

(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
2022 年 1 月 20 日 (20.01.2022)



(10) 国际公布号
WO 2022/011892 A1

(51) 国际专利分类号:
G06K 9/62 (2006.01)

(21) 国际申请号: PCT/CN2020/125972

(22) 国际申请日: 2020 年 11 月 2 日 (02.11.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010681178.2 2020 年 7 月 15 日 (15.07.2020) CN

(71) 申请人: 北京市商汤科技开发有限公司(BEIJING SENSETIME TECHNOLOGY DEVELOPMENT CO., LTD.) [CN/CN]; 中国北京市海淀区北四环西路 58 号 11 层 1101-1117 室, Beijing 100080 (CN)。

(72) 发明人: 窦浩轩(DOU, Haoxuan); 中国北京市海淀区北四环西路 58 号 11 层 1101-1117 室, Beijing 100080 (CN)。王意如(WANG, Yiru); 中国北京市海淀区北四环西路 58 号 11 层 1101-1117 室, Beijing 100080 (CN)。甘伟豪(GAN, Weihao); 中国北京市海淀区北四环西路 58 号 11 层 1101-1117 室, Beijing 100080 (CN)。路少卿(LU, Shaoqing); 中国北京市海淀区北四环西路 58 号 11 层 1101-1117 室, Beijing 100080 (CN)。王伟(WU, Wei); 中国北京市海淀区北四环西路 58 号 11 层 1101-1117 室, Beijing 100080 (CN)。闫俊杰(YAN, Junjie); 中国北京市海淀区北四环西路 58 号 11 层 1101-1117 室, Beijing 100080 (CN)。

(74) 代理人: 北京派特恩知识产权代理有限公司(CHINA PAT INTELLECTUAL PROPERTY

(54) Title: NETWORK TRAINING METHOD AND APPARATUS, TARGET DETECTION METHOD AND APPARATUS, AND ELECTRONIC DEVICE

(54) 发明名称: 网络训练方法及装置、目标检测方法及装置和电子设备

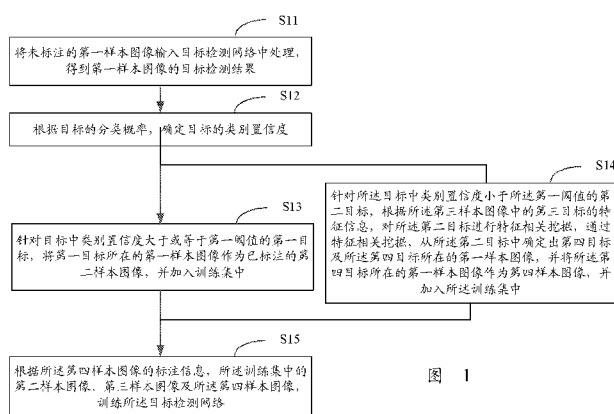


图 1

- S11 Input an unannotated first sample image into a target detection network for processing, so as to obtain a target detection result of the first sample image
- S12 Determine a category confidence of a target according to a classification probability of the target
- S13 For a first target having a category confidence greater than or equal to a first threshold, use the first sample image where the first target is located as an annotated second sample image, and add the annotated second sample image to a training set
- S14 For a second target having a category confidence less than the first threshold, perform feature-related mining on the second target according to feature information of a third target in a third sample image; determine, from the second target by means of feature-related mining, a fourth target and a first sample image where the fourth target is located; and use the first sample image where the fourth target is located as a fourth sample image, and add the fourth sample image into the training set
- S15 Train the target detection network according to annotation information of the fourth sample image, the second sample image, the third sample image, and the fourth sample image in the training set

(57) Abstract: Embodiments of the present disclosure relate to a network training method and apparatus, a target detection method and apparatus, and an electronic device. The network training method comprises: inputting an unannotated sample image into a target detection network to obtain a target detection result, the result comprising an image area, feature information, and classification probability of a target; determining a category confidence of the target according to the classification probability of the target; for a first target having a category confidence greater than or equal to a threshold, using a sample image where the first target is located as an

OFFICE); 中国北京市海淀区海淀南路21号中关村知识产权大厦B座2层, 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, IT, 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, 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))。

annotated image, and adding the annotated image into a training set; for a second target having a category confidence less than a first threshold, performing feature-related mining on the second target, determining a fourth target, and adding a sample image where the fourth target is located into the training set; and training the target detection network according to the training set.

(57) 摘要: 本公开实施例涉及一种网络训练方法及装置、目标检测方法及装置和电子设备。该网络训练方法包括: 将未标注的样本图像输入目标检测网络, 得到目标检测结果, 该结果包括目标的图像区域、特征信息及分类概率; 根据目标的分类概率, 确定目标的类别置信度; 针对类别置信度大于或等于阈值的第一目标, 将第一目标所在的样本图像作为已标注图像加入训练集; 针对类别置信度小于第一阈值的第二目标, 对第二目标进行特征相关挖掘, 确定出第四目标, 将其所在的样本图像加入训练集; 根据训练集训练目标检测网络。

网络训练方法及装置、目标检测方法及装置和电子设备

相关申请的交叉引用

本申请基于申请号为 202010681178.2、申请日为 2020 年 07 月 15 日的中国专利申请提出，并要求该中国专利申请的优先权，该中国专利申请的全部内容在此引入本申请作为参考。

技术领域

本公开涉及计算机技术领域，尤其涉及一种网络训练方法及装置、目标检测方法及装置和电子设备。

背景技术

计算机视觉是人工智能技术的重要方向，在计算机视觉处理中，通常需要对图像或视频中的目标（例如行人、物体等）进行检测。大规模长尾数据的目标检测在很多领域有重要应用，例如在城市监控中的异常物体检测，异常行为检测和突发事件报警等。然而，由于长尾数据的数据量巨大，以及严重的正负样本不均衡现象，即大部分数据图片为背景图，仅有小部分图片中含有可检测的目标，导致相关技术的目标检测方式对长尾数据的目标检测效果较差。

发明内容

本公开实施例提出了一种网络训练及目标检测技术方案。

根据本公开实施例的一方面，提供了一种网络训练方法，包括：

将未标注的第一样本图像输入目标检测网络中处理，得到所述第一样本图像的目标检测结果，所述目标检测结果包括所述第一样本图像中目标的图像区域、特征信息及分类概率；根据所述目标的分类概率，确定所述目标的类别置信度；针对所述目标中类别置信度大于或等于第一阈值的第二目标，将所述第一目标所在的第一样本图像作为已标注的第二样本图像，并加入训练集中，其中，所述第二样本图像的标注信息包括所述第一目标的图像区域及与所述第一目标的类别置信度对应的类别，所述训练集中包括已标注的第三样本图像；针对所述目标中类别置信度小于所述第一阈值的第二目标，根据所述第三样本图像中的第三目标的特征信息，对所述第二目标进行特征相关挖掘，通过特征相关挖掘，从所述第二目标中确定出第四目标及所述第四目标所在的第一样本图像，并将所述第四目标所在的第一样本图像作为第四样本图像，并加入所述训练集中；根据所述第四样本图像的标注信息，所述训练集中的第二样本图像、第三样本图像及所述第四样本图像，训练所述目标检测网络。

在一种可能的实现方式中，所述根据所述第四样本图像的标注信息，所述训练集中的第二样本图像、第三样本图像及所述第四样本图像，训练所述目标检测网络，包括：

根据所述训练集的正样本图像中目标的类别，分别确定从各个类别的正样本图像中采样的第一数量，所述正样本图像为图像中包括目标的样本图像；根据各个类别的正样本图像中采样的第一数量，对各个类别的正样本图像进行采样，得到多个第五样本图像；对所述训练集的负样本图像进行采样，得到多个第六样本图像，所述负样本图像为图像中不包括目标的样本图像；根据所述第五样本图像及所述第六样本图像，训练所述目标检测网络。

在一种可能的实现方式中，所述根据所述第三样本图像中的第三目标的特征信息，对所述第二目标进行特征相关挖掘，通过特征相关挖掘，从所述第二目标中确定出第四目标及所述第四目标所在的第一样本图像，包括：根据所述第二目标的分类概率，确定所述第

二目标的信息熵；根据所述第二目标的类别置信度及信息熵，从所述第二目标中选择出第五目标；根据所述第三样本图像中的第三目标的类别以及待挖掘的样本图像的总数量，分别确定各个类别待挖掘的样本图像的第二数量；根据所述第三样本图像中的第三目标的特征信息，所述第五目标的特征信息以及各个类别待挖掘的样本图像的第二数量，从所述第五目标中确定出第四目标及所述第四目标所在的第一样本图像。

在一种可能的实现方式中，根据所述第二目标的类别置信度及信息熵，从所述第二目标中选择出第五目标，包括：根据所述第二目标的类别置信度及信息熵，分别对所述第二目标进行排序，选择出第三数量的第六目标和第四数量的第七目标；对所述第六目标和所述第七目标进行合并，得到所述第五目标。

在一种可能的实现方式中，根据所述第三样本图像中的第三目标的类别以及待挖掘的样本图像的总数量，分别确定各个类别待挖掘的样本图像的第二数量，包括：根据所述第三样本图像中的第三目标的类别，确定各个类别的第三目标的比例；根据各个类别的第三目标的比例，确定各个类别的抽样比重；根据各个类别的抽样比重，分别确定各个类别待挖掘的样本图像的第二数量。

在一种可能的实现方式中，根据所述第三样本图像中的第三目标的特征信息，所述第五目标的特征信息以及各个类别待挖掘的样本图像的第二数量，从所述第五目标中确定出第四目标及所述第四目标所在的第一样本图像，包括：根据第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离，分别确定所述第一类别的第三目标中与各个第五目标距离最小的第三目标，并作为第八目标，所述第一类别为第三目标的类别中的任意一个；将所述第八目标中距离最大的目标，确定为第四目标。

在一种可能的实现方式中，根据所述第三样本图像中的第三目标的特征信息，所述第五目标的特征信息以及各个类别待挖掘的样本图像的第二数量，从所述第五目标中确定出第四目标及所述第四目标所在的第一样本图像，还包括：将确定出的第四目标添加到所述第一类别的第三目标中，并将所述确定出的第四目标从未标注的第五目标中移除。

在一种可能的实现方式中，所述方法还包括：将所述第三样本图像输入所述目标检测网络中处理，得到所述第三样本图像中的第三目标的特征信息。

在一种可能的实现方式中，在所述将未标注的第一样本图像输入目标检测网络中处理，得到所述第一样本图像的目标检测结果的步骤之前，所述方法还包括：

通过已标注的第三样本图像对所述目标检测网络进行预训练。

在一种可能的实现方式中，所述第一样本图像包括长尾图像。

根据本公开实施例的一方面，提供了一种目标检测方法，该方法包括：将待处理图像输入目标检测网络中处理，得到所述待处理图像的目标检测结果，所述目标检测结果包括所述待处理图像中目标的位置和类别，所述目标检测网络是根据上述的网络训练方法训练得到的。

根据本公开实施例的一方面，提供了一种网络训练装置，包括：

目标检测部分，被配置为将未标注的第一样本图像输入目标检测网络中处理，得到所述第一样本图像的目标检测结果，所述目标检测结果包括所述第一样本图像中目标的图像区域、特征信息及分类概率；

置信度确定部分，被配置为根据所述目标的分类概率，确定所述目标的类别置信度；

标注部分，被配置为针对所述目标中类别置信度大于或等于第一阈值的第一目标，将所述第一目标所在的第一样本图像作为已标注的第二样本图像，并加入训练集中，其中，所述第二样本图像的标注信息包括所述第一目标的图像区域及与所述第一目标的类别置信度对应的类别，所述训练集中包括已标注的第三样本图像；

特征挖掘部分，被配置为针对所述目标中类别置信度小于所述第一阈值的第二目标，根据所述第三样本图像中的第三目标的特征信息，对所述第二目标进行特征相关挖掘，通

过特征相关挖掘,从所述第二目标中确定出第四目标及所述第四目标所在的第一样本图像,并将所述第四目标所在的第一样本图像作为第四样本图像,并加入所述训练集中;

训练部分,被配置为根据所述第四样本图像的标注信息,所述训练集中的第二样本图像、第三样本图像及所述第四样本图像,训练所述目标检测网络。

5 在一种可能的实现方式中,所述训练部分包括:采样数量确定子部分,被配置为根据所述训练集的正样本图像中目标的类别,分别确定从各个类别的正样本图像中采样的第一数量,所述正样本图像为图像中包括目标的样本图像;第一采样子部分,被配置为根据各个类别的正样本图像中采样的第一数量,对各个类别的正样本图像进行采样,得到多个第五样本图像;第二采样子部分,被配置为对所述训练集的负样本图像进行采样,得到多个第六样本图像,所述负样本图像为图像中不包括目标的样本图像;训练子部分,被配置为根据所述第五样本图像及所述第六样本图像,训练所述目标检测网络。

10 在一种可能的实现方式中,所述特征挖掘部分包括:信息熵确定子部分,被配置为根据所述第二目标的分类概率,确定所述第二目标的信息熵;目标选择子部分,被配置为根据所述第二目标的类别置信度及信息熵,从所述第二目标中选择出第五目标;挖掘数量确定子部分,被配置为根据所述第三样本图像中的第三目标的类别以及待挖掘的样本图像的总数量,分别确定各个类别待挖掘的样本图像的第二数量;目标及图像确定子部分,被配置为根据所述第三样本图像中的第三目标的特征信息,所述第五目标的特征信息以及各个类别待挖掘的样本图像的第二数量,从所述第五目标中确定出第四目标及所述第四目标所在的第一样本图像。

20 在一种可能的实现方式中,所述目标选择子部分被配置为:根据所述第二目标的类别置信度及信息熵,分别对所述第二目标进行排序,选择出第三数量的第六目标和第四数量的第七目标;对所述第六目标和所述第七目标进行合并,得到所述第五目标。

25 在一种可能的实现方式中,所述挖掘数量确定子部分被配置为:根据所述第三样本图像中的第三目标的类别,确定各个类别的第三目标的比例;根据各个类别的第三目标的比例,确定各个类别的抽样比重;根据各个类别的抽样比重,分别确定各个类别待挖掘的样本图像的第二数量。

30 在一种可能的实现方式中,所述目标及图像确定子部分被配置为:根据第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离,分别确定所述第一类别的第三目标中与各个第五目标距离最小的第三目标,并作为第八目标,所述第一类别为第三目标的类别中的任意一个;将所述第八目标中距离最大的目标,确定为第四目标。

在一种可能的实现方式中,所述目标及图像确定子部分还被配置为:将确定出的第四目标添加到所述第一类别的第三目标中,并将所述确定出的第四目标从未标注的第五目标中移除。

35 在一种可能的实现方式中,所述装置还包括:特征提取部分,被配置为将所述第三样本图像输入所述目标检测网络中处理,得到所述第三样本图像中的第三目标的特征信息。

在一种可能的实现方式中,在所述目标检测部分之前,所述装置还包括:预训练部分,被配置为通过已标注的第三样本图像对所述目标检测网络进行预训练。

在一种可能的实现方式中,所述第一样本图像包括长尾图像。

40 根据本公开实施例的一方面,提供了一种目标检测装置,所述装置包括:检测处理部分,被配置为将待处理图像输入目标检测网络中处理,得到所述待处理图像的目标检测结果,所述目标检测结果包括所述待处理图像中目标的位置和类别,所述目标检测网络是根据上述的网络训练方法训练得到的。

45 在一种可能的实现方式中,在所述根据所述训练集的正样本图像中目标的类别,分别确定从各个类别的正样本图像中采样的第一数量的步骤之前,包括:对训练集中的正样本图像和负样本图像进行采样,得到数量相同或相近的正样本图像和负样本图像。

在一种可能的实现方式中，所述待挖掘的样本图像的总数量为所述第一样本图像的总数量的5%~25%。

在一种可能的实现方式中，所述对所述第六目标和所述第七目标进行合并，得到所述第五目标，包括：去除所述第六目标中与所述第七目标相同的目标，得到所述第六目标中与所述第七目标不同的剩余目标；将所述剩余目标和所述第七目标作为所述第五目标。

在一种可能的实现方式中，所述方法还包括：在所述第一类别的所述第四样本图像的数量达到所述第一类别的待挖掘的样本图像的第二数量时，结束对所述第一类别的特征相关挖掘。

在一种可能的实现方式中，在所述根据第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离，分别确定所述第一类别的第三目标中与各个第五目标距离最小的第三目标，并作为第八目标的步骤之后，所述方法还包括：在所述第四目标所在的第一样本图像的数量达到所述第一类别的待挖掘的样本图像的第二数量时，结束对所述第八目标的确定。

在一种可能的实现方式中，在所述根据第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离，分别确定所述第一类别的第三目标中与各个第五目标距离最小的第三目标，并作为第八目标的步骤之后，所述方法还包括：在所述第四目标所在的第一样本图像的数量未达到所述第一类别的待挖掘的样本图像的第二数量，且存储所述第五目标的特征信息的集合为空时，结束对所述第八目标的确定。

在一种可能的实现方式中，所述将所述第三样本图像输入所述目标检测网络中处理，得到所述第三样本图像中的第三目标的特征信息，包括：将所述第三样本图像，输入所述目标检测网络中，得到所述目标检测网络的隐藏层输出的特征向量；将所述特征向量确定为所述第三目标的特征信息。

根据本公开实施例的一方面，提供了一种电子设备，包括：处理器；被配置为存储处理器可执行指令的存储器；其中，所述处理器被配置为调用所述存储器存储的指令，以执行上述网络训练方法，或执行上述目标检测方法。

根据本公开实施例的一方面，提供了一种计算机可读存储介质，其上存储有计算机程序指令，所述计算机程序指令被处理器执行时实现网络训练方法，或实现上述目标检测方法。

根据本公开的实施例，能够通过目标检测网络获取未标注样本图像的目标检测结果；根据目标检测结果分别进行伪标注和特征相关挖掘，标注并收集高价值的样本图像，加入训练集；根据扩充后的训练集训练目标检测网络，从而扩充训练集中的正样本数据数量，缓解正负样本之间的不平衡问题，提高了目标检测网络的训练效果。

应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，而非限制本公开实施例。根据下面参考附图对示例性实施例的详细说明，本公开实施例的其它特征及方面将变得清楚。

附图说明

此处的附图被并入说明书中并构成本说明书的一部分，这些附图示出了符合本公开的实施例，并与说明书一起用于说明本公开实施例的技术方案。

图1示出根据本公开实施例的网络训练方法的流程图。

图2示出根据本公开实施例的网络训练方法的处理过程的示意图。

图3示出根据本公开实施例的网络训练装置的框图。

图4示出根据本公开实施例的一种电子设备的框图。

图5示出根据本公开实施例的一种电子设备的框图。

具体实施方式

以下将参考附图详细说明本公开的各种示例性实施例、特征和方面。附图中相同的附图标记表示功能相同或相似的元件。尽管在附图中示出了实施例的各种方面，但是除非特别指出，不必按比例绘制附图。

5 在这里专用的词“示例性”意为“用作例子、实施例或说明性”。这里作为“示例性”所说明的任何实施例不必解释为优于或好于其它实施例。

本文中术语“和/或”，仅仅是一种描述关联对象的关联关系，表示可以存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况。另外，本文中术语“至少一种”表示多种中的任意一种或多种中的至少两种的任意组合，例如，10 包括 A、B、C 中的至少一种，可以表示包括从 A、B 和 C 构成的集合中选择的任意一个或多个元素。

另外，为了更好地说明本公开实施例，在下文的具体实施方式中给出了众多的具体细节。本领域技术人员应当理解，没有某些具体细节，本公开实施例同样可以实施。在一些实例中，对于本领域技术人员熟知的方法、手段、元件和电路未作详细描述，以便于凸显15 本公开实施例的主旨。

图 1 示出根据本公开实施例的网络训练方法的流程图，如图 1 所示，所述网络训练方法包括：

在步骤 S11 中，将未标注的第一样本图像输入目标检测网络中处理，得到所述第一样本图像的目标检测结果，所述目标检测结果包括所述第一样本图像中目标的图像区域、特征信息及分类概率；20

在步骤 S12 中，根据所述目标的分类概率，确定所述目标的类别置信度；

在步骤 S13 中，针对所述目标中类别置信度大于或等于第一阈值的第二目标，将所述第一目标所在的第一样本图像作为已标注的第二样本图像，并加入训练集中，其中，所述第二样本图像的标注信息包括所述第一目标的图像区域及与所述第一目标的类别置信度25 对应的类别，所述训练集中包括已标注的第三样本图像；

在步骤 S14 中，针对所述目标中类别置信度小于所述第一阈值的第二目标，根据所述第三样本图像中的第三目标的特征信息，对所述第二目标进行特征相关挖掘，通过特征相关挖掘，从所述第二目标中确定出第四目标及所述第四目标所在的第一样本图像，并将所述第四目标所在的第一样本图像作为第四样本图像，并加入所述训练集中；30

在步骤 S15 中，根据所述第四样本图像的标注信息，所述训练集中的第二样本图像、第三样本图像及所述第四样本图像，训练所述目标检测网络。

在一种可能的实现方式中，所述方法可以由终端设备或服务器等电子设备执行，终端设备可以为用户设备（User Equipment, UE）、移动设备、用户终端、终端、蜂窝电话、无绳电话、个人数字处理（Personal Digital Assistant, PDA）、手持设备、计算设备、车载设备、可穿戴设备等，所述方法可以通过处理器调用存储器中存储的计算机可读指令的方式来实现。或者，可通过服务器执行所述方法。35

举例来说，第一样本图像可以是通过图像采集设备（例如摄像头）采集的图像。第一样本图像可包括大规模的长尾（Long-tailed）图像，也即大部分图像为背景图像，小部分图像中包括可检测的目标。可检测的目标可例如包括人体、人脸、车辆、物体等。例如在安防领域中，可通过摄像头采集某一地理区域的图像，可能仅有小部分时间有人经过该地理区域，从而采集到的图像大部分为背景图像，仅小部分图像中包括人脸和/或人体。在该情况下，采集到的多个图像即可组成长尾数据集。本公开实施例对第一样本图像的获取方式及第一样本图像中目标的类别不作限制。40

在一种可能的实现方式中，可预先设置有目标检测网络，用于检测图像中目标的位置（也即检测框）和类别。该目标检测网络可例如为卷积神经网络，本公开实施例对目标检45

测网络的网络结构不作限制。

在一种可能的实现方式中，在步骤 S11 之前，该方法还包括：通过已标注的第三样本图像对所述目标检测网络进行预训练。也就是说，可预设训练集，该训练集中包括已标注的第三样本图像，第三样本图像的标注信息可包括图像中目标的检测框和类别。根据该训练集，可采用相关技术中的方式对目标检测网络进行预训练，以使该目标检测网络具有一定的检测精度。

然而，预训练后的目标检测网络对大规模长尾图像的检测效果较差，因此，可通过主动学习的方式，采用未标注的第一样本图像进一步训练目标检测网络。

在一种可能的实现方式中，在步骤 S11 中，可将未标注的第一样本图像输入目标检测网络中处理，得到第一样本图像的目标检测结果。该目标检测结果可包括第一样本图像中目标的图像区域、特征信息及分类概率。目标所在的图像区域可为图像中的检测框；目标的特征信息可例如为目标检测网络的隐藏层（例如卷积层）输出的特征向量；目标的分类概率可表示该目标属于各个类别的分类后验概率。

在一种可能的实现方式中，第一样本图像中的目标也可称为实例，每个第一样本图像中可能检测出一个或多个目标。在实际处理中，检测到的目标的数量量级可能是图像数量量级的几倍到几十倍。

在一种可能的实现方式中，在步骤 S12 中，根据目标的分类概率，可求取分类概率的最大值，确定为该目标的类别置信度。

在一种可能的实现方式中，在步骤 S13 中，针对类别置信度大于或等于第一阈值的目标（可称为第一目标），可将该第一目标所在的第一样本图像作为已标注的样本图像（可称为第二样本图像），并加入训练集中。将第一目标的图像区域作为标注的图像区域，将与该第一目标的类别置信度对应的类别作为该第一目标的标注类别。同一个第二样本图像可能被该第二样本图像中的多个第一目标标注多次。其中，第一阈值例如为 0.99，本公开实施例对第一阈值的取值不作限制。

在一种可能的实现方式中，步骤 S13 的处理过程可称为伪标注（pseudo-labeling）。也即，将置信度较高的目标所在的图像作为高价值的样本，将目标检测推理结果直接作为目标的标注结果。通过这种方式，可以扩充训练集中正样本数据的数量，以解决正样本收集困难的问题。

在一种可能的实现方式中，在步骤 S14 中，针对类别置信度小于第一阈值的目标（可称为第二目标），可根据训练集中已标注的第三样本图像中目标（可称为第三目标）的特征信息，对第二目标进行特征相关挖掘，从第二目标中挖掘出满足要求的目标（可称为第四目标）。例如，可计算第三目标的特征信息与第二目标的特征信息之间的距离或相关度，根据距离或相关度选择出预设数量的目标，并将选择出的预设数量的目标作为第四目标。

在一种可能的实现方式中，可将挖掘到的第四目标所在的第一样本图像作为第四样本图像，并加入所述训练集中，从而完成特征相关挖掘的处理过程。通过这种方式，能够进一步扩充训练集中样本数据的数量。

在一种可能的实现方式中，可通过人工标注的方式获取第四样本图像的标注信息，例如人工确定第四样本图像中目标的检测框和类别。本公开实施例对此不作限制。

在一种可能的实现方式中，在步骤 S15 中，在得到第四样本图像的标注信息后，可根据训练集中的第二样本图像、第三样本图像及第四样本图像，训练目标检测网络。

根据本公开的实施例，能够通过目标检测网络获取未标注样本图像的目标检测结果；根据目标检测结果分别进行伪标注和特征相关挖掘，标注并收集高价值的样本图像，加入训练集；根据扩充后的训练集训练目标检测网络，从而扩充训练集中的正样本数据数量，缓解正负样本之间的不平衡问题，提高了目标检测网络的训练效果。

在一种可能的实现方式中，经过步骤 S11 处理，得到各个第一样本图像的目标检测结

果；经过 S12 处理，得到各个第一样本图像中目标的类别置信度。在步骤 S13 中，可将类别置信度大于或等于第一阈值的第一目标所在的样本图像加入训练集，通过伪标注方式得到已标注的第二样本图像；在步骤 S14 中，可对类别置信度小于第一阈值的第二目标进行挖掘。

5 在一种可能的实现方式中，步骤 S14 可包括：

 根据所述第二目标的分类概率，确定所述第二目标的信息熵；

 根据所述第二目标的类别置信度及信息熵，从所述第二目标中选择出第五目标；

 根据所述第三样本图像中的第三目标的类别以及待挖掘的样本图像的总数量，分别确定各个类别待挖掘的样本图像的第二数量；

10 根据所述第三样本图像中的第三目标的特征信息，所述第五目标的特征信息以及各个类别待挖掘的样本图像的第二数量，从所述第五目标中确定出第四目标及所述第四目标所在的第一样本图像。

 举例来说，根据第二目标的分类概率，可计算得到第二目标的信息熵，用于表示第二目标的不确定程度，也即，第二目标的信息熵越大，则第二目标的不确定程度越大；反之，
15 第二目标的信息熵越小，则第二目标的不确定程度越小。本公开实施例对信息熵的计算方式不作限制。

 在一种可能的实现方式中，根据第二目标的类别置信度及信息熵，可分别从多个第二目标中选择出满足一定条件的目标（可称为第五目标），例如选择出类别置信度较大的目标、信息熵较大的目标等。

20 在一种可能的实现方式中，根据所述第二目标的类别置信度及信息熵，从所述第二目标中选择出第五目标的步骤，可包括：

 根据所述第二目标的类别置信度及信息熵，分别对所述第二目标进行排序，选择出第三数量的第六目标和第四数量的第七目标；

 对所述第六目标和所述第七目标进行合并，得到所述第五目标。

25 也就是说，根据第二目标的类别置信度，对多个第二目标进行排序；根据排序结果，从多个第二目标中选择出预设的第三数量的目标（可称为第六目标）。类似地，根据第二目标的信息熵，对多个第二目标进行排序；根据排序结果，从多个第二目标中选择出预设的第四数量的目标（可称为第七目标）。其中，第三数量和第四数量可分别为 3K，K 表示待挖掘的样本图像的数量，K 例如取值为 10000。在实际处理中，K 的取值可能为未标注的第一样本图像的总数量的 5%~25%。本公开实施例对 K 的取值，以及第三数量和第四数量与 K 之间的数量关系均不作限制。

 应当理解，本领域技术人员可根据实际情况设置待挖掘的样本图像的数量 K、第三数量及第四数量的取值，且第三数量和第四数量可以不同，本公开实施例对此不作限制。

35 在一种可能的实现方式中，可将选取的第六目标和第七目标合并，将合并后的多个目标作为第五目标，以便去除其中可能存在的重复目标。在实际处理中，可得到大约 6K 个第五目标。

 上述的处理方式可称为自举法（bootstrapping），通过这种方式，可从第二目标中同时选取一定数量的、可能性较高的正样本和负样本，以便后续进行特征相关挖掘，从而降低特征相关挖掘的计算量，提高处理效率。

40 在一种可能的实现方式中，根据所述第三样本图像中的第三目标的类别以及待挖掘的样本图像的总数量，分别确定各个类别待挖掘的样本图像的第二数量的步骤，可包括：

 根据所述第三样本图像中的第三目标的类别，确定各个类别的第三目标的比例；

 根据各个类别的第三目标的比例，确定各个类别的抽样比重；

 根据各个类别的抽样比重，分别确定各个类别待挖掘的样本图像的第二数量。

45 举例来说，根据训练集中已有标注的第三样本图像中的第三目标的类别，

可确定各个类别的第三目标的比例 f_c ；根据该比例 f_c ，可通过如下公式计算出各个类别的抽样比重 $\hat{r}_c$ 。

$$r_c = \max(f_c, t \cdot \exp(f_c / t - 1)) \quad (1)$$

$$\hat{r}_c = \frac{r_c}{\sum_{i \in C} r_i} \quad (2)$$

5 在公式(1)和(2)中， r_c 表示类别 c 的抽样值； t 为超参数，取值例如为0.1； C 表示类别数量； r_i 表示 C 个类别中第 i 个类别的抽样值。

通过公式(1)和(2)的处理，可提高比例较小的类别所对应的抽样比重，并降低比例较大的类别所对应的抽样比重，从而缓解不同类别的样本之间的数量不平衡的问题，以便提高网络的训练效果。

10 在一种可能的实现方式中，根据各个类别的抽样比重 $\hat{r}_c$ 以及待挖掘的样本图像的总数量(K 个)，可确定出各个类别待挖掘的样本图像的第二数量。进而可根据第二数量进行特征相关挖掘。

在一种可能的实现方式中，所述方法还包括：将所述第三样本图像输入所述目标检测网络中处理，得到所述第三样本图像中的第三目标的特征信息。

15 也就是说，可将训练集中已标注的第三样本图像，输入到目标检测网络中，由目标检测网络的隐藏层(例如卷积层)输出该第三样本图像的特征信息，例如特征向量。通过这种方式，可得到第三样本图像的特征，以便于后续的特征相关挖掘。

在一种可能的实现方式中，根据所述第三样本图像中的第三目标的特征信息，所述第五目标的特征信息以及各个类别待挖掘的样本图像的第二数量，从所述第五目标中确定出第四目标及所述第四目标所在的第一样本图像，包括：

20 根据第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离，分别确定所述第一类别的第三目标中与各个第五目标距离最小的第三目标，并作为第八目标，所述第一类别为第三目标的类别中的任意一个；

将所述第八目标中距离最大的目标，确定为第四目标。

25 举例来说，在确定各个类别待挖掘的样本图像的第二数量后，可采用 k 中心(k -center)方式，从第五目标所在的样本图像中挖掘对应数量的样本图像。针对第三目标的多个类别中的任一类别(可称为第一类别)，可计算该第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离，该距离可例如为欧氏距离。对于任意一个第五目标，可确定第一类别中的第三目标中与该第五目标距离最小的第三目标，从而可分别确定与各个第五目标距离最小的第三目标，可称为第八目标。

30 在一种可能的实现方式中，可从各个第八目标中，选择出距离最大的一个目标，确定为本次的特征相关挖掘得到的第四目标。如下公式所示：

$$u = \arg \max_{j \in \hat{D}_{UL}} \min_{l \in \hat{D}_L^c} \text{dist}(f_j, g_l) \quad (3)$$

35 在公式(3)中， u 表示特征相关挖掘得到的第四目标； $\text{dist}(f_j, g_l)$ 表示第 j 个第五目标的特征信息 f_j 与第一类别 c 的第 l 个第三目标的特征信息 g_l 之间的距离； $\hat{D}_{UL}$ 表示第五目标的特征信息的集合； $\hat{D}_L^c$ 表示第一类别 c 的第三目标的特征信息的集合。

在一种可能的实现方式中，可确定该第四目标所在的第一样本图像，并将该样本图像加入训练集中，作为第四样本图像，从而完成本次的特征相关挖掘过程。

40 在一种可能的实现方式中，根据所述第三样本图像中的第三目标的特征信息，所述第五目标的特征信息以及各个类别待挖掘的样本图像的第二数量，从所述第五目标中确定出第四目标及所述第四目标所在的第一样本图像的步骤，还包括：

将确定出的第四目标添加到所述第一类别的第三目标中，并将所述确定出的第四目标从未标注的第五目标中移除。

也就是说，将该次特征相关挖掘得到的第四目标作为已标注目标，并将该第四目标从未标注目标中移除。在该情况下，可将该第四目标的特征信息加入到第一类别c的第三目标的特征信息的集合 $\hat{D}_L^c$ 中，从第五目标的特征信息的集合 $\hat{D}_{UL}$ 中移除。这样，在下次的特征相关挖掘中，可以通过公式（3）对更新后的两个集合进行挖掘，重复上述过程。

在一种可能的实现方式中，在第一类别的第四样本图像的数量达到第一类别的第二数量，或未达到第二数量且第五目标耗尽（集合 $\hat{D}_{UL}$ 为空）时，可完成该第一类别的特征相关挖掘。

通过这种方式，可分别对各个类别进行特征相关挖掘，最终得到足够数量的第四样本图像（通常为K个样本图像），从而进一步扩充训练集中的样本图像的数量，并缓解正负样本之间的不均衡。

在一种可能的实现方式中，可对挖掘到的第四样本图像进行人工标注（human annotation），得到第四样本图像的标注信息。由于第四样本图像中可能同时存在正样本图像（也即图像中包括目标的第四样本图像）和负样本图像（也即图像中不包括目标的第四样本图像），因此，第四样本图像的标注信息可包括图像是正样本图像或负样本图像的样本类别信息，正样本图像中目标所在的图像框及目标的类别。

在一种可能的实现方式中，在完成人工标注后，可在步骤S15中根据所述第四样本图像的标注信息，所述训练集中的第二样本图像、第三样本图像及第四样本图像，训练目标检测网络。

其中，步骤S15可包括：根据所述训练集的正样本图像中目标的类别，分别确定从各个类别的正样本图像中采样的第一数量，所述正样本图像为图像中包括目标的样本图像；

根据各个类别的正样本图像中采样的第一数量，对各个类别的正样本图像进行采样，得到多个第五样本图像；

对所述训练集的负样本图像进行采样，得到多个第六样本图像，所述负样本图像为图像中不包括目标的样本图像；

根据所述第五样本图像及所述第六样本图像，训练所述目标检测网络。

举例来说，可通过重采样（resampling）的方式来训练目标检测网络，通过重采样来增加数据中出现频次较低的数据的采样频率，来改善网络对于这些数据的性能，进一步改善正负样本之间的不均衡。

在一种可能的实现方式中，可分别对训练集（包括第二样本图像、第三样本图像及第四样本图像）中的正样本图像和负样本图像进行采样，以使采样后正样本图像和负样本图像的数量相同或相近。

在一种可能的实现方式中，对于正样本图像，可预设正样本图像的采样总数量。根据训练集中的正样本图像中目标的类别，分别确定从各个类别的正样本图像中采样的第一数量。

与前面的处理过程类似，根据正样本图像中目标的类别，可确定各个类别的目标的比例；根据该比例，可通过如下公式计算出各个类别的抽样比重：

$$R_h = \begin{cases} t_1 * e^{\frac{q_h - 1}{t_1}} & 0 < q_h \leq t_1 \\ q_h & t_1 < q_h < 1 \end{cases} \quad (4)$$

公式（4）中， R_h 表示第h个类别的正样本图像的抽样比重； q_h 表示第h个类别的目标的比例； t_1 为超参数，取值例如为0.1。

通过公式（4）的处理，可提高比例较小的类别所对应的抽样比重，并降低比例较大

的类别所对应的抽样比重，从而缓解不同类别的正样本图像之间的数量不平衡，以便提高网络的训练效果。

在一种可能的实现方式中，根据各个类别的正样本图像的抽样比重以及正样本图像的采样总数量，可确定各个类别的正样本图像的第一数量。

5 在一种可能的实现方式中，对于任意一个类别，可根据该类别的第一数量，在该类别的正样本图像中随机采样出第一数量的正样本图像，作为第五样本图像。对各个类别的正样本图像分别进行采样，可得到采样总数量的第五样本图像。

10 在一种可能的实现方式中，对于负样本图像，可根据预设的采样总数量对训练集中负样本图像直接进行随机采样，得到采样总数量的第六样本图像。负样本图像的该采样总数量可与正样本图像的采样总数量相同或不同，本公开实施例对此不作限制。

15 在一种可能的实现方式中，可根据第五样本图像及第六样本图像，训练目标检测网络。也即，将第五及第六样本图像分别输入目标检测网络，得到第五及第六样本图像的目标检测结果；根据目标检测结果及标注信息，确定目标检测网络的损失；根据损失反向调整目标检测网络的参数；经过多轮迭代，在满足预设条件（例如网络收敛）时，得到训练后的目标检测网络。

通过这种方式，可显著提高训练后的目标检测网络对于长尾图像的检测效果。

在一种可能的实现方式中，在步骤S11之前，通过已标注的第三样本图像对所述目标检测网络进行预训练的步骤，也可以采用上述的重采样训练方式进行，从而提高目标检测网络的预训练效果。

20 在实际应用中，可重复步骤S11-S15的整个处理过程，实现持续的增量训练。也就是说，当再次采集到未标注的样本图像时，可将本次训练后的目标检测网络作为初始的目标检测网络，将本次扩展后的训练集作为初始的训练集，重复进行伪标注-特征相关挖掘-重采样训练的处理过程，从而持续提升目标检测网络的性能。

25 图2示出根据本公开实施例的网络训练方法的处理过程的示意图。如图2所示，数据源中包括大量未标注的第一样本图像20，将第一样本图像20输入目标检测网络中进行预测（predict），得到各个第一样本图像20的目标检测结果21，包括第一样本图像中目标的图像区域（未示出）、特征向量及分类概率。

30 如图2所示，在该示例中，目标检测网络可包括CNN主干网络211、特征图金字塔网络（FPN）212以及全连接网络213，全连接网络213例如为bbox head。第一样本图像20输入到目标检测网络后，经由CNN主干网络211及FPN 212处理，得到第一样本图像的特征图214，特征图214经由全连接网络213处理，得到目标检测结果21。

35 在该示例中，可根据目标的分类概率，确定目标的类别置信度；对于类别置信度大于或等于第一阈值（例如为0.99）的第一目标，确定出这些第一目标所在的第一样本图像，作为第二样本图像22，并对第二样本图像22进行伪标注，也即将第一目标的图像区域及与第一目标的类别置信度对应的类别作为第二样本图像22的标注信息。将已标注的第二样本图像22加入到训练集25中，从而实现对训练集中正样本的扩充。

40 在该示例中，对于类别置信度小于第一阈值的第二目标，通过自举法选择出一定数量的第五目标，得到第五目标所在的样本图像23。根据训练集中已标注的第三样本图像中的第三目标的特征向量（未示出），对第五目标进行特征相关挖掘，确定出第四目标以及第四目标所在的第一样本图像，作为第四样本图像24。对第四样本图像24进行人工标注，加入训练集25中，从而实现对训练集中已标注图像的进一步扩充。

45 在该示例中，经两次扩充后，训练集25中包括已标注的第二样本图像、第三样本图像及第四样本图像。对训练集25进行重采样，平衡正负样本的数量，以及不同类别的正样本的数量，得到重采样后的训练集26；进而根据重采样后的训练集26，训练目标检测网络，从而完成整个处理过程。

根据本公开的实施例，还提供了一种目标检测方法，该方法包括：

将待处理图像输入目标检测网络中处理，得到所述待处理图像的目标检测结果，所述目标检测结果包括所述待处理图像中目标的位置和类别，所述目标检测网络是根据上述的网络训练方法训练得到的。

也就是说，可将上述方法训练得到的目标检测网络进行部署，实现待处理图像的目标检测。待处理图像可例如为图像采集设备（例如摄像头）采集的图像，图像中可能包括待检测的目标，例如，人体、人脸、车辆、物体等。本公开实施例对此不作限制。

在一种可能的实现方式中，可将待处理图像输入目标检测网络中处理，得到所述待处理图像的目标检测结果。该目标检测结果包括待处理图像中目标的位置和类别，例如待处理图像中人脸所在的检测框和人脸对应的身份。

通过这种方式，可提高目标检测的检测精度，实现大规模长尾图像数据的目标检测。

根据本公开实施例的网络训练方法，利用主动学习挖掘方法来对潜在的无标注数据进行挖掘，利用半监督学习方法来对辅助对无标注数据进行标注，扩充正样本数据的数量，从而解决大规模长尾检测中数据规模大且正样本收集困难的问题，并且，在一定程度上缓解了正负样本之间的不均衡的问题。在有限的标注与计算资源环境下有效提升了模型性能。

根据本公开实施例的网络训练方法，采用重采样的方式训练目标检测网络，能够解决正负样本不均衡对网络训练的负面影响，并缓解正样本不同类别之间不均衡对网络训练的负面影响，使得目标检测网络在训练时能够有效收敛并提高网络性能。

根据本公开实施例的网络训练方法，利用主动学习方法，可在巨量的未标注数据中，挖掘对于模型提升有帮助的潜在高价值样本，能够在有限的标注与计算资源环境下有效提升模型性能，大量节省深度学习模型应用在新的业务上所需的人力以及计算成本；利用重采样方法，能够有效在样本不均衡情况下训练目标检测网络，无需过多人工调参干预，节省深度学习模型应用在新的业务上所需的人力成本。

根据本公开实施例的网络训练方法，能够应用于智能视频分析，安防等领域中，在有限的人工以及计算资源下，可以使用本方法在线上对智能视频分析或智能监控中潜在的目标进行检测，并对应用的检测网络进行快速迭代提升，用较小的人力和计算成本快速达到业务所需的性能要求，并能够在之后持续提升网络性能。

本公开实施例的网络训练方法，可以应用于线上智能视频分析或智能监控中，以在有限的人工以及计算资源下，在线上对于智能视频分析或者智能监控中潜在的目标检测应用进行快速迭代提升，从而用较小的人力和计算成本快速达到业务所需的性能要求，并能在之后继续持续提升模型性能。

可以理解，本公开实施例提及的上述各个方法实施例，在不违背原理逻辑的情况下，均可以彼此相互结合形成结合后的实施例，限于篇幅，本公开实施例不再赘述。本领域技术人员可以理解，在具体实施方式的上述方法中，各步骤的具体执行顺序应当以其功能和可能的内在逻辑确定。

此外，本公开实施例还提供了网络训练装置、目标检测装置、电子设备、计算机可读存储介质、程序，上述均可用来实现本公开实施例提供的任一种网络训练方法或目标检测方法，相应技术方案和描述和参见方法部分的相应记载，不再赘述。

图3示出根据本公开实施例的网络训练装置的框图，所述装置包括处理器（图3中未示出），所述处理器用于执行存储器（图3中未示出）中存储的程序部分；如图3所示，存储器中存储的程序部分包括：

目标检测部分31，被配置为将未标注的第一样本图像输入目标检测网络中处理，得到所述第一样本图像的目标检测结果，所述目标检测结果包括所述第一样本图像中目标的图像区域、特征信息及分类概率；

置信度确定部分32，被配置为根据所述目标的分类概率，确定所述目标的类别置信度；

标注部分33, 被配置为针对所述目标中类别置信度大于或等于第一阈值的第二目标, 将所述第二目标所在的第一样本图像作为已标注的第二样本图像, 并加入训练集中, 其中, 所述第二样本图像的标注信息包括所述第二目标的图像区域及与所述第二目标的类别置信度对应的类别, 所述训练集中包括已标注的第三样本图像;

特征挖掘部分34, 被配置为针对所述目标中类别置信度小于所述第一阈值的第二目标, 根据所述第三样本图像中的第三目标的特征信息, 对所述第二目标进行特征相关挖掘, 通过特征相关挖掘, 从所述第二目标中确定出第四目标及所述第四目标所在的第一样本图像, 并将所述第四目标所在的第一样本图像作为第四样本图像, 并加入所述训练集中;

训练部分35, 被配置为根据所述第四样本图像的标注信息, 所述训练集中的第二样本图像、第三样本图像及所述第四样本图像, 训练所述目标检测网络。

在一种可能的实现方式中, 所述训练部分包括: 采样数量确定子部分, 被配置为根据所述训练集的正样本图像中目标的类别, 分别确定从各个类别的正样本图像中采样的第一数量, 所述正样本图像为图像中包括目标的样本图像; 第一采样子部分, 被配置为根据各个类别的正样本图像中采样的第一数量, 对各个类别的正样本图像进行采样, 得到多个第五样本图像; 第二采样子部分, 被配置为对所述训练集的负样本图像进行采样, 得到多个第六样本图像, 所述负样本图像为图像中不包括目标的样本图像; 训练子部分, 被配置为根据所述第五样本图像及所述第六样本图像, 训练所述目标检测网络。

在一种可能的实现方式中, 所述特征挖掘部分包括: 信息熵确定子部分, 被配置为根据所述第二目标的分类概率, 确定所述第二目标的信息熵; 目标选择子部分, 被配置为根据所述第二目标的类别置信度及信息熵, 从所述第二目标中选择出第五目标; 挖掘数量确定子部分, 被配置为根据所述第三样本图像中的第三目标的类别以及待挖掘的样本图像的总数量, 分别确定各个类别待挖掘的样本图像的第二数量; 目标及图像确定子部分, 被配置为根据所述第三样本图像中的第三目标的特征信息, 所述第五目标的特征信息以及各个类别待挖掘的样本图像的第二数量, 从所述第五目标中确定出第四目标及所述第四目标所在的第一样本图像。

在一种可能的实现方式中, 所述目标选择子部分被配置为: 根据所述第二目标的类别置信度及信息熵, 分别对所述第二目标进行排序, 选择出第三数量的第六目标和第四数量的第七目标; 对所述第六目标和所述第七目标进行合并, 得到所述第五目标。

在一种可能的实现方式中, 所述挖掘数量确定子部分被配置为: 根据所述第三样本图像中的第三目标的类别, 确定各个类别的第三目标的比例; 根据各个类别的第三目标的比例, 确定各个类别的抽样比重; 根据各个类别的抽样比重, 分别确定各个类别待挖掘的样本图像的第二数量。

在一种可能的实现方式中, 所述目标及图像确定子部分被配置为: 根据第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离, 分别确定所述第一类别的第三目标中与各个第五目标距离最小的第三目标, 并作为第八目标, 所述第一类别为第三目标的类别中的任意一个; 将所述第八目标中距离最大的目标, 确定为第四目标。

在一种可能的实现方式中, 所述目标及图像确定子部分还被配置为: 将确定出的第四目标添加到所述第一类别的第三目标中, 并将所述确定出的第四目标从未标注的第五目标中移除。

在一种可能的实现方式中, 所述装置还包括: 特征提取部分, 被配置为将所述第三样本图像输入所述目标检测网络中处理, 得到所述第三样本图像中的第三目标的特征信息。

在一种可能的实现方式中, 所述装置还包括: 预训练部分, 被配置为通过已标注的第三样本图像对所述目标检测网络进行预训练。

在一种可能的实现方式中, 所述第一样本图像包括长尾图像。

在一种可能的实现方式中, 采样数量确定子部分, 还被配置为: 在所述根据所述训练

集的正样本图像中目标的类别，分别确定从各个类别的正样本图像中采样的第一数量之前，对训练集中的正样本图像和负样本图像进行采样，得到数量相同或相近的正样本图像和负样本图像。

5 在一种可能的实现方式中，所述待挖掘的样本图像的总数量为所述第一样本图像的总数量的5%~25%。

在一种可能的实现方式中，所述目标选择子部分，还被配置为：去除所述第六目标中与所述第七目标相同的目标，得到所述第六目标中与所述第七目标不同的剩余目标；将所述剩余目标和所述第七目标作为所述第五目标。

10 在一种可能的实现方式中，所述方法还包括：在所述根据第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离，分别确定所述第一类别的第三目标中与各个第五目标距离最小的第三目标，并作为第八目标之后，在所述第四目标所在的第一样本图像的数量达到所述第一类别的待挖掘的样本图像的第二数量时，结束对所述第八目标的确定。

15 在一种可能的实现方式中，所述目标及图像确定子部分还被配置为：在所述根据第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离，分别确定所述第一类别的第三目标中与各个第五目标距离最小的第三目标，并作为第八目标之后，在所述第四目标所在的第一样本图像的数量未达到所述第一类别的待挖掘的样本图像的第二数量，且存储所述第五目标的特征信息的集合为空时，结束对所述第八目标的确定。

20 在一种可能的实现方式中，所述特征提取部分，还被配置为：将所述第三样本图像，输入所述目标检测网络中，得到所述目标检测网络的隐藏层输出的特征向量；将所述特征向量确定为所述第三目标的特征信息。

25 根据本公开的一方面，提供了一种目标检测装置，所述装置包括：检测处理部分，被配置为将待处理图像输入目标检测网络中处理，得到所述待处理图像的目标检测结果，所述目标检测结果包括所述待处理图像中目标的位置和类别，所述目标检测网络是根据上述的网络训练方法训练得到的。

在一些实施例中，本公开实施例提供的装置具有的功能或包含的部分可以被配置为执行上文方法实施例描述的方法，其具体实现可以参照上文方法实施例的描述，为了简洁，这里不再赘述。

30 在一些实施例中，“部分”还可以是部分电路、部分处理器、部分程序或软件等等，当然也可以是单元，还可以是模块也可以是非模块化的。

本公开实施例还提出一种计算机可读存储介质，其上存储有计算机程序指令，所述计算机程序指令被处理器执行时实现上述方法。计算机可读存储介质可以是非易失性计算机可读存储介质。

35 本公开实施例还提出一种电子设备，包括：处理器；被配置为存储处理器可执行指令的存储器；其中，所述处理器被配置为调用所述存储器存储的指令，以执行上述方法。

本公开实施例还提供了一种计算机程序产品，包括计算机可读代码，当计算机可读代码在设备上运行时，设备中的处理器执行被配置为实现如上任一实施例提供的网络训练方法或目标检测方法的指令。

40 本公开实施例还提供了另一种计算机程序产品，被配置为存储计算机可读指令，指令被执行时使得计算机执行上述任一实施例提供的网络训练方法或目标检测方法的检测操作。

电子设备可以被提供为终端、服务器或其它形态的设备。

图4示出根据本公开实施例的一种电子设备800的框图。例如，电子设备800可以是移动电话，计算机，数字广播终端，消息收发设备，游戏控制台，平板设备，医疗设备，健身设备，个人数字助理等终端。

45 参照图4，电子设备800可以包括以下一个或多个组件：处理组件802，存储器804，电

源组件806, 多媒体组件808, 音频组件810, 输入/输出(I/O)的接口812, 传感器组件814, 以及通信组件816。

处理组件802通常控制电子设备800的整体操作, 诸如与显示, 电话呼叫, 数据通信, 相机操作和记录操作相关联的操作。处理组件802可以包括一个或多个处理器820来执行指令, 以完成上述的方法的全部或部分步骤。此外, 处理组件802可以包括一个或多个模块, 便于处理组件802和其他组件之间的交互。例如, 处理组件802可以包括多媒体模块, 以方便多媒体组件808和处理组件802之间的交互。

存储器804被配置为存储各种类型的数据以支持在电子设备800的操作。这些数据的示例包括用于在电子设备800上操作的任何应用程序或方法的指令, 联系人数据, 电话簿数据, 消息, 图片, 视频等。存储器804可以由任何类型的易失性或非易失性存储设备或者它们的组合实现, 如静态随机存取存储器(SRAM), 电可擦除可编程只读存储器(EEPROM), 可擦除可编程只读存储器(EPROM), 可编程只读存储器(PROM), 只读存储器(ROM), 磁存储器, 快闪存储器, 磁盘或光盘。

电源组件806为电子设备800的各种组件提供电力。电源组件806可以包括电源管理系统, 一个或多个电源, 及其他与为电子设备800生成、管理和分配电力相关联的组件。

多媒体组件808包括在所述电子设备800和用户之间的提供一个输出接口的屏幕。在一些实施例中, 屏幕可以包括液晶显示器(LCD)和触摸面板(TP)。如果屏幕包括触摸面板, 屏幕可以被实现为触摸屏, 以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界, 而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中, 多媒体组件808包括一个前置摄像头和/或后置摄像头。当电子设备800处于操作模式, 如拍摄模式或视频模式时, 前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

音频组件810被配置为输出和/或输入音频信号。例如, 音频组件810包括一个麦克风(MIC), 当电子设备800处于操作模式, 如呼叫模式、记录模式和语音识别模式时, 麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器804或经由通信组件816发送。在一些实施例中, 音频组件810还包括一个扬声器, 用于输出音频信号。

I/O接口812为处理组件802和外围接口模块之间提供接口, 上述外围接口模块可以是键盘, 点击轮, 按钮等。这些按钮可包括但不限于: 主页按钮、音量按钮、启动按钮和锁定按钮。

传感器组件814包括一个或多个传感器, 用于为电子设备800提供各个方面的状态评估。例如, 传感器组件814可以检测到电子设备800的打开/关闭状态, 组件的相对定位, 例如所述组件为电子设备800的显示器和小键盘, 传感器组件814还可以检测电子设备800或电子设备800一个组件的位置改变, 用户与电子设备800接触的存在或不存在, 电子设备800方位或加速/减速和电子设备800的温度变化。传感器组件814可以包括接近传感器, 被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件814还可以包括光传感器, 如互补金属氧化物半导体(CMOS)或电荷耦合装置(CCD)图像传感器, 用于在成像应用中使用。在一些实施例中, 该传感器组件814还可以包括加速度传感器, 陀螺仪传感器, 磁传感器, 压力传感器或温度传感器。

通信组件816被配置为便于电子设备800和其他设备之间有线或无线方式的通信。电子设备800可以接入基于通信标准的无线网络, 如无线网络(WiFi), 第二代移动通信技术(2G)或第三代移动通信技术(3G), 或它们的组合。在一个示例性实施例中, 通信组件816经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中, 所述通信组件816还包括近场通信(NFC)模块, 以促进短程通信。例如, 在NFC模块可基于射频识别(RFID)技术, 红外数据协会(IrDA)技术, 超宽带(UWB)技术, 蓝牙(BT)

技术和其他技术来实现。

在示例性实施例中，电子设备800可以被一个或多个应用专用集成电路（ASIC）、数字信号处理器（DSP）、数字信号处理设备（DSPD）、可编程逻辑器件（PLD）、现场可编程门阵列（FPGA）、控制器、微控制器、微处理器或其他电子元件实现，用于执行上述方法。

5 在示例性实施例中，还提供了一种非易失性计算机可读存储介质，例如包括计算机程序指令的存储器804，上述计算机程序指令可由电子设备800的处理器820执行以完成上述方法。

图5示出根据本公开实施例的一种电子设备1900的框图。例如，电子设备1900可以被提供为一服务器。参照图5，电子设备1900包括处理组件1922，其进一步包括一个或多个
10 处理器，以及由存储器1932所代表的存储器资源，被配置为存储可由处理组件1922的执行的指令，例如应用程序。存储器1932中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外，处理组件1922被配置为执行指令，以执行上述方法。

电子设备1900还可以包括一个电源组件1926被配置为执行电子设备1900的电源管理，一个有线或无线网络接口1950被配置为将电子设备1900连接到网络，和一个输入输出（I/O）
15 接口1958。电子设备1900可以操作基于存储在存储器1932的操作系统，例如微软服务器操作系统（Windows Server™），苹果公司推出的基于图形用户界面操作系统（Mac OS X™），多用户多进程的计算机操作系统（Unix™），自由和开放源代码的类Unix操作系统（Linux™），开放源代码的类Unix操作系统（FreeBSD™）或类似。

在示例性实施例中，还提供了一种非易失性计算机可读存储介质，例如包括计算机程序指令的存储器1932，上述计算机程序指令可由电子设备1900的处理组件1922执行以完成
20 上述方法。

本公开实施例可以是系统、方法和/或计算机程序产品。计算机程序产品可以包括计算机可读存储介质，其上载有用于使处理器实现本公开实施例的各个方面的计算机可读程序指令。

25 计算机可读存储介质可以是保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是——但不限于——电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子（非穷举的列表）包括：便携式计算机盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM或闪存）、静态随机存取存储器（SRAM）、便携式压缩盘只读存储器（CD-ROM）、数字多功能盘（DVD）、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身，诸如无线电波或者
30 其他自由传播的电磁波、通过波导或其他传输媒介传播的电磁波（例如，通过光纤电缆的光脉冲）、或者通过电线传输的电信号。

35 这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备，或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计算机可读程序指令，并转发该计算机可读程序指令，以供存储在各个计算/处理设备中的
40 计算机可读存储介质中。

用于执行本公开操作的计算机程序指令可以是汇编指令、指令集架构（ISA）指令、机器指令、机器相关指令、微代码、固件指令、状态设置数据、或者以一种或多种编程语言的任意组合编写的源代码或目标代码，所述编程语言包括面向对象的编程语言——诸如 Smalltalk、C++等，以及常规的过程式编程语言——诸如“C”语言或类似的编程语言。计算机
45 可读程序指令可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独

立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)——连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。在一些实施例中，通过利用计算机可读程序指令的状态信息来个性化定制电子电路，例如可编程逻辑电路、现场可编程门阵列（FPGA）或可编程逻辑阵列（PLA），该电子电路可以执行计算机可读程序指令，从而实现本公开的各个方面。

这里参照根据本公开实施例的方法、装置（系统）和计算机程序产品的流程图和/或框图描述了本公开的各个方面。应当理解，流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合，都可以由计算机可读程序指令实现。

这些计算机可读程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器，从而生产出一种机器，使得这些指令在通过计算机或其它可编程数据处理装置的处理器执行时，产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中，这些指令使得计算机、可编程数据处理装置和/或其他设备以特定方式工作，从而，存储有指令的计算机可读介质则包括一个制造品，其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

也可以把计算机可读程序指令加载到计算机、其它可编程数据处理装置、或其它设备上，使得在计算机、其它可编程数据处理装置或其它设备上执行一系列操作步骤，以产生计算机实现的过程，从而使得在计算机、其它可编程数据处理装置、或其它设备上执行的指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

附图中的流程图和框图显示了根据本公开的多个实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段或指令的一部分，所述模块、程序段或指令的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个连续的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或动作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

该计算机程序产品可以具体通过硬件、软件或其结合的方式实现。在一个可选实施例中，所述计算机程序产品具体体现为计算机存储介质，在另一个可选实施例中，计算机程序产品具体体现为软件产品，例如软件开发包(Software Development Kit, SDK)等等。

以上已经描述了本公开的各实施例，上述说明是示例性的，并非穷尽性的，并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择，旨在最好地解释各实施例的原理、实际应用或对市场中的技术的改进，或者使本技术领域的其它普通技术人员能理解本文披露的各实施例。

工业实用性

本公开实施例涉及一种网络训练方法及装置、目标检测方法及装置和电子设备。该网络训练方法包括：将未标注的样本图像输入目标检测网络中处理，得到目标检测结果，该结果包括目标的图像区域、特征信息及分类概率；根据目标的分类概率，确定目标的类别置信度；针对类别置信度大于或等于阈值的第一目标，将第一目标所在的样本图像作为已标注图像并加入训练集；针对类别置信度小于第一阈值的第二目标，对第二目标进行特征相关挖掘，从第二目标中确定出第四目标，将其所在的样本图像并加入训练集；根据训练

集中的样本图像训练目标检测网络。本公开实施例可提高目标检测网络的训练效果。

权利要求书

1、一种网络训练方法，包括：

将未标注的第一样本图像输入目标检测网络中处理，得到所述第一样本图像的目标检测结果，所述目标检测结果包括所述第一样本图像中目标的图像区域、特征信息及分类概率；

根据所述目标的分类概率，确定所述目标的类别置信度；

针对所述目标中类别置信度大于或等于第一阈值的所述第一目标，将所述第一目标所在的第一样本图像作为已标注的第二样本图像，并加入训练集中，其中，所述第二样本图像的标注信息包括所述第一目标的图像区域及与所述第一目标的类别置信度对应的类别，所述训练集中包括已标注的第三样本图像；

针对所述目标中类别置信度小于所述第一阈值的第二目标，根据所述第三样本图像中的第三目标的特征信息，对所述第二目标进行特征相关挖掘，通过特征相关挖掘，从所述第二目标中确定出第四目标及所述第四目标所在的第一样本图像，并将所述第四目标所在的第一样本图像作为第四样本图像，并加入所述训练集中；

根据所述第四样本图像的标注信息，所述训练集中的第二样本图像、第三样本图像及所述第四样本图像，训练所述目标检测网络。

2、根据权利要求1所述的方法，其中，所述根据所述第四样本图像的标注信息，所述训练集中的第二样本图像、第三样本图像及所述第四样本图像，训练所述目标检测网络，包括：

根据所述训练集的正样本图像中目标的类别，分别确定从各个类别的正样本图像中采样的第一数量，所述正样本图像为图像中包括目标的样本图像；

根据各个类别的正样本图像中采样的第一数量，对各个类别的正样本图像进行采样，得到多个第五样本图像；

对所述训练集的负样本图像进行采样，得到多个第六样本图像，所述负样本图像为图像中不包括目标的样本图像；

根据所述第五样本图像及所述第六样本图像，训练所述目标检测网络。

3、根据权利要求1或2所述的方法，其中，所述根据所述第三样本图像中的第三目标的特征信息，对所述第二目标进行特征相关挖掘，通过特征相关挖掘，从所述第二目标中确定出第四目标及所述第四目标所在的第一样本图像，包括：

根据所述第二目标的分类概率，确定所述第二目标的信息熵；

根据所述第二目标的类别置信度及信息熵，从所述第二目标中选择出第五目标；

根据所述第三样本图像中的第三目标的类别以及待挖掘的样本图像的总数量，分别确定各个类别待挖掘的样本图像的第二数量；

根据所述第三样本图像中的第三目标的特征信息，所述第五目标的特征信息以及各个类别待挖掘的样本图像的第二数量，从所述第五目标中确定出第四目标及所述第四目标所在的第一样本图像。

4、根据权利要求3所述的方法，其中，根据所述第二目标的类别置信度及信息熵，从所述第二目标中选择出第五目标，包括：

根据所述第二目标的类别置信度及信息熵，分别对所述第二目标进行排序，选择出第三数量的第六目标和第四数量的第七目标；

对所述第六目标和所述第七目标进行合并，得到所述第五目标。

5、根据权利要求3或4所述的方法，其中，根据所述第三样本图像中的第三目标的类别以及待挖掘的样本图像的总数量，分别确定各个类别待挖掘的样本图像的第二数量，包括：

根据所述第三样本图像中的第三目标的类别，确定各个类别的第三目标的比例；

根据各个类别的第三目标的比例，确定各个类别的抽样比重；

根据各个类别的抽样比重，分别确定各个类别待挖掘的样本图像的第二数量。

- 5 6、根据权利要求3-5中任意一项所述的方法，其中，根据所述第三样本图像中的第三目标的特征信息，所述第五目标的特征信息以及各个类别待挖掘的样本图像的第二数量，从所述第五目标中确定出第四目标及所述第四目标所在的第一样本图像，包括：

根据第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离，分别确定所述第一类别的第三目标中与各个第五目标距离最小的第三目标，并作为第八目标，所述第一类别为第三目标的类别中的任意一个；

- 10 将所述第八目标中距离最大的目标，确定为第四目标。

7、根据权利要求6所述的方法，其中，根据所述第三样本图像中的第三目标的特征信息，所述第五目标的特征信息以及各个类别待挖掘的样本图像的第二数量，从所述第五目标中确定出第四目标及所述第四目标所在的第一样本图像，还包括：

- 15 将确定出的第四目标添加到所述第一类别的第三目标中，并将所述确定出的第四目标从未标注的第五目标中移除。

8、根据权利要求1-7中任意一项所述的方法，其中，所述方法还包括：

将所述第三样本图像输入所述目标检测网络中处理，得到所述第三样本图像中的第三目标的特征信息。

- 20 9、根据权利要求1-8中任意一项所述的方法，其中，在所述将未标注的第一样本图像输入目标检测网络中处理，得到所述第一样本图像的目标检测结果的步骤之前，所述方法还包括：

通过已标注的第三样本图像对所述目标检测网络进行预训练。

10、根据权利要求1-9中任意一项所述的方法，其中，所述第一样本图像包括长尾图像。

- 25 11、根据权利要求2所述的方法，其中，在所述根据所述训练集的正样本图像中目标的类别，分别确定从各个类别的正样本图像中采样的第一数量的步骤之前，所述方法包括：

对训练集中的正样本图像和负样本图像进行采样，得到数量相同或相近的正样本图像和负样本图像。

12、根据权利要求3所述的方法，其中，所述待挖掘的样本图像的总数量为所述第一样本图像的总数量的5%~25%。

- 30 13、根据权利要求4所述的方法，其中，所述对所述第六目标和所述第七目标进行合并，得到所述第五目标，包括：

去除所述第六目标中与所述第七目标相同的目标，得到所述第六目标中与所述第七目标不同的剩余目标；

将所述剩余目标和所述第七目标作为所述第五目标。

- 35 14、根据权利要求6所述的方法，其中，在所述根据第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离，分别确定所述第一类别的第三目标中与各个第五目标距离最小的第三目标，并作为第八目标的步骤之后，所述方法还包括：

在所述第四目标所在的第一样本图像的数量达到所述第一类别的待挖掘的样本图像的第二数量时，结束对所述第八目标的确定。

- 40 15、根据权利要求7所述的方法，其中，在所述根据第一类别的第三目标的特征信息与各个第五目标的特征信息之间的距离，分别确定所述第一类别的第三目标中与各个第五目标距离最小的第三目标，并作为第八目标的步骤之后，所述方法还包括：

在所述第四目标所在的第一样本图像的数量未达到所述第一类别的待挖掘的样本图像的第二数量，且存储所述第五目标的特征信息的集合为空时，结束对所述第八目标的确定。

45

16、根据权利要求8所述的方法，其中，所述将所述第三样本图像输入所述目标检测网络中处理，得到所述第三样本图像中的第三目标的特征信息，包括：

将所述第三样本图像，输入所述目标检测网络中，得到所述目标检测网络的隐藏层输出的特征向量；

5 将所述特征向量确定为所述第三目标的特征信息。

17、一种目标检测方法，所述方法包括：

将待处理图像输入目标检测网络中处理，得到所述待处理图像的目标检测结果，所述目标检测结果包括所述待处理图像中目标的位置和类别，

所述目标检测网络是根据权利要求1-10中任意一项的网络训练方法训练得到的。

10 18、一种网络训练装置，包括：

目标检测部分，被配置为将未标注的第一样本图像输入目标检测网络中处理，得到所述第一样本图像的目标检测结果，所述目标检测结果包括所述第一样本图像中目标的图像区域、特征信息及分类概率；

置信度确定部分，被配置为根据所述目标的分类概率，确定所述目标的类别置信度；

15 标注部分，被配置为针对所述目标中类别置信度大于或等于第一阈值的所述第一目标，将所述第一目标所在的第一样本图像作为已标注的第二样本图像，并加入训练集中，其中，所述第二样本图像的标注信息包括所述第一目标的图像区域及与所述第一目标的类别置信度对应的类别，所述训练集中包括已标注的第三样本图像；

20 特征挖掘部分，被配置为针对所述目标中类别置信度小于所述第一阈值的第二目标，根据所述第三样本图像中的第三目标的特征信息，对所述第二目标进行特征相关挖掘，通过特征相关挖掘，从所述第二目标中确定出第四目标及所述第四目标所在的第一样本图像，并将所述第四目标所在的第一样本图像作为第四样本图像，并加入所述训练集中；

训练部分，被配置为根据所述第四样本图像的标注信息，所述训练集中的第二样本图像、第三样本图像及所述第四样本图像，训练所述目标检测网络。

25 19、一种电子设备，包括：

处理器；

被配置为存储处理器可执行指令的存储器；

其中，所述处理器被配置为调用所述存储器存储的指令，以执行权利要求1至16中任意一项所述的网络训练方法，或执行权利要求17所述的目标检测方法。

30 20、一种计算机可读存储介质，其上存储有计算机程序指令，所述计算机程序指令被处理器执行时实现权利要求1至16中任意一项所述的方法网络训练方法，或实现权利要求17所述的目标检测方法。

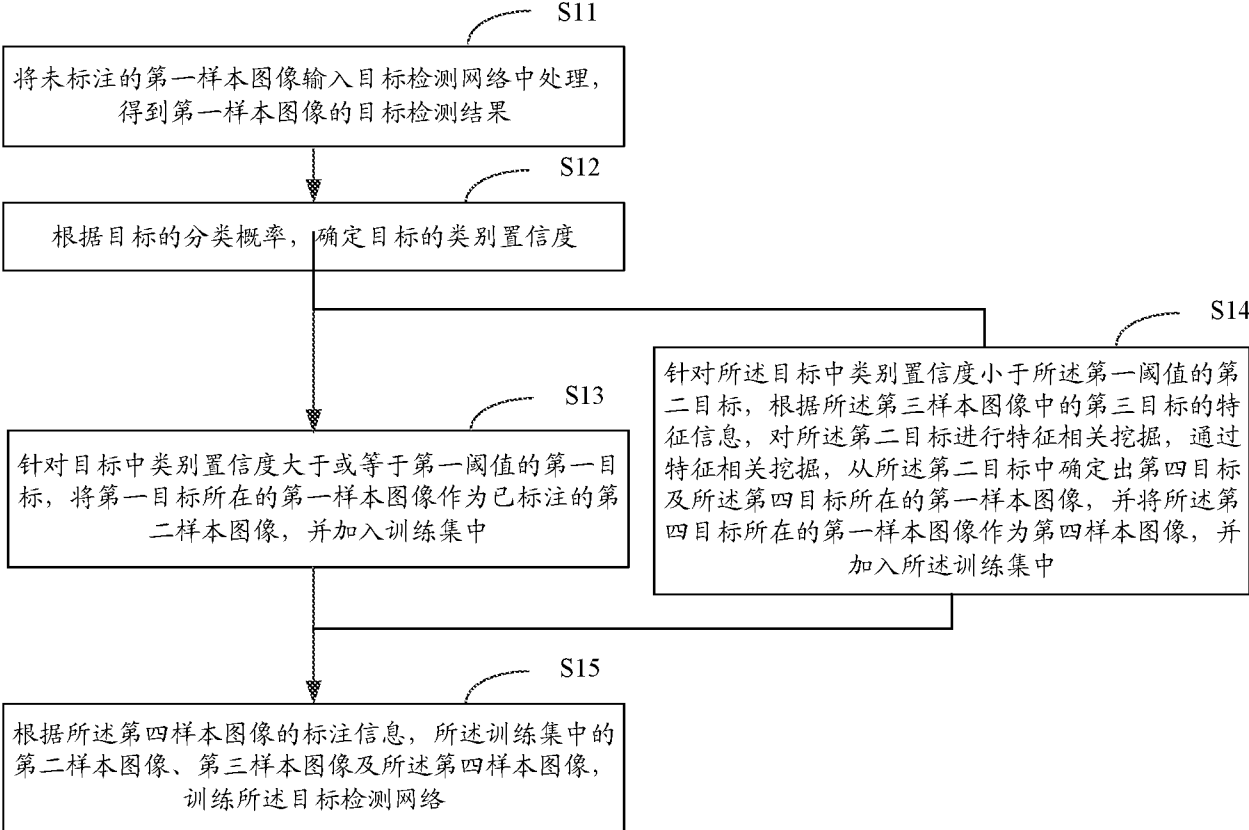


图 1

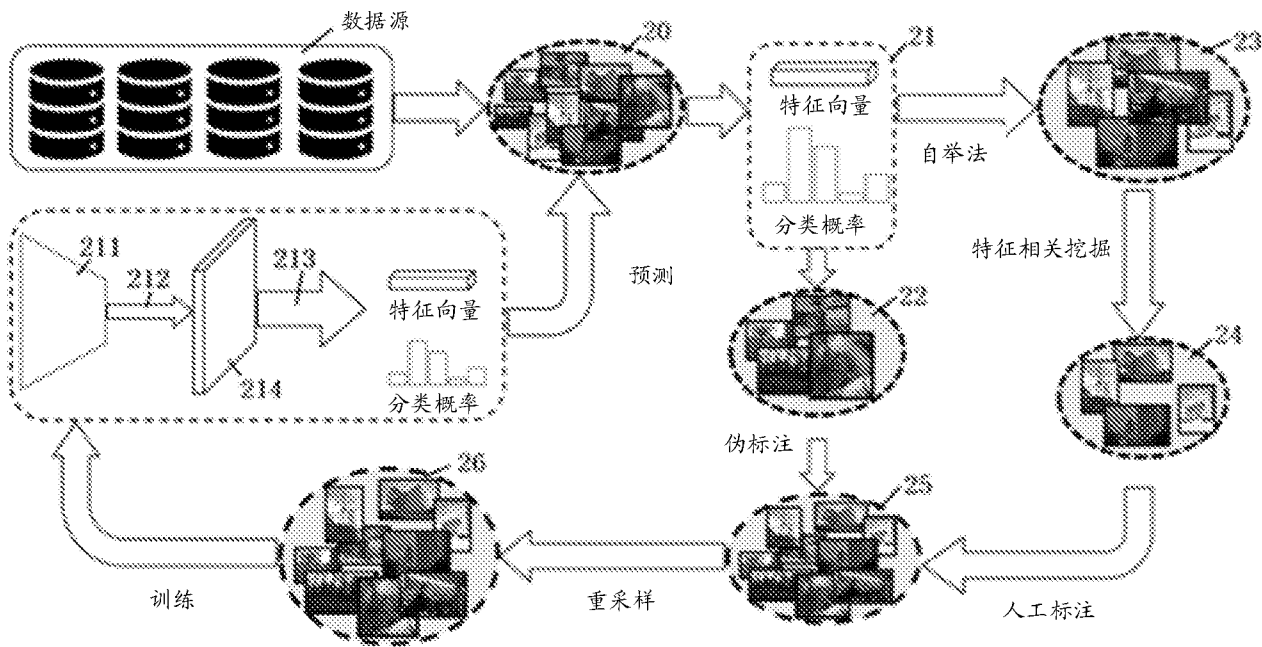


图 2

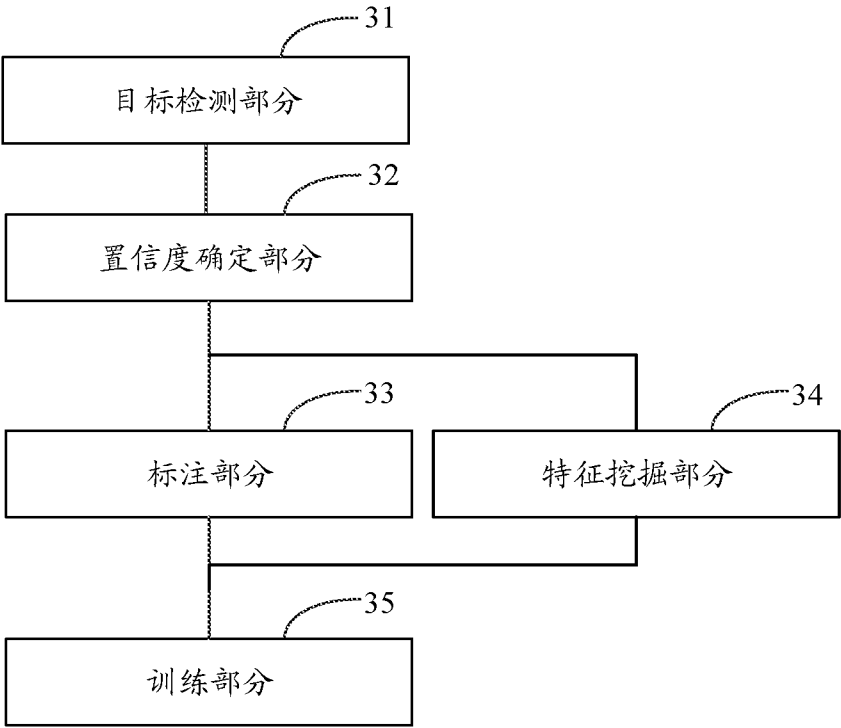


图 3

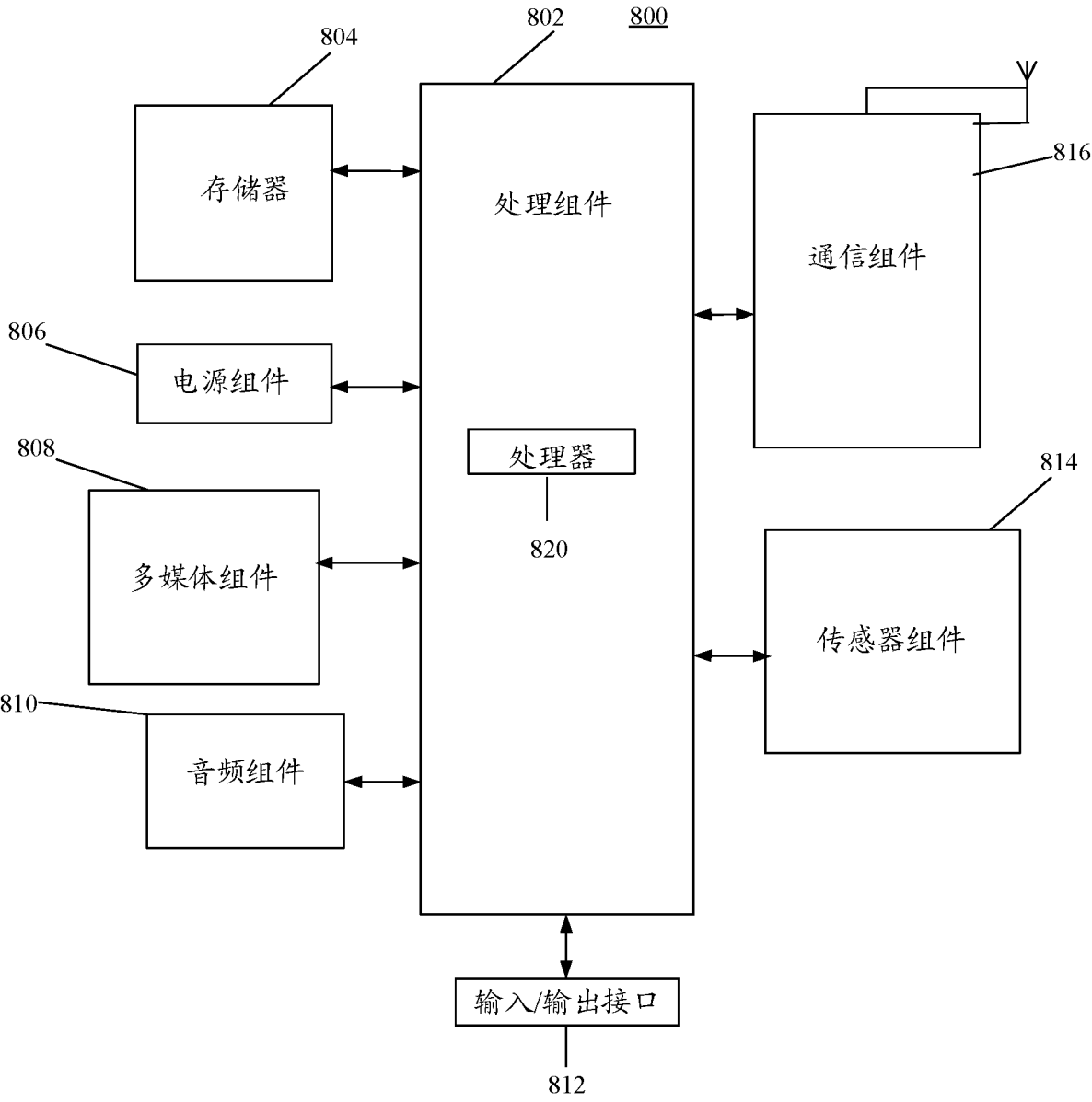


图 4

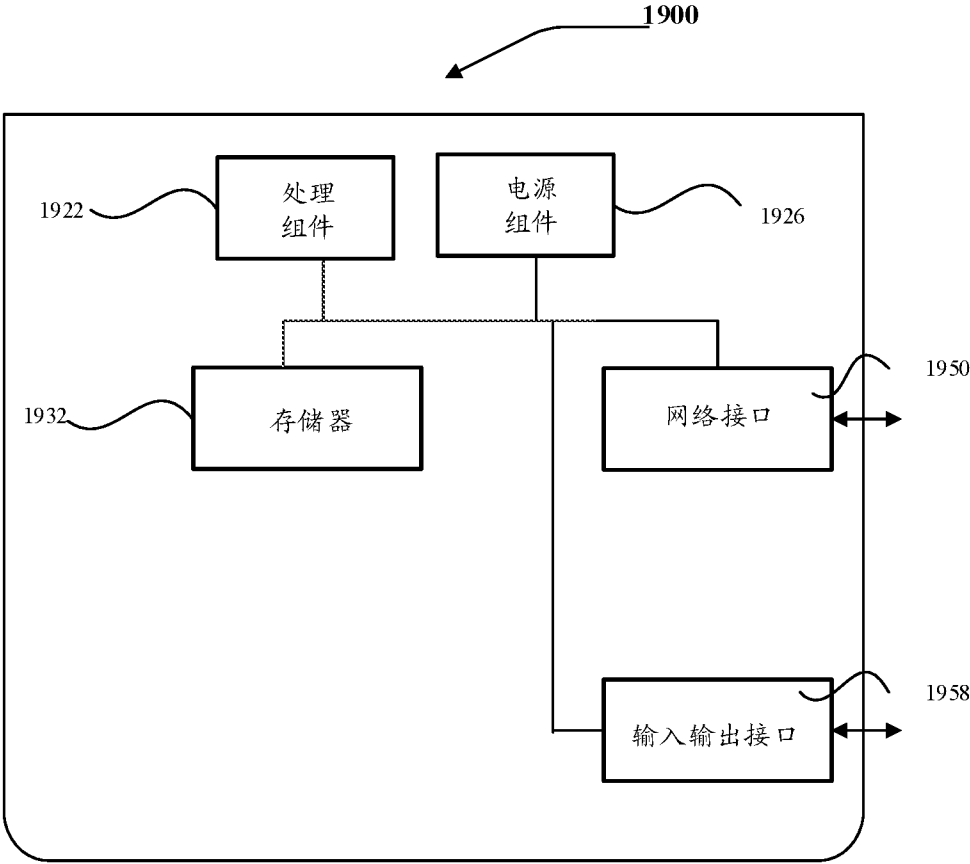


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/125972

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)

CNABS, SIPOABS, CNTXT, CNKI, DWPI, USTXT: 训练, 半监督, 标记, 标注, 置信度, 阈值, 特征, 相似度, 距离, 熵;
training, semi-supervised, label, annotation, confidence, threshold, feature, similarity, distance, entropy.**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 104318242 A (AIR FORCE ENGINEERING UNIVERSITY, PLA) 28 January 2015 (2015-01-28) description, paragraphs [0107]-[0222], and figure 1	1-20
Y	CN 108985334 A (GUANGZHOU DEEP DOMAIN INFORMATION TECHNOLOGY CO., LTD.) 11 December 2018 (2018-12-11) description, paragraphs [0051]-[0067], and figures 1-4	1-20
A	CN 104156438 A (DEZHOU UNIVERSITY) 19 November 2014 (2014-11-19) entire document	1-20
A	US 2019073447 A1 (IBM) 07 March 2019 (2019-03-07) entire document	1-20
A	US 2019378044 A1 (GROUPON, INC.) 12 December 2019 (2019-12-12) entire document	1-20



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

08 April 2021

Date of mailing of the international search report

16 April 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/125972

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	104318242	A	28 January 2015	None			
CN	108985334	A	11 December 2018	None			
CN	104156438	A	19 November 2014	None			
US	2019073447	A1	07 March 2019	US	10606982	B2	31 March 2020
US	2019378044	A1	12 December 2019	None			

国际检索报告

国际申请号

PCT/CN2020/125972

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))

CNABS, SIP0ABS, CNTXT, CNKI, DWPI, USTXT: 训练, 半监督, 标记, 标注, 置信度, 阈值, 特征, 相似度, 距离, 熵; training, semi-supervised, label, annotation, confidence, threshold, feature, similarity, distance, entropy.

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
Y	CN 104318242 A (中国人民解放军空军工程大学) 2015年 1月 28日 (2015 - 01 - 28) 说明书[0107]-[0222]段, 图1	1-20
Y	CN 108985334 A (广州深域信息科技有限公司) 2018年 12月 11日 (2018 - 12 - 11) 说明书[0051]-[0067]段, 图1-4	1-20
A	CN 104156438 A (德州学院) 2014年 11月 19日 (2014 - 11 - 19) 全文	1-20
A	US 2019073447 A1 (IBM) 2019年 3月 7日 (2019 - 03 - 07) 全文	1-20
A	US 2019378044 A1 (GROUPON, INC.) 2019年 12月 12日 (2019 - 12 - 12) 全文	1-20

☐ 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

2021年 4月 8日

国际检索报告邮寄日期

2021年 4月 16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

受权官员

张会

电话号码 62412328

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/125972

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	104318242	A	2015年 1月 28日	无	
CN	108985334	A	2018年 12月 11日	无	
CN	104156438	A	2014年 11月 19日	无	
US	2019073447	A1	2019年 3月 7日	US 10606982 B2	2020年 3月 31日
US	2019378044	A1	2019年 12月 12日	无	