr> <td>A</td> <td>CN 109271960 A (燕山大学) 2019年 1月 25日 (2019 - 01 - 25) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>WO 2018052586 A1 (KONICA MINOLTA LABORATORY U.S.A., INC.) 2018年 3月 22日 (2018 - 03 - 22) 全文</td> <td>1-10</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 109858461 A (苏州大学) 2019年 6月 7日 (2019 - 06 - 07) 说明书第[0007]-[0037]段, 权利要求1-10	1-10	X	唐斯琪 等. "一种多列特征图融合的深度人群计数算法" 《郑州大学学报(理学版)》, 第50卷, 第2期, 2018年 6月 30日 (2018 - 06 - 30), 第2部分	1, 6-7	Y	唐斯琪 等. "一种多列特征图融合的深度人群计数算法" 《郑州大学学报(理学版)》, 第50卷, 第2期, 2018年 6月 30日 (2018 - 06 - 30), 第2部分	2-4, 8-10	Y	CN 109214337 A (苏州大学) 2019年 1月 15日 (2019 - 01 - 15) 说明书第[0053]-[0098]段	2-4, 8-10	A	CN 109271960 A (燕山大学) 2019年 1月 25日 (2019 - 01 - 25) 全文	1-10	A	WO 2018052586 A1 (KONICA MINOLTA LABORATORY U.S.A., INC.) 2018年 3月 22日 (2018 - 03 - 22) 全文	1-10
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 109858461 A (苏州大学) 2019年 6月 7日 (2019 - 06 - 07) 说明书第[0007]-[0037]段, 权利要求1-10	1-10																					
X	唐斯琪 等. "一种多列特征图融合的深度人群计数算法" 《郑州大学学报(理学版)》, 第50卷, 第2期, 2018年 6月 30日 (2018 - 06 - 30), 第2部分	1, 6-7																					
Y	唐斯琪 等. "一种多列特征图融合的深度人群计数算法" 《郑州大学学报(理学版)》, 第50卷, 第2期, 2018年 6月 30日 (2018 - 06 - 30), 第2部分	2-4, 8-10																					
Y	CN 109214337 A (苏州大学) 2019年 1月 15日 (2019 - 01 - 15) 说明书第[0053]-[0098]段	2-4, 8-10																					
A	CN 109271960 A (燕山大学) 2019年 1月 25日 (2019 - 01 - 25) 全文	1-10																					
A	WO 2018052586 A1 (KONICA MINOLTA LABORATORY U.S.A., INC.) 2018年 3月 22日 (2018 - 03 - 22) 全文	1-10																					
☒ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																							
* 引用文件的具体类型: "A" 认为不特别相关的表示了现有技术一般状态的文件 "E" 在国际申请日的当天或之后公布的在先申请或专利 "L" 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) "O" 涉及口头公开、使用、展览或其他方式公开的文件 "P" 公布日先于国际申请日但迟于所要求的优先权日的文件 "T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 "X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 "Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 "&" 同族专利的文件																							
国际检索实际完成的日期 2020年 5月 7日		国际检索报告邮寄日期 2020年 5月 19日																					
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 丁文勃 电话号码 86-(10)-53961330																					

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 109101930 A (华中科技大学) 2018年 12月 28日 (2018 - 12 - 28) 全文	1-10

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/075795

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	109858461	A	2019年 6月 7日	无	
CN	109214337	A	2019年 1月 15日	无	
CN	109271960	A	2019年 1月 25日	无	
WO	2018052586	A1	2018年 3月 22日	US 2019236411 A1	2019年 8月 1日
CN	109101930	A	2018年 12月 28日	无	