

(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
2018年6月21日 (21.06.2018)



(10) 国际公布号
WO 2018/107760 A1

- (51) 国际专利分类号:
G06K 9/62 (2006.01)
- (21) 国际申请号: PCT/CN2017/094016
- (22) 国际申请日: 2017年7月24日 (24.07.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201611166903.2 2016年12月16日 (16.12.2016) CN
- (71) 申请人: 北京大学深圳研究生院(PEKING UNIVERSITY SHENZHEN GRADUATE SCHOOL) [CN/CN]; 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。

(72) 发明人: 王文敏(WANG, Wenmin); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 松鸿蒙(SONG, Hongmeng); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 王荣刚(WANG, Ronggang); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 李革(LI, Ge); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 董胜富(DONG, Shengfu); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 王振宇(WANG, Zhenyu); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 李英(LI, Ying); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 赵辉(ZHAO, Hui); 中

(54) Title: COLLABORATIVE DEEP NETWORK MODEL METHOD FOR PEDESTRIAN DETECTION

(54) 发明名称: 一种用于行人检测的协同式深度网络模型方法

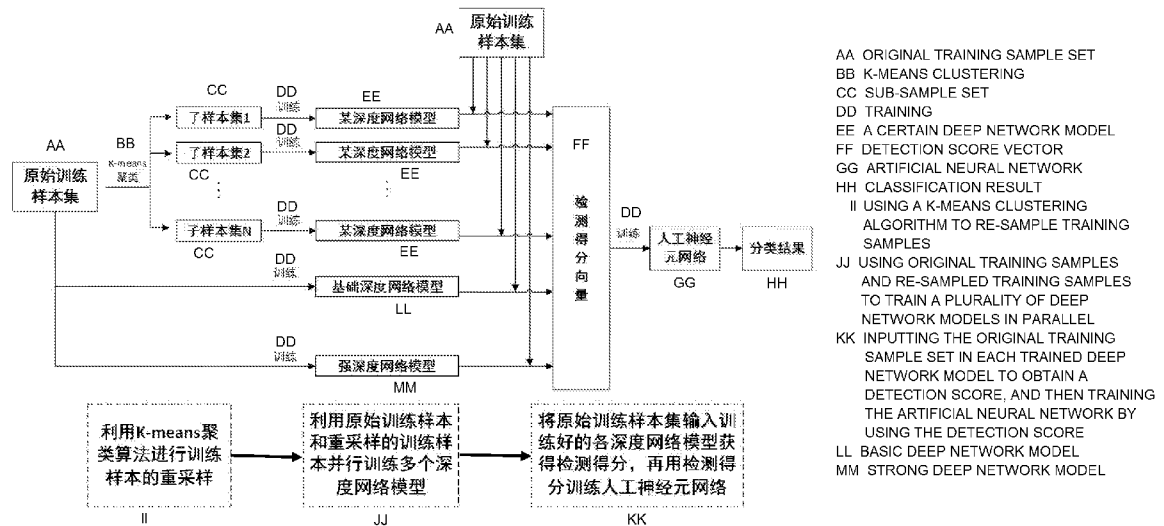


图1

(57) Abstract: Disclosed in the present invention is a collaborative deep network model method for pedestrian detection. The method comprises: constructing a new collaborative multi-model learning framework to complete a classification process during pedestrian detection; and using an artificial neural network to integrate determination results of sub-classifiers in a collaborative model, and training the network by means of the method for machine learning, so that information fed back by sub-classifiers can be more effectively synthesized. Provided is a re-sampling method based on a K-means clustering algorithm. The re-sampling method enhances the classification effect of each classifier in the collaborative model, and thus improves the overall classification effect. In the present invention, by building a collaborative deep network model, different types of training data sets obtained by means of a clustering algorithm are used for training a plurality of deep network models in parallel, and then classification results, on deep network models, of an original data set are integrated and comprehensively analysed by means of the artificial neural network, so as to realise more accurate sample

国广东省深圳市南山区西丽镇丽水路深圳大学城
北大校区, Guangdong 518055 (CN)。 高文(GAO,
Wen); 中国广东省深圳市南山区西丽镇丽水路深
圳大学城北大校区, Guangdong 518055 (CN)。

(74) 代理人: 北京万象新悦知识产权代理事务
所(普通合伙)(BEIJING WANXIANGXINYUE
INTELLECTUAL PROPERTY OFFICE); 中国北京
市海淀区颐和园路1号北京资源燕园宾馆
六层1625号, Beijing 100080 (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,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, 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))。

classification.

(57) 摘要: 本发明公布了一种用于行人检测的协同式深度网络模型方法, 包括: 构建一种新的协同式多模型学习框架来完成行人检测中的分类过程; 利用人工神经网络来整合协同式模型中各子分类器的判决结果, 并借用机器学习的方法训练该网络, 能够更加有效地综合各分类器反馈的信息; 提出一种基于K-means聚类算法的重采样方法, 增强协同模型中各分类器的分类效果, 进而提升整体分类效果。本发明通过建立协同式深度网络模型, 将由聚类算法得到的不同类型的训练数据集用于并行地训练多个深度网络模型, 再通过人工神经网络将原始数据集在各个深度网络模型上的分类结果进行整合和综合分析, 以实现更加精准的样本分类。

一种用于行人检测的协同式深度网络模型方法

技术领域

本发明属于信息技术领域，涉及数字文件内容保护技术，尤其涉及一种用于行人检测的协同式深度网络模型方法。

背景技术

无论是自然科学研究还是社会科学研究，人类自身一直是其中最重要的关注对象，因此，利用计算机分析人类活动一直是计算机视觉领域中最热门的研究课题。而人类动作识别和事件检测等智能化行为的分析都需要以快速准确的行人检测为前提，因而行人检测技术的研究日益受到学术界和工业界的广泛关注，对现代视频监控、智能机器人和无人驾驶技术的发展均具有深远意义。一般来说，行人检测是指判断给定图像或视频帧中是否包含行人，如果包含，并标出行人位置的过程，因而可以分为样本的分类和定位两大部分。其中，快速准确的样本分类是行人检测技术的前提和关键。

现有的行人检测技术已较为成熟，其基本框架来源于 HOG+SVM 模型。其中，行人样本的分类大致分为样本收集（候选框提取）、预处理、特征提取、分类器训练和测试五个主要步骤。其中，特征提取和分类器训练是影响检测性能的关键。按照分类过程所使用分类器数目的多少，行人检测算法可以分为单分类器模型算法和多分类器模型算法。在第一类中，基于人工特征的检测算法和基于深度学习的检测算法占据主流。前者先用人工定义的方法从图像中提取特征，包括 HOG 特征、LBP 特征、ACF 等，再送入常用分类器中进行训练，最后利用训练好的模型去区分行人和非行人样本。后来，深度学习算法的出现大大提高了特征提取的准确性和分类器的检测性能，但仍然受限于单个分类器在特征学习能力上的局限性，其结果仍未达到理想水平。第二类算法成功突破了单分类器模型的瓶颈，它们使用多个分类器去学习样本特征，并将各分类结果整合在一起，实现了更加精准的分类效果，如常用的基于部件模型的分类算法、级联模型和集成模型等。基于部件的模型往往从样本的各个部分出发，提取局部特征，训练局部分类器，最后再整合分析，对于遮挡较为严重的样本能够较好地学习到有用的轮廓特征，从而大大提升了检测性能；级联模型则使用 Boosting 的思想，将多个分类器顺序排列，然后用不同的样本逐一训练这些分类器，使得它们具备不同的分类能力，其中后一分类器的设计需依据前一分类器的分类结果，最终级联所有的弱分类器形成一个强分类器，从分类器互补的角度改进了检测算法的性能；集成模型则是多分类器模型中较为罕

见的一类，在行人检测领域尚未被人关注和使用，它采用并行的方式来整合多个分类器，让所有子分类器共同完成最终的决策，这与人类社会协同合作的决策方式极其相似。目前在行人检测领域成功得到验证和应用的集成模型仅有集成 CNN（Convolutional Neural Network，卷积神经网络）模型，它并行地训练多个 CNN 模型，然后将每个样本输入各 CNN 模型，最后取输出的得分中最大、最小或平均值作为该样本的最终分类结果。基于集成 CNN 模型的分类型算法具体过程包括：

- 1) 准备训练数据集，用滑动窗口法从原始数据集中提取多尺度的行人和非行人候选框；
- 2) 将每个样本依次送入 CNN 进行卷积和 pooling 计算，得到特征图；
- 3) 采用不同的 dropout 比例设置全连接层，然后将从原始样本中提取的特征图送入训练不同的 CNN 模型；
- 4) 将测试数据集依次送入训练好的模型中进行行人样本的分类，每个样本均得到若干检测得分，计算得分的最大、最小或平均值作为最终的判别得分。

可见，上述原始基于集成 CNN 模型的分类型算法仅仅集成 CNN 模型。该模型比基于部件的模型和级联模型更具科学性，可以解决行人形变和遮挡问题，但是，它在整合方式上存在弊端，不具备一般性和理论依据，样本分类的精准性不足。

发明内容

为了克服上述现有技术的不足，本发明提供一种用于行人检测的协同式深度网络模型方法（Collaborative Deep Network，简称 CDN），主要针对行人检测技术中的分类过程，设计一种协同式深度网络模型，将由 K-means 聚类算法得到的不同类型的训练数据集用于并行地训练多个深度网络，再通过人工神经网络将原始数据集在这些深度网络上的分类结果进行整合和综合分析，以实现更加精准的样本分类，可用于行人检测中行人样本分类。

本发明的原理是：本发明对基于集成 CNN 模型的分类型算法进行改进，基于集成 CNN 模型的分类型算法的思想是利用全连接层中 dropout 的结点的不同来构建多个不同的 CNN 网络，用样本训练每个 CNN 网络，最终对每个网络的输出取最大、最小和平均值，进而完成分类。本发明借鉴了人类社会分工协作的工作模式，先并行地训练多个不同的深度网络模型作为成员，然后训练一个类似决策者的人工神经网络去学习数据集中的每个样本在各个子分类器上的分类结果信息，使其学会综合分析这些信息并得到更加准确的结论。同时，为了增强团队每个成员的能力，即各个子分类器的分类效果，进而提升整个分类模型的分类型能力，本发明还提出了一种基于 K-means 聚类算法的重采样方法，先将原始数据集中提取出来的候选样

本框按照某种特征进行聚类，得到不同类别的行人样本和非行人样本，再送入不同的检测模型进行学习，以使得分类器能够学到更加集中的样本特性。需要注意的是，本发明实际上提供了一种新的集成模型框架，使用多种深度网络模型，其中使用的深度网络模型本身效果越好，其集成后的效果也越显著。总之，本发明采用的协同式深度网络模型方法能够集成多个不同类型的深度网络模型，并且使用人工神经网络取代现有的原始算法中计算最大、最小或平均值的方法来整合不同深度网络的分类结果，为了提升每个子深度网络的分类能力，还设计了基于 K-means 聚类算法的重采样技术，从而达到实现更加准确的行人分类的效果。

本发明提供的技术方案是：

一种用于行人检测的协同式深度网络模型方法（简称 CDN），通过建立协同式深度网络模型，将由聚类算法得到的不同类型的训练数据集用于并行地训练多个深度网络模型，再通过人工神经网络将原始数据集在各个深度网络模型上的分类结果进行整合和综合分析，以实现更加精准的样本分类；包括如下步骤：

- 1) 采用基于 K-means 聚类算法的重采样方法，将原始训练样本数据集按照不同特征划分为不同的子样本集；
- 2) 选取多个深度网络模型分别作为子分类器，利用上述子样本集并行地训练所述多个深度网络模型，得到多个训练好的子分类器；
- 3) 将原始训练样本数据集同时送入所述多个训练好的子分类器中，得到检测得分，将所述检测得分连接成检测得分向量，利用所述检测得分向量训练一个人工神经网络，得到训练好的协同式深度网络模型；
- 4) 将测试数据集输入到所述训练好的协同式深度网络模型中对行人样本进行分类，得到行人样本分类。

针对上述协同式深度网络模型方法，进一步地，步骤 1) 所述基于 K-means 聚类算法的重采样方法，具体采用 K-means 聚类算法自动分割原始训练样本数据集，通过重采样得到多个具备不同特性的训练子样本集，使得训练子样本集之间差异较大而训练子样本集内差异较小；包括如下步骤：

- 11) 对原始训练样本数据集 D 中的每一个样本，提取得到特征向量 $X = \{x_i\}_{i=1}^n$ ， n 为样本总数；
- 12) 随机选择 k 个样本的特征向量作为聚类中心，记为 $\{c_j\}_{j=1}^k$ ；

13) 通过式 1 计算得到每个特征向量与每个聚类中心的距离 d_{ij} :

$$d_{ij} = \|x_i - c_j\|_2 \quad (\text{式 1})$$

式 1 中, x_i 为原始训练样本集 D 中的每个样本的特征向量; c_j 为每个聚类中心;

14) 对于每个特征向量 x_i , 依据其与各个聚类中心的距离大小, 将该特征向量划分至彼此间距离最近的中心向量所在的类;

15) 通过式 2 更新每类的中心向量:

$$c_j = \frac{1}{|C_j|} \sum_{i \in C_j} x_i \quad (\text{式 2})$$

$$C_j = \{i | \forall m \neq j, d_{im} > d_{ij}\}$$

其中, C_j 表示每个类别中所包含的所有样本标号;

16) 当 C_j 不再变化时, 停止聚类过程, 得到不同的子样本集; 否则返回步骤 13)。

针对上述协同式深度网络模型方法, 进一步地, 所述特征向量为三通道特征, 每个样本对应三个特征向量, 将三个通道中的每一个特征向量分别对原始训练样本进行聚类; 将步骤 12) 所述分类个数 k 的模型初始化值设为 2, 代表对应样本包含行人和非行人两类的属性。

针对上述协同式深度网络模型方法, 进一步地, 步骤 2) 作为子分类器的深度网络模型包括基础深度网络模型、强深度学习检测器和用弱训练样本训练的强深度学习检测器。其中, 基础深度网络模型包括卷积神经网络模型和感知器模型; 所述强深度学习检测器包括联合深度学习模型。

针对上述协同式深度网络模型方法, 进一步地, 步骤 3) 所述人工神经网络的前馈模型为式 3:

$$net_j = \sum_{i=1}^{i=n} w_{ij} x_i + b_j, j = 1, 2, \dots, m$$

$$y_j = f(net_j), j = 1, 2, \dots, m \quad (\text{式 3})$$

其中, x_i 代表输入层第 i 个结点的值, w_{ij} 代表输入层第 i 个结点到输出层第 j 个结点的连接权重, n 是输入层结点个数, b_j 代表输出层第 j 个结点的偏置;

可采用反向传播算法训练所述人工神经网络。

针对上述协同式深度网络模型方法, 进一步地, 在本发明实施例中, 步骤 2) 选取联合

深度学习模型为子分类器，原始训练样本数据集采用 Caltech 行人数据库和 ETH 行人数据库中的原始图像，对该子分类器的训练过程分为以下几步：

第一步，先将原始图像变换到 YUV 颜色空间，提取得到三通道特征；

第二步，将得到的三通道特征输入联合深度学习模型，对输入在三通道特征进行两次卷积变换和一次 pooling 操作，得到多个部件检测图；

第三步，通过式 4 对部件检测图进行形变处理，计算得到各部件的得分 s_p ：

$$B_p = M_p + \sum_{n=1}^N c_{np} D_{np}$$

$$s_p = \max_{(x,y)} b_p^{(x,y)}$$
(式 4)

其中， M_p 代表第 p 个部件检测图； D_{np} 和 c_{np} 分别代表第 p 个部件对应的第 n 个形变图及其权重； $b_p^{(x,y)}$ 对应 B_p 中 (x, y) 位置上的元素；

第四步，通过式 5 将所有部件的得分构成部件得分向量 s ：

$$s = (s_p)_{p=1}^{20} \quad (\text{式 5})$$

再用所述部件得分向量训练视觉推理和分类网络，即完成上述子分类器的训练，得到训练好的子分类器。

上述训练过程中，第一步所述三通道特征，其中，第一通道特征对应原始图像的 Y 通道特征；将第二通道特征分为四部分：左上角、右上角、左下角、右下角，左上角、右上角、左下角分别对应尺度缩小为原来一半的原始图片的 Y、U、V 三个通道的特征图，右下角用 0 填充；将第三通道特征也分为四部分，反映原始图片的边缘信息，左上角、右上角、左下角分别是由原始图片的 Y、U、V 三个通道的特征图经过 Sobel 算子变换并进行尺寸放缩后得到的边缘图，右下角由上述三个边缘图中每个位置取幅值最大的像素值构成。在本发明实施例中，第二步使用二十个卷积核，计算得到二十个部件检测图，为部件的特征图。第三步中采用的形变图为与部件检测图尺寸相同的二维矩阵，每个矩阵中的值均在 0 到 255 之间。

与现有技术相比，本发明的有益效果是：

本发明针对计算机视觉中的行人检测领域，现有行人检测技术中使用的分类算法所存在的局限性，提出了一种新的协同式深度网络模型算法。先由 K-means 聚类算法从原始数据集

中分割出不同类型的训练子数据集，再用这些数据集并行地训练多个深度网络模型，最后通过人工神经网络将全部样本在这些训练好的深度网络模型上的分类结果整合在一起，综合分析以实现更加精准的样本分类。本发明在多个实验数据集上相比其他算法都有更好的效果。

与现有技术相比，本发明的核心优点体现在以下方面：

（一）构建了一种新的协同式多模型学习框架来完成行人检测中的分类过程，框架内并行地训练若干不同的深度网络模型，最后整合各模型的分类结果来共同做出决策，有效避免了单一分类器在特征提取和特征学习上的局限性。

构建一种新的协同式多模型学习框架是现有的行人检测技术中基本没有考虑过的，本发明提供的多个深度网络模型协作的分类模式可以有效弥补单一深度模型在特征提取和特征学习上的局限性，尤其是对于形变和遮挡较为严重的行人样本，充分发挥每个分类器的特长，消除检测误差，从而实现更加准确的行人检测。

（二）提出利用人工神经网络来整合协同式模型中各子分类器的判决结果，并借用机器学习的方法训练该网络，能够更加有效地综合各分类器反馈的信息。

利用人工神经网络来整合协同式模型中的各个深度网络的策略比现有的利用平均或取最大值等人工规则更贴近人脑决策方式，更具科学性，能够实现更好的检测效果。

（三）提出了一种基于 K-means 聚类算法的重采样技术，先将原始数据集中提取出来的候选样本框按照某种特征进行聚类，得到不同类型的行人样本和非行人样本，用以训练不同的检测模型，利用 K-means 聚类算法进行样本的重采样可以使每个子分类器能学到不同且更加集中的行人特征，以使得各分类器均能够学到更加专一的样本特征，增强某一类样本的分类能力，增强协同模型中各分类器的分类效果，进而提升整体分类效果。

附图说明

图 1 是本发明提供的协同式深度网络模型方法的流程框图。

图 2 本发明实施例中以联合深度模型为子分类器的训练过程的流程框图。

具体实施方式

下面结合附图，通过实施例进一步描述本发明，但不以任何方式限制本发明的范围。

本发明提出了一种用于行人检测的协同式深度网络模型算法（简称 CDN），该算法是一种样本分类方法，不包含行人检测过程中的候选框提取和行人定位过程；主要针对行人检测技术中的分类过程，设计一种协同式深度网络模型，将由 K-means 聚类算法得到的不同类型

的训练数据集用于并行地训练多个深度网络，再通过人工神经网络将原始数据集在这些深度网络上的分类结果进行整合和综合分析，以实现更加精准的样本分类，可用于行人检测中行人样本分类；图 1 是本发明提供的协同式深度网络模型方法的流程框图，主要包括如下步骤：

1) 采用基于 K-means 聚类算法的重采样技术准备训练样本数据集，将原始样本数据集按照某种特征的不同划分为不同的子样本集；

2) 选取多个深度网络模型作为子分类器，并利用原始训练样本集和重采样的上述子样本集分别并行地训练它们；

其中每个深度网络模型均采用原模型本身的训练方法进行训练；

3) 将原始样本数据集同时送入多个训练好的子分类器中，得到检测得分，然后将这些得分连接成向量去训练一个神经网络，最终得到训练好的协同式深度网络模型；

4) 将测试数据集输入训练好的模型中进行行人样本的分类。

步骤 1) 中，使用基于 K-means 聚类算法的重采样技术准备训练样本数据集，即采用 K-means 聚类算法自动分割原始样本数据集，以重采样得到多个具备不同特性的训练子样本集，使得不同样本集间差异较大而样本集内差异较小，进而使得多个深度网络能够学习到不同但更加集中的某类特征，并增强区分特定样本的能力。具体步骤如下：

11) 对原始训练样本集 D 中的每一个样本，提取得到特征向量 $X = \{x_i\}_{i=1}^n$ ， n 为样本总数；

12) 随机选择 k 个样本的特征向量作为聚类中心，记为 $\{c_j\}_{j=1}^k$ ；

13) 通过式 1 计算得到每个特征向量与每个聚类中心的距离 d_{ij} ：

$$d_{ij} = \|x_i - c_j\|_2 \quad (\text{式 1})$$

式 1 中， x_i 为原始训练样本集 D 中的每个样本的特征向量； c_j 为每个聚类中心；

14) 对于每个特征向量 x_i ，依据其与各个聚类中心的距离 d_{ij} 的大小，将该特征向量划分至彼此间距离最近的中心向量所在的类；

15) 通过式 2 更新每类的中心向量：

$$c_j = \frac{1}{|C_j|} \sum_{i \in C_j} x_i \quad (\text{式 2})$$

$$C_j = \{i \mid \forall m \neq j, d_{im} > d_{ij}\}$$

其中， C_j 表示每个类别中所包含的所有样本标号；

16) 如果 C_j 不再变化, 则停止聚类过程, 得到不同的子样本集; 否则返回步骤 13)。

由于这里提取的特征向量为三通道特征, 故每个样本应对应三个特征向量, 且在重采样过程中, 需依据这三个通道中的每一个分别对原始训练样本进行聚类。另外, 考虑到 K-means 算法需要事先规定分类个数 k , 本模型初始化 k 等于 2, 以对应样本所具有的包含行人和非行人两类的属性。

步骤 2) 中, 使用上述聚类得到的训练子样本集和原始数据集分别去训练多个不同的深度网络模型, 每个模型的训练过程采用多线程并行计算。这里可以使用的深度网络模型包括以下三类:

2A) 基础深度网络模型, 如 CNN、感知器模型等, 这些模型已经具有较好的特征学习和分类能力, 但不足以完成存在大量行人形变和遮挡的复杂场景下的行人检测任务, 故需要互相配合才能更好地避免自身存在的检测误差;

2B) 强深度学习检测器, 如上述联合深度学习模型等, 这些模型比基础深度网络模型更能准确快速地检测出图像中的行人, 且对于复杂场景已有较好的应对措施, 在 CDN 中加入强深度学习检测器可以有效保证整体模型的检测效果, 并进一步提升强深度学习检测器的检测性能;

2C) 用弱训练样本训练的强深度学习检测器, 这一类模型较为特殊, 它们具有较强的分类能力, 但被用某一类的训练样本训练, 进而对某类样本有极致的识别能力, 通过将具备检测不同类型样本能力的分类器进行整合, CDN 模型会比普通的多模型融合具有更全面的检测能力。

步骤 3) 中, 在得到多个训练好的深度网络模型后, 将原始训练数据集中的各个样本同时送入这些深度模型进行分类识别, 并得到若干检测得分 (每个深度网络模型均得出分类检测的得分), 将这些输出得分组成向量作为每个样本的观测信息去训练一个人工神经网络, 以得到一个协同式分类模型, 即协同式深度网络模型, 该模型中内嵌多个不同的深度网络子分类器, 可在综合学习各样本的初步分类信息以后能够有效利用这些信息去完成更加准确的行人检测任务。人工神经网络的前馈模型为式 3:

$$\begin{aligned} net_j &= \sum_{i=1}^{i=n} w_{ij} x_i + b_j, j = 1, 2, \dots, m \\ y_j &= f(net_j), j = 1, 2, \dots, m \end{aligned} \quad (\text{式 3})$$

其中, x_i 代表输入层第 i 个结点的值, w_{ij} 代表输入层第 i 个结点到输出层第 j 个结点的连

接权重， n 是输入层结点个数， b_j 代表输出层第 j 个结点的偏置。

训练人工神经网络可采用 BP (Back Propagation, 即反向传播) 算法。

协同式分类模型是在综合学习各样本的初步分类信息以后得到的，能够有效利用各样本的初步分类信息去完成更加准确的行人检测任务。在线测试时，只需将测试数据集输入训练好的协同式分类模型中进行分类，即获得测试数据集的行人样本分类。

为了便于实验验证，以下实施例采用联合深度学习模型 (以下简称 UDN) 作为本发明的协同式深度网络模型选取的子分类器，在 Caltech 行人数据库和 ETH 行人数据库上 (原始数据库作为原始训练样本数据集，先将原始数据库中的图像经过预处理得到矩形候选框图像后用于训练) 分别做了测试。UDN 模型成功将特征提取、形变处理、遮挡处理和分类四个环节融入到一个 CNN 模型当中，有效地解决了行人形变和遮挡等问题。

UDN 作为子分类器的训练过程的流程如图 2 所示。其中，输入图片定义为三通道的、大小为 84×28 的图像数据，经由 64 个 $9 \times 9 \times 3$ 的卷积核做卷积操作，再经过 4×4 的 pooling 操作，得到 64 个 19×5 的特征图，这些特征图再输入第二卷积层，经过 20 个设计好的处理形变的卷积核得到 20 个部件检测图，随后利用形变层计算部件检测得分，最后送入视觉推理和分类模型进行类别的估计。具体训练过程分为以下几步：

第一步，先将图像变换到 YUV 颜色空间，然后提取三通道特征，其中，第一通道特征对应原始图片的 Y 通道特征；第二通道特征分为 4 块，左上角、右上角、左下角三块分别对应原始图片的 Y、U、V 三个通道的特征图其尺度缩小为原来一半的结果，右下角那一块用 0 填充；同样，第三通道特征也分为四个部分，反映的是原始图片的边缘信息，其左上角、右上角、左下角分别是由原始图片的 Y、U、V 三个通道的特征图经过 Sobel 算子变换并进行尺寸放缩后得到的边缘图，右下角则由上述三个边缘图中每个位置取幅值最大的像素值构成。

第二步，对输入的三通道特征进行两次卷积变换和一次 pooling 操作，得到 20 个部件检测图；

在训练过程中，本实施例使用 20 个不同的卷积核，计算得到 20 个部件检测，为人体部件的特征图。

第三步，通过式 4 对部件检测图进行形变处理，计算得到各部件得分 s_p ：

$$B_p = M_p + \sum_{n=1}^N c_{np} D_{np}$$

$$s_p = \max_{(x,y)} b_p^{(x,y)}$$
(式 4)

其中, M_p 代表第 p 个部件检测图; D_{np} 和 c_{np} 分别代表第 p 个部件对应的第 n 个形变图及其权重; $b_p^{(x,y)}$ 对应 B_p 中 (x, y) 位置上的元素;

本实施例中设计的形变图为与部件检测图尺寸相同的二维矩阵, 每个矩阵中的值均在 0 到 255 之间, 包括四种矩阵, 第一种和第三种均将矩阵划分为从左至右排列的六个长条形区域, 每个区域的值相同, 且左边区域的值始终大于右边区域, 区别在于第一种的区域间的值的跳跃较小, 第三种区域间的值的跳跃较大; 第二种和第四种则将矩阵划分为从上至下排列的六个长条形区域, 每个区域中的值均相同, 区别在于第二种满足上边区域的值要大于下边区域的值, 而第四种则满足从上至下区域的值先递增后递减的变化规律; 这些形变图中具体参数值的预设可以不固定, 通过训练过程来优化这些参数。

第四步, 所有部件的得分构成部件得分向量 s , 如式 5; 再用部件得分向量训练视觉推理和分类网络, 训练方法采用标准 BP 算法; 即完成上述子分类器的训练, 得到训练好的子分类器。

$$s = (s_p)_{p=1}^{20} \quad (\text{式 } 5)$$

其中, 在联合深度模型中, 视觉推理和分类网络是一种人工神经网络, 网络中每个隐藏层接收的输入信息不仅来自于上层还来自于上上层。

实验结果表明, CDN 比行人检测领域其他较为先进的算法具有更好的性能和检测效果, 可以有效提升单一模型的分类能力; 其中, 行人检测领域其他较为先进的算法包括 HOG (Histogram of Oriented Gradient, 即梯度直方图)、HOGLBP (Histogram of Oriented Gradient and Local Binary Pattern, 即梯度直方图和局部二值化模式)、DPM (Deformable Part Models, 即形变部件模型)、DDM (Discriminative Deep Model, 即辨别深度模型)、ICF (Integral Channel Features, 即积分通道特征)、CNN (Convolutional Neural Network, 即卷积神经网络)、ACF (Aggregated Channel Features, 即聚合通道特征)、UDN (United Deep Model, 即联合深度模型)。对比结果如表 1 所示:

表 1 不同行人检测模型实验结果统计表

算法模型	Log 平均缺失率	
	Caltech 测试集	ETH 数据集
HOG	68%	64%
HOGLBP	68%	55%

DPM	63%	51%
Discriminative Deep Model	61%	47%
ICF	48%	50%
CNN	46%	—
ACF	43%	50%
UDN	42.54%	45.47%
CDN	40.49%	44.11%

表 2 不同设计的 CDN 模型在 Caltech 数据集上实验结果统计表

聚类特征设计	Log 平均缺失率
CDN-RGB-based	41.70%
CDN-ACF-based	40.51%
CDN	40.49%
聚类数目设计	Log 平均缺失率
CDN-6, 6, 6	42.40%
CDN-5, 5, 2	42.01%
CDN	40.49%
深度网络设计	Log 平均缺失率
CDN-UDN+CNN	42.07%
CDN-UDNs	41.08%
CDN	40.49%
集成方式设计	Log 平均缺失率
CDN-Max	42.71%
CDN-3-layer Deep Neural Network	41.99%
CDN-Mean	41.88%
CDN	40.49%

表 2 则给出了不同设计的 CDN 在 Caltech 数据集上的检测效果，可以发现，CDN 若基于三通道特征进行聚类且聚类数目均为 2，同时集成用聚类得到的子训练集和原始训练集分别训练的联合深度学习模型，并且采用人工神经网络进行结果的整合，会获得最好的检测性

能。

需要注意的是，公布实施例的目的在于帮助进一步理解本发明，但是本领域的技术人员可以理解：在不脱离本发明及所附权利要求的精神和范围内，各种替换和修改都是可能的。因此，本发明不应局限于实施例所公开的内容，本发明要求保护的范围以权利要求书界定的范围为准。

权 利 要 求 书

1、一种用于行人检测的协同式深度网络模型方法，通过建立协同式深度网络模型，将由聚类算法得到的不同类型的训练数据集用于并行地训练多个深度网络模型，再通过人工神经网络将原始数据集在各个深度网络模型上的分类结果进行整合和综合分析，以实现更加精准的样本分类；包括如下步骤：

1) 采用基于 K-means 聚类算法的重采样方法，将原始训练样本数据集按照不同特征划分为不同的子样本集；

2) 选取多个深度网络模型分别作为子分类器，利用上述子样本集并行地训练所述多个深度网络模型，得到多个训练好的子分类器；

3) 将原始训练样本数据集同时送入所述多个训练好的子分类器中，得到检测得分，将所述检测得分连接成检测得分向量，利用所述检测得分向量训练一个人工神经网络，得到训练好的协同式深度网络模型；

4) 将测试数据集输入到所述训练好的协同式深度网络模型中对行人样本进行分类，得到行人样本分类。

2、如权利要求 1 所述协同式深度网络模型方法，其特征是，步骤 1) 所述基于 K-means 聚类算法的重采样方法，具体采用 K-means 聚类算法自动分割原始训练样本数据集，通过重采样得到多个具备不同特性的训练子样本集，使得训练子样本集之间差异较大而训练子样本集内差异较小；包括如下步骤：

11) 对原始训练样本数据集 D 中的每一个样本，提取得到特征向量 $X=\{x_i\}_{i=1}^n$ ， n 为样本总数；

12) 随机选择 k 个样本的特征向量作为聚类中心，记为 $\{c_j\}_{j=1}^k$ ；

13) 通过式 1 计算得到每个特征向量与每个聚类中心的距离 d_{ij} ：

$$d_{ij} = \|x_i - c_j\|_2 \quad (\text{式 1})$$

式 1 中， x_i 为原始训练样本集 D 中的每个样本的特征向量； c_j 为每个聚类中心；

14) 对于每个特征向量 x_i ，依据其与各个聚类中心的距离大小，将该特征向量划分至彼此间距离最近的中心向量所在的类；

15) 通过式 2 更新每类的中心向量：

$$c_j = \frac{1}{|C_j|} \sum_{i \in C_j} x_i \quad (\text{式 } 2)$$

$$C_j = \{i | \forall m \neq j, d_{im} > d_{ij}\}$$

其中, C_j 表示每个类别中所包含的所有样本标号;

16) 当 C_j 不再变化时, 停止聚类过程, 得到不同的子样本集; 否则返回步骤 13)。

3、如权利要求 2 所述协同式深度网络模型方法, 其特征是, 所述特征向量为三通道特征, 每个样本对应三个特征向量, 将三个通道中的每一个特征向量分别对原始训练样本进行聚类; 将步骤 12) 所述分类个数 k 的模型初始化值设为 2, 代表对应样本包含行人和非行人两类的属性。

4、如权利要求 1 所述协同式深度网络模型方法, 其特征是, 步骤 2) 作为子分类器的深度网络模型包括基础深度网络模型、强深度学习检测器和用弱训练样本训练的强深度学习检测器。

5、如权利要求 4 所述协同式深度网络模型方法, 其特征是, 所述基础深度网络模型包括卷积神经网络模型和感知器模型; 所述强深度学习检测器包括联合深度学习模型。

6、如权利要求 1 所述协同式深度网络模型方法, 其特征是, 步骤 3) 所述人工神经网络的前馈模型为式 3:

$$net_j = \sum_{i=1}^{i=n} w_{ij} x_i + b_j, j = 1, 2, \dots, m$$

$$y_j = f(net_j), j = 1, 2, \dots, m \quad (\text{式 } 3)$$

其中, x_i 代表输入层第 i 个结点的值, w_{ij} 代表输入层第 i 个结点到输出层第 j 个结点的连接权重, n 是输入层结点个数, b_j 代表输出层第 j 个结点的偏置;

可采用反向传播算法训练所述人工神经网络。

7、如权利要求 1 所述协同式深度网络模型方法, 其特征是, 步骤 2) 选取联合深度学习模型为子分类器, 原始训练样本数据集采用 Caltech 行人数据库和 ETH 行人数据库中的原始图像, 对该子分类器的训练过程分为以下几步:

第一步, 先将原始图像变换到 YUV 颜色空间, 提取得到三通道特征;

第二步, 将得到的三通道特征输入联合深度学习模型, 对输入在三通道特征进行两次卷积变换和一次 pooling 操作, 得到多个部件检测图;

第三步, 通过式 4 对部件检测图进行形变处理, 计算得到各部件的得分 s_p :

$$\begin{aligned} B_p &= M_p + \sum_{n=1}^N c_{np} D_{np} \\ s_p &= \max_{(x,y)} b_p^{(x,y)} \end{aligned} \quad (\text{式 4})$$

其中, M_p 代表第 p 个部件检测图; D_{np} 和 c_{np} 分别代表第 p 个部件对应的第 n 个形变图及其权重; $b_p^{(x,y)}$ 对应 B_p 中 (x, y) 位置上的元素;

第四步, 通过式 5 将所有部件的得分构成部件得分向量 s :

$$s = (s_p)_{p=1}^{20} \quad (\text{式 5})$$

再用所述部件得分向量训练视觉推理和分类网络, 即完成上述子分类器的训练, 得到训练好的子分类器。

8、如权利要求 7 所述协同式深度网络模型方法, 其特征是, 第一步所述三通道特征中, 第一通道特征对应原始图像的 Y 通道特征; 将第二通道特征分为四部分: 左上角、右上角、左下角、右下角, 左上角、右上角、左下角分别对应尺度缩小为原来一半的原始图片的 Y、U、V 三个通道的特征图, 右下角用 0 填充; 将第三通道特征也分为四部分, 反映原始图片的边缘信息, 左上角、右上角、左下角分别是由原始图片的 Y、U、V 三个通道的特征图经过 Sobel 算子变换并进行尺寸放缩后得到的边缘图, 右下角由上述三个边缘图中每个位置取幅值最大的像素值构成。

9、如权利要求 7 所述协同式深度网络模型方法, 其特征是, 第二步使用二十个卷积核, 计算得到二十个部件检测图, 为部件的特征图。

10、如权利要求 7 所述协同式深度网络模型方法, 其特征是, 第三步中采用的形变图为与部件检测图尺寸相同的二维矩阵, 每个矩阵中的值均在 0 到 255 之间。

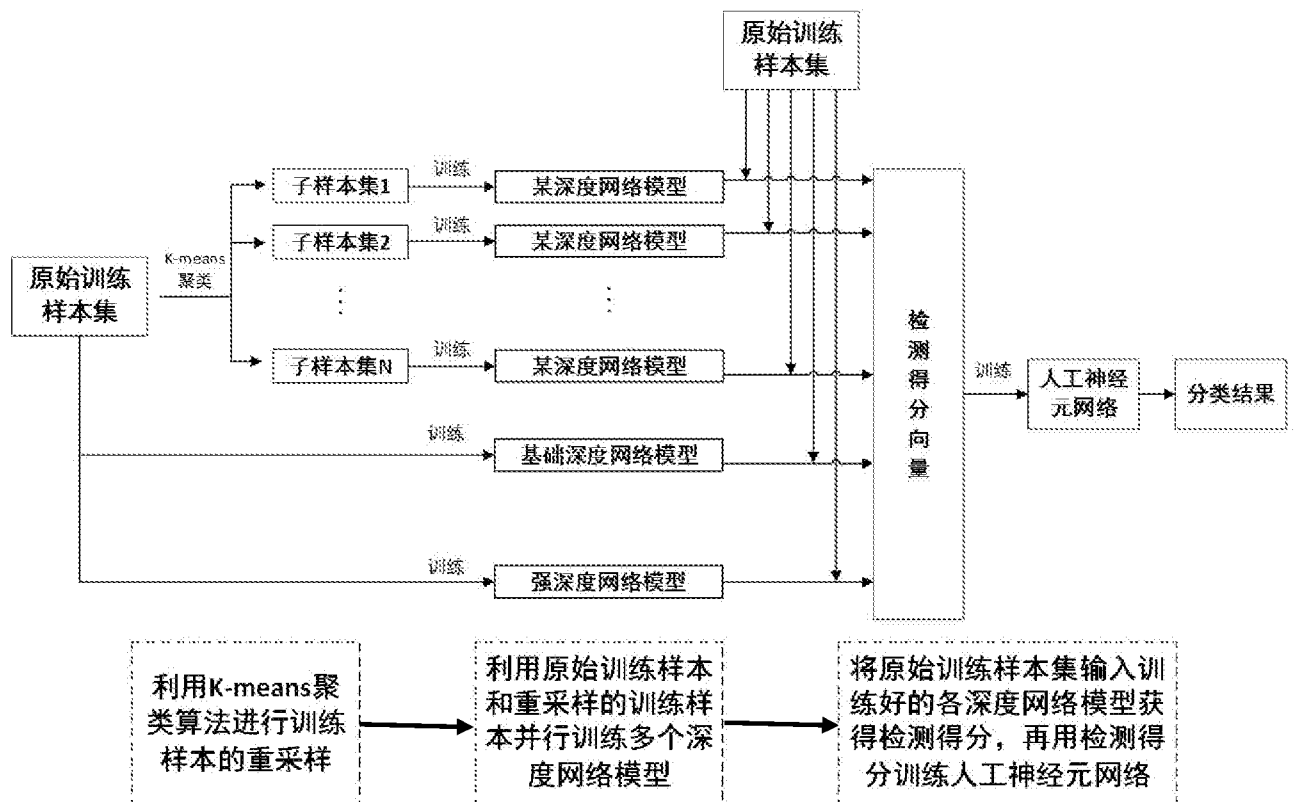


图 1

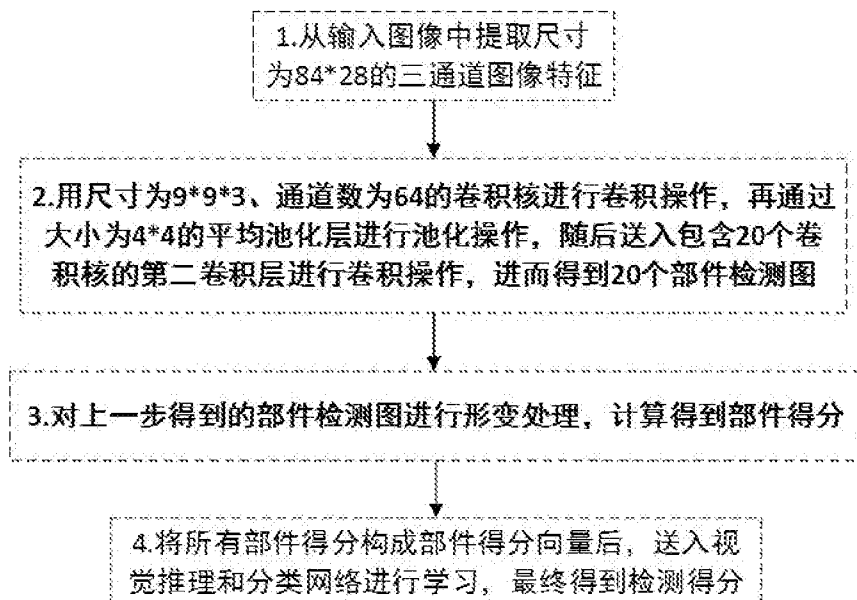


图 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/094016

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/62 (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

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)

CPRSABS, CNTXT, CNKI, DWPI, IEEE: 检测, 网络, 神经, 分类, 聚类, 均值, 样本, 子分类器, detect, network, neural, classify, cluster, means, sample, sub, classifier

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 106650806 A (PEKING UNIVERSITY SHENZHEN GRADUATE SCHOOL), 10 May 2017 (10.05.2017), claims 1-10	1-10
A	CN 1908960 A (INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES), 07 February 2007 (07.02.2007), entire document	1-10
A	CN 103489004 A (SOUTH CHINA UNIVERSITY OF TECHNOLOGY), 01 January 2014 (01.01.2014), entire document	1-10
A	US 8644624 B2 (SAMSUNG ELECTRONICS CO., LTD.), 04 February 2014 (04.02.2014), entire document	1-10

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 22 October 2017	Date of mailing of the international search report 03 November 2017
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer WANG, Xuelian Telephone No. (86-10) 62411747

International application No.
PCT/CN2017/094016

Form PCT/ISA/210 (patent family annex) (July 2009)

国际检索报告

国际申请号

PCT/CN2017/094016

A. 主题的分类

G06K 9/62 (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

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))

CPRSABS, CNTXT, CNKI, DWPI, IEEE: 检测, 网络, 神经, 分类, 聚类, 均值, 样本, 子分类器, detect, network, neural, classify, cluster, means, sample, sub, classifier

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 106650806 A (北京大学深圳研究生院) 2017年 5月 10日 (2017 - 05 - 10) 权利要求1-10	1-10
A	CN 1908960 A (中国科学院计算技术研究所) 2007年 2月 7日 (2007 - 02 - 07) 全文	1-10
A	CN 103489004 A (华南理工大学) 2014年 1月 1日 (2014 - 01 - 01) 全文	1-10
A	US 8644624 B2 (SAMSUNG ELECTRONICS CO LTD) 2014年 2月 4日 (2014 - 02 - 04) 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

2017年 10月 22日

国际检索报告邮寄日期

2017年 11月 3日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局 (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

受权官员

王雪莲

电话号码 (86-10) 62411747

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2017/094016

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	106650806	A	2017年 5月 10日	无			
CN	1908960	A	2007年 2月 7日	CN	100426314	C	2008年 10月 15日
CN	103489004	A	2014年 1月 1日	无			
US	8644624	B2	2014年 2月 4日	US	2011026840	A1	2011年 2月 3日

表 PCT/ISA/210 (同族专利附件) (2009年7月)