



(12) 发明专利

(10) 授权公告号 CN 110210561 B

(45) 授权公告日 2022.04.01

(21) 申请号 201910473459.6

G06V 10/44 (2022.01)

(22) 申请日 2019.05.31

G06N 3/08 (2006.01)

(65) 同一申请的已公布的文献号

审查员 雷皓婷

申请公布号 CN 110210561 A

(43) 申请公布日 2019.09.06

(73) 专利权人 北京市商汤科技开发有限公司

地址 100084 北京市海淀区中关村东路1号
院3号楼7层710-712房间

(72) 发明人 祝新革 庞江淼 杨策元 石建萍
林达华

(74) 专利代理机构 北京派特恩知识产权代理有
限公司 11270

代理人 李梅香 张颖玲

(51) Int. Cl.

G06K 9/62 (2022.01)

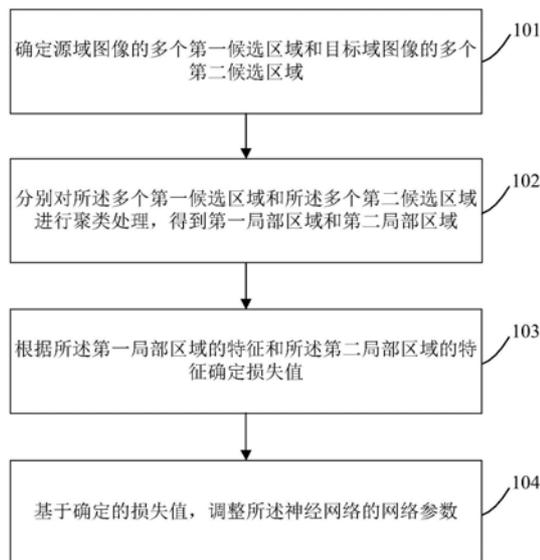
权利要求书3页 说明书16页 附图3页

(54) 发明名称

神经网络的训练方法、目标检测方法及装置、存储介质

(57) 摘要

本申请公开了一种神经网络的训练方法及装置、存储介质,其中,所述的方法包括:确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域;分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理,得到第一局部区域和第二局部区域;根据所述第一局部区域的特征和所述第二局部区域的特征确定损失值;基于确定的损失值,调整所述神经网络的网络参数。



1. 一种神经网络的训练方法,其特征在于,所述方法包括:
 - 确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域;
 - 分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理,得到第一局部区域和第二局部区域;
 - 根据各个第一局部区域的特征,重建各个第一局部区域的图像;
 - 根据各个第二局部区域的特征,重建各个第二局部区域的图像;其中,所述第一局部区域与所述第二局部区域的区域大小相等;
 - 从源域图像中抠取各个第一局部区域位置处的图像,得到真实的各个第一局部区域的图像;
 - 从目标域图像中抠取各个第二局部区域位置处图像,得到真实的各个第二局部区域的图像;
 - 对重建的各个第一局部区域的图像、真实的各个第一局部区域的图像、重建的各个第二局部区域的图像以及真实的各个第二局部区域的图像进行真伪判别;
 - 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失以及真伪判别的损失确定损失值;
 - 基于确定的损失值,调整所述神经网络的网络参数。
2. 根据权利要求1所述的方法,其特征在于,所述方法还包括:
 - 利用所述神经网络对源域图像进行处理,得到处理结果;
 - 根据得到的处理结果和源域图像的标注结果,调整所述神经网络的网络参数。
3. 根据权利要求1所述的方法,其特征在于,所述确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域,包括:
 - 基于源域图像得到关于所述源域图像的第一特征表示,根据所述第一特征表示确定所述源域图像的多个第一候选区域;
 - 基于目标域图像得到关于所述目标域图像的第二特征表示,根据所述第二特征表示确定所述目标域图像的多个第二候选区域。
4. 根据权利要求3所述的方法,其特征在于,在根据所述第一特征表示确定所述源域图像的多个第一候选区域,根据所述第二特征表示确定所述目标域图像的多个第二候选区域之后,所述方法还包括:
 - 对所述源域图像的第一特征表示和所述源域图像的多个第一候选区域进行池化处理,以使得各个第一候选区域的特征的维度相同;
 - 对所述目标域图像的第二特征表示和所述目标域图像的多个第二候选区域进行池化处理,以使得各个第二候选区域的特征的维度相同;
 - 其中,在进行池化处理后,第一候选区域的特征的维度与第二候选区域的特征的维度相同。
5. 根据权利要求3或4所述的方法,其特征在于,所述分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理,得到第一局部区域和第二局部区域,包括:
 - 对所述多个第一候选区域进行聚类处理,得到K个第一聚类中心;其中,所述K为正整数;
 - 根据各个第一聚类中心,确定各个第一局部区域;

对所述多个第二候选区域进行聚类处理,得到K个第二聚类中心;

根据各个第二聚类中心,确定各个第二局部区域。

6. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

为各个第二局部区域赋予权重;其中,一个第二局部区域的权重表征该第二局部区域包含所述目标域图像中的目标的概率;

根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失以及真伪判别的损失确定所述损失值,包括:

根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失、真伪判别的损失以及各个第二局部区域的权重,确定所述损失值。

7. 一种目标检测方法,其特征在于,所述方法包括:

获取目标域图像;

利用目标检测网络对所述目标域图像进行处理,得到所述目标域图像中各个目标的位置和/或分类;

其中,所述目标检测网络采用权利要求1至6任一所述的方法训练。

8. 一种神经网络的训练装置,其特征在于,所述装置包括:

确定模块,用于确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域;

聚类模块,用于分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理,得到第一局部区域和第二局部区域;

对齐模块,用于根据各个第一局部区域的特征,重建各个第一局部区域的图像;根据各个第二局部区域的特征,重建各个第二局部区域的图像;其中,所述第一局部区域与所述第二局部区域的区域大小相等;从源域图像中抠取各个第一局部区域位置处的图像,得到真实的各个第一局部区域的图像;从目标域图像中抠取各个第二局部区域位置处图像,得到真实的各个第二局部区域的图像;对重建的各个第一局部区域的图像、真实的各个第一局部区域的图像、重建的各个第二局部区域的图像以及真实的各个第二局部区域的图像进行真伪判别;根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失以及真伪判别的损失确定损失值;

训练模块,用于基于确定的损失值,调整所述神经网络的网络参数。

9. 根据权利要求8所述的装置,其特征在于,所述训练模块,还用于:

利用所述神经网络对源域图像进行处理,得到处理结果;

根据得到的处理结果和源域图像的标注结果,调整所述神经网络的网络参数。

10. 根据权利要求8所述的装置,其特征在于,所述确定模块,还用于:

基于源域图像得到关于所述源域图像的第一特征表示,根据所述第一特征表示确定所述源域图像的多个第一候选区域;

基于目标域图像得到关于所述目标域图像的第二特征表示,根据所述第二特征表示确定所述目标域图像的多个第二候选区域。

11. 根据权利要求10所述的装置,其特征在于,所述装置还包括:

池化模块,用于:

在根据所述第一特征表示确定所述源域图像的多个第一候选区域,根据所述第二特征

表示确定所述目标域图像的多个第二候选区域之后，

对所述源域图像的第一特征表示和所述源域图像的多个第一候选区域进行池化处理，以使得各个第一候选区域的特征的维度相同；

对所述目标域图像的第二特征表示和所述目标域图像的多个第二候选区域进行池化处理，以使得各个第二候选区域的特征的维度相同；

其中，在进行池化处理后，第一候选区域的特征的维度与第二候选区域的特征的维度相同。

12. 根据权利要求10或11所述的装置，其特征在于，所述聚类模块，还用于：

对所述多个第一候选区域进行聚类处理，得到K个第一聚类中心；其中，所述K为正整数；

根据各个第一聚类中心，确定各个第一局部区域；

对所述多个第二候选区域进行聚类处理，得到K个第二聚类中心；

根据各个第二聚类中心，确定各个第二局部区域。

13. 根据权利要求8所述的装置，其特征在于，所述装置还包括权重分配模块，用于：

为各个第二局部区域赋予权重；其中，一个第二局部区域的权重表征该第二局部区域包含所述目标域图像中的目标的概率；

所述对齐模块，还用于：

根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失、真伪判别的损失以及各个第二局部区域的权重，确定所述损失值。

14. 一种目标检测装置，其特征在于，所述装置包括：

获取模块，用于获取目标域图像；

检测模块，用于利用目标检测网络对所述目标域图像进行处理，得到所述目标域图像中各个目标的位置和/或分类；

其中，所述目标检测网络采用权利要求1至6任一所述的方法训练。

15. 一种神经网络的训练装置，所述装置包括：存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述程序时实现权利要求1至6任一项所述的神经网络的训练方法。

16. 一种目标检测装置，所述装置包括：存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述程序时实现权利要求7所述的目标检测方法。

17. 一种存储介质，所述存储介质存储有计算机程序，所述计算机程序被处理器执行时，能够使得所述处理器执行权利要求1至6任一项所述的神经网络的训练方法。

18. 一种存储介质，所述存储介质存储有计算机程序，所述计算机程序被处理器执行时，能够使得所述处理器执行权利要求7所述的目标检测方法。

神经网络的训练方法、目标检测方法及装置、存储介质

技术领域

[0001] 本申请涉及计算机视觉技术领域，具体涉及一种神经网络的训练方法及装置、存储介质，目标检测方法及装置、存储介质。

背景技术

[0002] 目前的物体检测方法往往是在单一环境下的数据集中进行训练，大量的数据使得这些方法能够在数据集中取得不错的效果，但是同时也限制了他们的泛化能力，即在面临不同的环境时，检测效果往往会大打折扣。在物体分类任务中，通过域适应方法来增强泛化能力已经是一种常见做法，但是在物体检测任务中，这些分类中的域适应方法的效果不尽如人意。

发明内容

[0003] 本申请提供一种神经网络的训练方法及应用训练好的神经网络进行目标检测的技术方案。

[0004] 第一方面，本申请实施例提供了一种神经网络的训练方法，所述方法包括：

[0005] 确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域；

[0006] 分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理，得到第一局部区域和第二局部区域；

[0007] 根据所述第一局部区域的特征和所述第二局部区域的特征确定损失值；

[0008] 基于确定的损失值，调整所述神经网络的网络参数。

[0009] 上述方案中，可选地，所述方法还包括：

[0010] 利用所述神经网络对源域图像进行处理，得到处理结果；

[0011] 根据得到的处理结果和源域图像的标注结果，调整所述神经网络的网络参数。

[0012] 上述方案中，可选地，所述确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域，包括：

[0013] 基于源域图像得到关于所述源域图像的第一特征表示，根据所述第一特征表示确定所述源域图像的多个第一候选区域；

[0014] 基于目标域图像得到关于所述目标域图像的第二特征表示，根据所述第二特征表示确定所述目标域图像的多个第二候选区域。

[0015] 上述方案中，可选地，在根据所述第一特征表示确定所述源域图像的多个第一候选区域，根据所述第二特征表示确定所述目标域图像的多个第二候选区域之后，所述方法还包括：

[0016] 对所述源域图像的第一特征表示和所述源域图像的多个第一候选区域进行池化处理，以使得各个第一候选区域的特征的维度相同；

[0017] 对所述目标域图像的第二特征表示和所述目标域图像的多个第二候选区域进行池化处理，以使得各个第二候选区域的特征的维度相同；

[0018] 其中,在进行池化处理,第一候选区域的特征的维度与第二候选区域的特征的维度相同。

[0019] 上述方案中,可选地,所述分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理,得到第一局部区域和第二局部区域,包括:

[0020] 对所述多个第一候选区域进行聚类处理,得到K个第一聚类中心;其中,所述K为正整数;

[0021] 根据各个第一聚类中心,确定各个第一局部区域;

[0022] 对所述多个第二候选区域进行聚类处理,得到K个第二聚类中心;

[0023] 根据各个第二聚类中心,确定各个第二局部区域;

[0024] 根据所述第一局部区域的特征和所述第二局部区域的特征确定损失值,包括:

[0025] 根据各个第一局部区域的特征和各个第二局部区域的特征确定损失值。

[0026] 上述方案中,可选地,根据各个第一局部区域的特征和各个第二局部区域的特征确定损失值,包括:

[0027] 根据各个第一局部区域的特征,重建各个第一局部区域的图像;

[0028] 根据各个第二局部区域的特征,重建各个第二局部区域的图像;其中,所述第一局部区域与第二局部区域的区域大小相等;

[0029] 从源域图像中抠取各个第一局部区域位置处的图像,得到真实的各个第一局部区域的图像;

[0030] 从目标域图像中抠取各个第二局部区域位置处图像,得到真实的各个第二局部区域的图像;

[0031] 对重建的各个第一局部区域的图像、真实的各个第一局部区域的图像、重建的各个第二局部区域的图像以及真实的各个第二局部区域的图像进行真伪判别;

[0032] 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失以及真伪判别的损失确定所述损失值

[0033] 上述方案中,可选地,所述方法还包括:

[0034] 为各个第二局部区域赋予权重;其中,一个第二局部区域的权重表征该第二局部区域包含所述目标域图像中的目标的概率;

[0035] 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失以及真伪判别的损失确定所述损失值,包括:

[0036] 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失、真伪判别的损失以及各个第二局部区域的权重,确定所述损失值。

[0037] 第二方面,本申请实施例提供了一种目标检测方法,所述方法包括:

[0038] 获取目标域图像;

[0039] 利用目标检测网络对所述目标域图像进行处理,得到所述目标域图像中各个目标的位置和/或分类;

[0040] 其中,所述目标检测网络采用上文所述的神经网络训练方法训练。

[0041] 第三方面,本申请实施例提供了一种神经网络的训练装置,所述装置包括:

[0042] 确定模块,用于确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域;

- [0043] 聚类模块,用于分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理,得到第一局部区域和第二局部区域;
- [0044] 对齐模块,用于根据所述第一局部区域的特征和所述第二局部区域的特征确定损失值;
- [0045] 训练模块,用于基于确定的损失值,调整所述神经网络的网络参数。
- [0046] 上述方案中,可选地,所述训练模块,还用于:
- [0047] 利用所述神经网络对源域图像进行处理,得到处理结果;
- [0048] 根据得到的处理结果和源域图像的标注结果,调整所述神经网络的网络参数。
- [0049] 上述方案中,可选地,所述确定模块,还用于:
- [0050] 基于源域图像得到关于所述源域图像的第一特征表示,根据所述第一特征表示确定所述源域图像的多个第一候选区域;
- [0051] 基于目标域图像得到关于所述目标域图像的第二特征表示,根据所述第二特征表示确定所述目标域图像的多个第二候选区域。
- [0052] 上述方案中,可选地,所述装置还包括:
- [0053] 池化模块,用于:
- [0054] 在根据所述第一特征表示确定所述源域图像的多个第一候选区域,根据所述第二特征表示确定所述目标域图像的多个第二候选区域之后,
- [0055] 对所述源域图像的第一特征表示和所述源域图像的多个第一候选区域进行池化处理,以使得各个第一候选区域的特征的维度相同;
- [0056] 对所述目标域图像的第二特征表示和所述目标域图像的多个第二候选区域进行池化处理,以使得各个第二候选区域的特征的维度相同;
- [0057] 其中,在进行池化处理后,第一候选区域的特征的维度与第二候选区域的特征的维度相同。
- [0058] 上述方案中,可选地,所述聚类模块,还用于:
- [0059] 对所述多个第一候选区域进行聚类处理,得到K个第一聚类中心;其中,所述K为正整数;
- [0060] 根据各个第一聚类中心,确定各个第一局部区域;
- [0061] 对所述多个第二候选区域进行聚类处理,得到K个第二聚类中心;
- [0062] 根据各个第二聚类中心,确定各个第二局部区域;
- [0063] 所述对齐模块,还用于:
- [0064] 根据各个第一局部区域的特征和各个第二局部区域的特征确定损失值。
- [0065] 上述方案中,可选地,所述对齐模块,还用于:
- [0066] 根据各个第一局部区域的特征,重建各个第一局部区域的图像;
- [0067] 根据各个第二局部区域的特征,重建各个第二局部区域的图像;其中,所述第一局部区域与所述第二局部区域的区域大小相等;
- [0068] 从源域图像中抠取各个第一局部区域位置处的图像,得到真实的各个第一局部区域的图像;
- [0069] 从目标域图像中抠取各个第二局部区域位置处图像,得到真实的各个第二局部区域的图像;

[0070] 对重建的各个第一局部区域的图像、真实的各个第一局部区域的图像、重建的各个第二局部区域的图像以及真实的各个第二局部区域的图像进行真伪判别；

[0071] 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失以及真伪判别的损失确定所述损失值。

[0072] 上述方案中,可选地,所述装置还包括权重分配模块,用于:

[0073] 为各个第二局部区域赋予权重;其中,一个第二局部区域的权重表征该第二局部区域包含所述目标域图像中的目标的概率;

[0074] 所述对齐模块,还用于:

[0075] 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失、真伪判别的损失以及各个第二局部区域的权重,确定所述损失值。

[0076] 第四方面,本申请实施例提供了一种目标检测装置,所述装置包括:

[0077] 获取模块,用于获取目标域图像;

[0078] 检测模块,用于利用目标检测网络对所述目标域图像进行处理,得到所述目标域图像中各个目标的位置和/或分类;

[0079] 其中,所述目标检测网络采用权利要求上文所述的方法训练。

[0080] 第五方面,本申请实施例提供了一种神经网络的训练装置,所述装置包括:存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现本申请实施例所述的神经网络的训练方法的步骤。

[0081] 第六方面,本申请实施例提供了一种存储介质,所述存储介质存储有计算机程序,所述计算机程序被处理器执行时,使得所述处理器执行本申请实施例所述的神经网络的训练方法的步骤。

[0082] 第七方面,本申请实施例提供了一种目标检测装置,所述装置包括:存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现本申请实施例所述的目标检测方法的步骤。

[0083] 第八方面,本申请实施例提供了一种存储介质,所述存储介质存储有计算机程序,所述计算机程序被处理器执行时,使得所述处理器执行本申请实施例所述的目标检测方法的步骤。

[0084] 本申请提供的技术方案,确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域;分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理,得到第一局部区域和第二局部区域;根据所述第一局部区域的特征和所述第二局部区域的特征确定损失值;基于确定的损失值,调整所述神经网络的网络参数。由于图像的特征是通过神经网络提取的,神经网络从源域图像中提取到的特征和从目标域中提取到的特征越接近,那么神经网络就越能提取到源域图像和目标域图像中的共性的分布。如此,相对于现有的域适应方法要将整个图像进行对齐处理(所谓对齐也就是使得神经网络能够从源域图像中提取到的特征和能够从目标域中提取到的特征尽量接近)时会引入大量噪音而言,本申请能通过仅将局部区域图像的对齐,从而避免了对整个图像对齐时由大量背景信息所引入的噪音,并促进神经网络方法的泛化能力,即通过源域图像和目标域图像中包含目标图像的区域的对齐,使得物体检测网络能够提取到源域图像和目标域图像中的共性的分布,提高神经网络对目标域图像的适应能力,进而提高神经网络在目标域中的检测效果。

附图说明

[0085] 图1为本申请实施例提供的一种神经网络的训练方法的实现流程示意图；

[0086] 图2为本申请实施例提供的基于局部对齐的自适应物体检测神经网络架构示意图；

[0087] 图3为本申请实施例提供的聚类方案和加权估计的工作流程示意图，图3(a)→3(b)表示聚类操作的过程，3(b)→3(c)表示加权估计的过程；

[0088] 图4为本申请实施例提供的一种神经网络的训练装置的组成结构示意图。

具体实施方式

[0089] 为了更好的解释本申请，下面，先介绍一些现有技术中的物体检测方法。

[0090] 在过去几年中，深度学习的进步显著地推动了计算机视觉中各种任务的发展，如目标检测和语义分割。然而，应该注意的是，这种显著的进展在很大程度上依赖于大规模的训练数据。虽然已经存在多个公共基准，但这些基准涵盖的场景范围非常有限。在实际部署中，环境条件(如成像传感器、天气和照明)的变化可能会导致结果大幅下降。

[0091] 解决这个问题一个自然想法是随着领域的变化获得新的训练数据。不幸的是，由于大规模标注需要巨大的成本，这种方法在实践中并不总是可行的。物体检测或实例分割的成本特别高，因为它需要详细的标注，例如单个物体上的边界框或掩码。另一个吸引人的选择是无监督的域适应，即将在标准数据集上训练的模型适应新的域(通常称为目标域)，但不需要标注目标域的数据。沿着这条线发展了各种方法，在图像分类和语义分割上显示出令人鼓舞的结果。然而，如何有效地应用到物体检测上仍然是一个广泛存在的问题。

[0092] 最先进的物体检测器通常在公共数据集上进行训练。当应用到不同的领域时，如果图像有显著差异，并且相应的标注结果不可用(或获取成本高昂)，则会导致检测效果显著下降。一个自然的补救办法是通过对齐两个域上的图像表示来调整物体检测网络。例如，这可以通过对抗性学习来实现，并且在图像分类等任务中被证明是有效的。然而，我们发现在物体检测中，这种方法的改进是非常有限的。一个重要的原因是，传统的域自适应方法力求将复杂背景图像作为一个整体进行对齐，而物体检测本质上侧重于可能包含感兴趣物体的局部区域。

[0093] 基于此，本申请提出了一种能够进行域的自适应的目标检测的方案。

[0094] 下面结合附图和具体实施例对本申请的目标检测的方案中用到的神经网络的训练方法进行详细阐述。

[0095] 本申请实施例提供一种神经网络的训练方法，如图1所示，所述方法主要包括：

[0096] 步骤101、确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域。

[0097] 本申请中，考虑一个涉及两个领域的问题，源域和目标域。具体地，对于源域，可提供带标注结果的图像样本；对于目标域，只提供图像样本，对图像样本没有进行标注。如此，期望训练一个神经网络，能够从两个域的图像中提取到接近的特征，从而将神经网络很好地推广到目标域；而该神经网络事先已经在标准数据集上进行过训练。

[0098] 这里，所述源域图像和所述目标域图像可以是在同一环境条件或不同环境条件所采集到的关于同一对象或同一地点的图像，但本申请实施例对具体如何获取源域图像和目标域图像的方式不作限定。需要说明的是，所述对象可以是静态的，还可以是动态的。

[0099] 例如,所述源域图像可以是在第一环境条件下采集的关于第一地点的图像,所述目标域图像可以是在第二环境条件下采集的关于第一地点的图像,其中,所述第一地点为广场1,第一环境条件为晴天环境,第二环境条件为雾状环境

[0100] 例如,所述源域图像可以是在第一环境条件下采集的关于第一对象的图像,所述目标域图像可以是在第二环境条件下采集的关于第一对象的图像,其中,所述第一对象为车辆,第一环境条件为晴天环境,第二环境条件为雨天环境。

[0101] 在一些可选实现方式中,所述确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域,包括:

[0102] 基于源域图像得到关于所述源域图像的第一特征表示,根据所述第一特征表示确定所述源域图像的多个第一候选区域;

[0103] 基于目标域图像得到关于所述目标域图像的第二特征表示,根据所述第二特征表示确定所述目标域图像的多个第二候选区域。

[0104] 本申请实施例中,为便后续对局部区域的对齐处理,候选区域的区域大小是固定的。但候选区域可以随意地在图像上进行分配。

[0105] 在一些实现方式中,步骤101可以利用主干网络(Backbone Network)和区域生成网络(RPN,Region Proposal Network)来实现。例如,先将所述源域图像和所述目标域图像输入到主干网络中进行处理,通过所述主干网络得到关于所述源域图像的第一特征表示和关于所述目标域图像的第二特征表示;再将所述第一特征表示和所述第二特征表示输入至RPN网络,通过所述RPN网络得到所述源域图像的第一候选区域和所述目标域图像的第二候选区域。

[0106] 上述方案中,可选地,在根据所述第一特征表示确定所述源域图像的多个第一候选区域,根据所述第二特征表示确定所述目标域图像的多个第二候选区域之后,所述方法还包括:

[0107] 对所述源域图像的第一特征表示和所述源域图像的多个第一候选区域进行池化处理,以使得各个第一候选区域的特征的维度相同;

[0108] 对所述目标域图像的第二特征表示和所述目标域图像的多个第二候选区域进行池化处理,以使得各个第二候选区域的特征的维度相同;

[0109] 其中,在进行池化处理后,第一候选区域的特征的维度与第二候选区域的特征的维度相同。

[0110] 将各个第一候选区域的特征和各个第二候选区域的特征映射相同之后,可以方便后续的损失值的确定,以及对目标域图像中的目标的分类和定位。

[0111] 步骤102、分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理,得到第一局部区域和第二局部区域。

[0112] 本实施例中,所述聚类处理是根据候选区域的中心的位置进行聚类。

[0113] 在一些可选实现方式中,所述分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理,得到第一局部区域和第二局部区域,包括:

[0114] 对所述多个第一候选区域进行聚类处理,得到K个第一聚类中心;其中,所述K为正整数;

[0115] 根据各个第一聚类中心,确定各个第一局部区域;

- [0116] 对所述多个第二候选区域进行聚类处理,得到K个第二聚类中心;
- [0117] 根据各个第二聚类中心,确定各个第二局部区域。
- [0118] 具体地,通过K-均值聚类算法对第一候选区域图像进行分割,得到K个第一聚类中心;通过K-均值聚类算法对第二候选区域图像进行分割,得到K个第二聚类中心。
- [0119] 这里,K可以是人为设置的,实际操作中,K的最大值可以由图像中的目标的最大数目确定,比如,图像中最多有8个目标,则k可以是2,4,8。
- [0120] 在一些可选实现方式中,所述根据各个第一聚类中心,确定各个第一局部区域,包括:
- [0121] 以每个第一聚类中心为中心,设定范围内的区域构成一个第一局部区域。
- [0122] 同理,在一些可选实现方式中,所述根据各个第二聚类中心,确定各个第二局部区域,包括:
- [0123] 以每个第二聚类中心为中心,设定范围内的区域构成一个第二局部区域。
- [0124] 其中,K个第一聚类中心可以确定K个第一局部区域,K个第二聚类中心可以确定K个第二局部区域。
- [0125] 这里,所述设定范围可根据实际情况进行设定或调整,但构成第一局部区域时设定范围的大小和都成第二局部区域时设定范围的大小相等。
- [0126] 步骤103、根据所述第一局部区域的特征和所述第二局部区域的特征确定损失值。
- [0127] 在一些可选实施方式中,所述根据所述第一局部区域的特征和所述第二局部区域的特征确定损失值,包括:
- [0128] 根据各个第一局部区域的特征和各个第二局部区域的特征确定损失值。
- [0129] 在一些可选实施方式中,根据各个第一局部区域的特征和各个第二局部区域的特征确定损失值,包括:
- [0130] 根据各个第一局部区域的特征,重建各个第一局部区域的图像;
- [0131] 根据各个第二局部区域的特征,重建各个第二局部区域的图像;其中,所述第一局部区域与所述第二局部区域的区域大小相等;
- [0132] 从源域图像中抠取各个第一局部区域位置处的图像,得到真实的各个第一局部区域的图像;
- [0133] 从目标域图像中抠取各个第二局部区域位置处图像,得到真实的各个第二局部区域的图像;
- [0134] 对重建的各个第一局部区域的图像、真实的各个第一局部区域的图像、重建的各个第二局部区域的图像以及真实的各个第二局部区域的图像进行真伪判别;
- [0135] 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失以及真伪判别的损失确定所述损失值。
- [0136] 其中,对重建的各个第一局部区域的图像、真实的各个第一局部区域的图像、重建的各个第二局部区域的图像以及真实的各个第二局部区域的图像进行真伪判别,可以分为域内的真伪判别和域间的真伪判别。其中,域内的真伪判别包括,源域的判别器对重建的各个第一局部区域的图像及其对应的真实的第一局部区域的图像进行真伪判别,以及目标域的判别器对重建的各个第二局部区域的图像及其对应的真实的第二局部区域的图像进行真伪判别。域间的真伪判别包括,将重建的各个第一局部区域的图像作为各个真实的第二

局部区域的图像输入至目标域的判别器,从而使得目标域的判别器对重建的各个第一局部区域的图像与对应的重建的第二局部区域的图像进行真伪判别,以及将重建的各个第二局部区域的图像作为各个真实的第一局部区域的图像输入至源域的判别器,从而使得源域的判别器对重建的各个第二局部区域的图像与对应的重建的第一局部区域的图像进行真伪判别。

[0137] 上述方案中,可选地,所述方法还包括:

[0138] 为各个第二局部区域赋予权重;其中,一个第二局部区域的权重表征该第二局部区域包含所述目标域图像中的目标的概率。

[0139] 这里,可以采用权重分配网络来赋予权重,权重分配网络可以采用源域图像及其标注结果进行训练,训练完成后,权重分配网络可以为每个第二局部区域分配权重。

[0140] 具体地,根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失以及真伪判别的损失确定所述损失值,包括:

[0141] 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失、真伪判别的损失以及各个第二局部区域的权重,确定所述损失值。

[0142] 具体的,在为各个第二局部区域赋予权重之后,在确定损失值时,重建各个第二局部区域的图像的损失要乘以对应的第二局部区域的权重,目标域的判别器进行真伪判别时的损失也要乘以相关的第二局部区域的权重。

[0143] 上述步骤103在采用对抗网络实现时,相当于对所述第一局部区域和所述第二局部区域进行对齐处理。

[0144] 在一些实现方式中,步骤103可以利用生成器(Generator)、判别器(Discriminator)和权重分配器(Weighting Estimator)来实现。

[0145] 例如,各个第一局部区域的特征(也就是源域图像的第一特征表示中的部分)作为第一生成器的输入,通过所述第一生成器得到重建的各个第一局部区域的图像;所各个第二局部区域的特征(也就是目标域图像的第二特征表示中的部分)作为第二生成器的输入,通过所述第二生成器得到重建的各个第二局部区域的图像。

[0146] 例如,将重建的各个第一局部区域的图像以及真实的各个第一局部区域的图像作为第一判别器的输入,通过所述第一判别器来判断每个第一局部区域的图像所属区域类别,并为不同区域类别的第一局部区域的图像标注相应的标签;所述标签包括表征经从源域图像中抠图得到的真实图像和表征经重建得到的重建图像。可选地,将从源域图像中经抠图得到的第一局部区域的图像标注为source real,为这类第一局部区域图像贴上标签1;将源域的经重建得到的第一局部区域的图像标注为source fake,为这类第一局部区域的图像贴上标签0。

[0147] 同理,将重建的各个第二局部区域的图像以及真实的各个第二局部区域的图像作为第二判别器的输入,通过所述第二判别器来判断每个第二局部区域的图像所属区域类别,并为不同区域类别的第二局部区域的图像标注相应的标签;所述标签包括表征经从目标域图像中抠图得到的真实图像和表征经重建得到的重建图像。可选地,将从目标域图像中经抠图得到的第二局部区域的图像标注为target real,为这类第二局部区域图像贴上标签1;将目标域的经重建得到的第二局部区域的图像标注为target fake,为这类第二局部区域的图像贴上标签0。

[0148] 步骤104、基于确定的损失值,调整所述神经网络的网络参数。

[0149] 在一些可选实施方式中,根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失、真伪判别的损失以及各个第二局部区域的权重,确定所述损失值,基于确定的损失值,调整所述神经网络的网络参数。

[0150] 需要说明的是,本申请实施例对确定各个损失的具体实现不作限定。

[0151] 上述方案中,可选地,所述方法还包括:

[0152] 利用所述神经网络对源域图像进行处理,得到处理结果;

[0153] 根据得到的处理结果和源域图像的标注结果,调整所述神经网络的网络参数。

[0154] 这里,所述处理结果包括目标分类和/或目标位置。

[0155] 这里,可根据得到的处理结果和源域图像的标注结果记作检测处理过程的损失。

[0156] 具体地,根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失、真伪判别的损失以及各个第二局部区域的权重、检测处理过程的损失来确定所述损失值,基于确定的损失值,调整所述神经网络的网络参数。

[0157] 需要说明的是,刚开始训练时,第一局部区域中不一定包括目标,通过本申请所述的神经网络训练方法,得到的神经网络的泛化能力更强。比如,源域中不包括雾状城市景观,但是,通过本申请神经网络的训练方法得到的神经网络可以识别雾状城市景观中的物体。

[0158] 本申请所述技术方案可用于各种物体检测任务中,对检测场景不做限定,比如检测场景包括环境感知场景、辅助驾驶场景、跟踪场景等。

[0159] 本申请实施例提出的神经网络的训练方法,首先确定不同域的候选区域,然后通过聚类的方式,得到局部区域,再利用对抗网络对不同域的局部区域的图像进行对齐操作;如此,通过聚类的方式来只关注物体所在的区域,将背景等噪音筛选出去,这样的处理方式减小了对齐的难度,通过不同域的局部区域的图像的对齐处理,可以使神经网络更好地适应目标域的图像,在处理新环境下的数据时,检测效果更好,同时该方法不需要标注新环境的数据。另外,该方法有很好的通用性,可以用到一系列的基于区域(region)的任务中,比如实例分割。

[0160] 在一种应用场景下,本申请实施例提供的神经网络的训练方法可以采用如图2所示的框架。该框架由两个关键组成部分组成:(1)一个区域挖掘组件,包括候选区域网络(RPN,Region Proposal Network)以及聚类,用于解决“在何处查找”的问题,通过对物体特征进行分组来选择局部区域;(2)一个区域级对齐组件,用于解决“如何对齐”的问题,该组件学习如何通过对抗性学习对齐不同域的局部区域的图像,将两个域的局部区域的图像进行对齐,以得到域不变的特征。特别地,对于该组件,分别使用两个生成器 G_s (源域的生成器)和 G_t (目标域的生成器)以重建第一局部区域的图像和第二局部区域的图像,然后引入一组判别器来减小真实的局部区域的图像、重建的局部区域的图像、真实的第二局部区域的图像和重建的第二局部区域的图像之间的差别。

[0161] 在RPN之后,我们得到了很多候选区域,形式为 $\{c_x, c_y, w, h\}$,其中 c_x 和 c_y 是中心坐标, w 是宽度, h 是高度。将K均值聚类方法应用到中心坐标系中,可以得到K均值聚类,其均值可作为聚类区域的聚类中心。在给定了每个局部区域的大小之后,确定了聚类中心(来自K-

平均值),则将自动得到局部区域。图3示出了一种聚类方案和加权估计的工作流程示意图。具体地,以K=4为例,图3(a)→3(b)表示聚类操作的过程,图3(b)→3(c)表示加权估计的过程。具体地,图3中(a)→(b)显示了K=4的区域聚类示例,图3(a)中的浅灰色矩形表示候选区域,图3(b)中的深灰色正方形表示聚类确定的局部区域。

[0162] 图2中的ROI表示对主干网提取的源域图像的特征以及RPN根据源域图像的特征确定的候选区域进行池化处理,从而将各个候选区域的特征映射为维度相同之后再再进行目标的分类和定位(即FC),在神经网络的训练过程中还可以根据聚类后形成的聚类中心对池化处理后的特征进行重分配,以便于后续的对抗网络中的生成器和判别器对局部区域的图像进行处理,也就是由生成器(G_s, G_t)重建局部区域的图像,判别器(D_s, D_t)判别输入的图像(真实的局部区域的图像和重建的局部区域的图像)的真伪。 D_w 附近的全连接层为权重分配网络,为各个第二局部区域分配权重,表示各个第二局部区域中包含目标的概率。

[0163] 表示 $\widehat{\Theta}$ 重图2所示的框架中,对抗网络的损失 \mathcal{L}_{adv} 为

$$[0164] \quad \mathcal{L}_{adv}(\mathbb{F}, G, D) = \mathcal{L}_{adv, D_s} + \mathcal{L}_{adv, D_t} + \mathcal{L}_{adv, G_s} + \mathcal{L}_{adv, G_t} + \mathcal{L}_{adv, \mathbb{F}} \quad (1)$$

[0165] 其中,公式(1)的等式右边的每一项都遵循一个标准的对抗公式:

$$[0166] \quad \mathcal{L}_{adv}(\widehat{\Theta}, \mathcal{P}; G, D) = \mathbb{E}[\log D(\mathcal{P})] + \mathbb{E}[\log(1 - D(G(\widehat{\Theta})))] \quad (2)$$

[0167] 其中, \mathcal{P} 表示表示根据聚类中心 Ψ 得到的真实图像区域, $\widehat{\Theta}$ 表示根据聚类中心对池化后的特征进行重分配后的特征。公式(1)的等式右边的第一项为源域的生成器重建第一局部区域的图像的损失,公式(1)的等式右边的第二项为目标域的生成器重建第二局部区域的图像的损失,公式(1)的等式右边的第三项为源域的判别器判别真假的损失,公式(1)的等式右边的第四项为目标域的判别器判别真假的损失,公式(1)的等式右边的第五项为跨域的对抗损失,包括了将重建的各个第一局部区域的图像作为各个真实的第二局部区域的图像输入至目标域的判别器,从而使得目标域的判别器对重建的各个第一局部区域的图像与对应的重建的第二局部区域的图像进行真伪判别的损失,以及将重建的各个第二局部区域的图像作为各个真实的第一局部区域的图像输入至源域的判别器,从而使得源域的判别器对重建的各个第二局部区域的图像与对应的重建的第一局部区域的图像进行真伪判别的损失。

[0168] 图2中的

[0169] 由于目标域上没有地面真值边界框,从目标图像上的RPN提取的候选区域往往无法覆盖感兴趣的物体,为此,我们可以利用源域中的地面真值边界框来指导目标域中的焦点。

[0170] 基于这个原因,引入权重分配网络 D_w 来根据目标域与源域的匹配程度来衡量第二局部区域。分配权重的权重分配网络的损失函数如公式(3)所示:

$$[0171] \quad \mathcal{L}_w(\widehat{\Theta}_s, \widehat{\Theta}_t; D_w) = \mathbb{E}[\log(D_w(\widehat{\Theta}_s))] + \mathbb{E}[\log(1 - D_w(\widehat{\Theta}_t))] \quad (3)$$

[0172] 其中, $\widehat{\Theta}_s$ 和 $\widehat{\Theta}_t$ 分别表示特征重分配后源域和目标域中的聚类区域(局部区域)的特征表示图3(c)中正方形框中的四个数字表示目标域的4个局部区域的权重。可以看出分

数越高表明目标区域更可能包含感兴趣的物体,并且更类似于源域的分布。由于 \mathcal{W}_t 为第二局部区域的指定权重,因此它仅适用于涉及目标域的参数:

$$\begin{aligned} \mathcal{W}_t \cdot \mathcal{L}_{adv}(\mathbb{F}, \mathbb{G}, \mathbb{D}) = & \mathcal{L}_{adv, D_s} + \mathcal{W}_t \mathcal{L}_{adv, D_t} + \mathcal{L}_{adv, G_s} + \mathcal{W}_t \mathcal{L}_{adv, G_t} \\ [0173] \quad & + \mathcal{L}_{adv, \mathbb{F}}^s + \mathcal{W}_t \mathcal{L}_{adv, \mathbb{F}}^t \end{aligned} \quad (4)$$

[0174] 神经网络训练过程中的总的损失函数如公式(5)所示:

$$[0175] \quad \min_{\mathbb{F}, \mathbb{G}, D_w} \max_D \mathcal{L}_{dec}(\mathbb{F}) + \mathcal{W}_t \cdot \mathcal{L}_{adv}(\mathbb{F}, \mathbb{G}, \mathbb{D}) + \mathcal{L}_w(D_w) \quad (5)$$

[0176] 其中 \mathcal{L}_{dec} 为检测任务的损失,即 $\mathcal{L}_{dec} = \mathcal{L}_{cls} + \mathcal{L}_{loc}$ 。 \mathcal{L}_{cls} 是分类损失, \mathcal{L}_{loc} 表示定位损失,也就是图2中的FC后的分类结果和定位结果产生的损失。 $\mathcal{L}_w(D_w)$ 为公式(3)的权重分配网络的损失。

[0177] 该方案的目标是缩小源域和目标域分布之间的差距,同时保持神经网络的检测性能。因此,总损失是两部分的结合,即检测损失和对抗损失。注意,这种对抗性损失是为了执行跨域功能。总损失公式为 $L_{\mathbb{F}} = \mathcal{L}_{dec, \mathbb{F}}^s + \lambda(\mathcal{L}_{adv, \mathbb{F}}^s + \mathcal{W}_t \mathcal{L}_{adv, \mathbb{F}}^t)$,其中 λ 表示影响因子。

[0178] 在反向传播的过程中,根据损失函数的值不断调整主干网、RPN以及对抗网络的参数,使得网络最优。

[0179] 在神经网络训练完成后,应用到具体的场景中时,可以只包括图2中的主干网、RPN、ROI以及FC部分。

[0180] 对应地,本申请实施例提供了一种目标检测方法,所述方法包括:

[0181] 获取目标域图像;

[0182] 利用目标检测网络对所述目标域图像进行处理,得到所述目标域图像中各个目标的位置和/或分类;

[0183] 其中,所述目标检测网络采用上文所述的神经网络训练方法训练。

[0184] 对应上述神经网络的训练方法,本申请实施例提供了一种神经网络的训练装置,如图4所示,所述装置包括:

[0185] 确定模块10,用于确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域;

[0186] 聚类模块20,用于分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理,得到第一局部区域和第二局部区域;

[0187] 对齐模块30,用于根据所述第一局部区域的特征和所述第二局部区域的特征确定损失值;

[0188] 训练模块40,用于基于确定的损失值,调整所述神经网络的网络参数。

[0189] 作为一种实施方式,可选地,所述训练模块40,还用于:

[0190] 利用所述神经网络对源域图像进行处理,得到处理结果;

[0191] 根据得到的处理结果和源域图像的标注结果,调整所述神经网络的网络参数。

[0192] 作为一种实施方式,可选地,所述确定模块10,还用于:

[0193] 基于源域图像得到关于所述源域图像的第一特征表示,根据所述第一特征表示确

定所述源域图像的多个第一候选区域；

[0194] 基于目标域图像得到关于所述目标域图像的第二特征表示,根据所述第二特征表示确定所述目标域图像的多个第二候选区域。

[0195] 上述方案中,可选地,所述装置还包括:

[0196] 池化模块50,用于:

[0197] 在根据所述第一特征表示确定所述源域图像的多个第一候选区域,根据所述第二特征表示确定所述目标域图像的多个第二候选区域之后,

[0198] 对所述源域图像的第一特征表示和所述源域图像的多个第一候选区域进行池化处理,以使得各个第一候选区域的特征的维度相同;

[0199] 对所述目标域图像的第二特征表示和所述目标域图像的多个第二候选区域进行池化处理,以使得各个第二候选区域的特征的维度相同;

[0200] 其中,在进行池化处理后,第一候选区域的特征的维度与第二候选区域的特征的维度相同。

[0201] 作为一种实施方式,可选地,所述聚类模块20,还用于:

[0202] 对所述多个第一候选区域进行聚类处理,得到K个第一聚类中心;其中,所述K为正整数;

[0203] 根据各个第一聚类中心,确定各个第一局部区域;

[0204] 对所述多个第二候选区域进行聚类处理,得到K个第二聚类中心;

[0205] 根据各个第二聚类中心,确定各个第二局部区域;

[0206] 所述对齐模块30,还用于:

[0207] 根据各个第一局部区域的特征和各个第二局部区域的特征确定损失值。

[0208] 上述方案中,可选地,所述对齐模块,还用于:

[0209] 根据各个第一局部区域的特征,重建各个第一局部区域的图像;

[0210] 根据各个第二局部区域的特征,重建各个第二局部区域的图像;其中,所述第一局部区域与所述第二局部区域的区域大小相等;

[0211] 从源域图像中抠取各个第一局部区域位置处的图像,得到真实的各个第一局部区域的图像;

[0212] 从目标域图像中抠取各个第二局部区域位置处图像,得到真实的各个第二局部区域的图像;

[0213] 对重建的各个第一局部区域的图像、真实的各个第一局部区域的图像、重建的各个第二局部区域的图像以及真实的各个第二局部区域的图像进行真伪判别;

[0214] 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失以及真伪判别的损失确定所述损失值。

[0215] 作为一种实施方式,可选地,本申请实施例提供的神经网络的训练装置还包括权重分配模块(图4中未示出),用于

[0216] 为各个第二局部区域赋予权重;其中,一个第二局部区域的权重表征该第二局部区域包含所述目标域图像中的目标的概率;

[0217] 所述对齐模块30,还用于:根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失、真伪判别的损失以及各个第二局部区域的权重,确定所述损失

值。

[0218] 本领域技术人员应当理解,图4中所示的神经网络的训练装置中的各处理模块的实现功能可参照前述神经网络的训练方法的相关描述而理解。本领域技术人员应当理解,图4所示的神经网络的训练装置中各处理单元的功能可通过运行于处理器上的程序而实现,也可通过具体的逻辑电路而实现。

[0219] 实际应用中,上述确定模块10、聚类模块20、对齐模块30、训练模块40和池化模块50的具体结构均可对应于处理器。所述处理器具体的结构可以为中央处理器(CPU, Central Processing Unit)、微处理器(MCU, Micro Controller Unit)、数字信号处理器(DSP, Digital Signal Processing)或可编程逻辑器件(PLC, Programmable Logic Controller)等具有处理功能的电子元器件或电子元器件的集合。其中,所述处理器包括可执行代码,所述可执行代码存储在存储介质中,所述处理器可以通过总线等通信接口与所述存储介质中相连,在执行具体的各单元的对应功能时,从所述存储介质中读取并运行所述可执行代码。所述存储介质用于存储所述可执行代码的部分优选为非瞬间存储介质。

[0220] 本申请实施例提供的神经网络的训练装置,能促进神经网络的泛化能力,得到更好的物体检测效果。

[0221] 本申请实施例还记载了一种神经网络的训练装置,所述装置包括:存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现前述任意一个技术方案提供的神经网络的训练方法。

[0222] 作为一种实施方式,所述处理器执行所述程序时实现:

[0223] 确定源域图像的多个第一候选区域和目标域图像的多个第二候选区域;

[0224] 分别对所述多个第一候选区域和所述多个第二候选区域进行聚类处理,得到第一局部区域和第二局部区域;

[0225] 根据所述第一局部区域的特征和所述第二局部区域的特征确定损失值;

[0226] 基于确定的损失值,调整所述神经网络的网络参数。

[0227] 作为一种实施方式,所述处理器执行所述程序时实现:利用所述神经网络对源域图像进行处理,得到处理结果;

[0228] 根据得到的处理结果和源域图像的标注结果,调整所述神经网络的网络参数。

[0229] 作为一种实施方式,所述处理器执行所述程序时实现:基于源域图像得到关于所述源域图像的第一特征表示,根据所述第一特征表示确定所述源域图像的多个第一候选区域;

[0230] 基于目标域图像得到关于所述目标域图像的第二特征表示,根据所述第二特征表示确定所述目标域图像的多个第二候选区域。

[0231] 作为一种实施方式,所述处理器执行所述程序时实现:在根据所述第一特征表示确定所述源域图像的多个第一候选区域,根据所述第二特征表示确定所述目标域图像的多个第二候选区域之后,

[0232] 对所述源域图像的第一特征表示和所述源域图像的多个第一候选区域进行池化处理,以使得各个第一候选区域的特征的维度相同;

[0233] 对所述目标域图像的第二特征表示和所述目标域图像的多个第二候选区域进行池化处理,以使得各个第二候选区域的特征的维度相同;

[0234] 其中,在进行池化处理后,第一候选区域的特征的维度与第二候选区域的特征的维度相同。

[0235] 作为一种实施方式,所述处理器执行所述程序时实现:对所述多个第一候选区域进行聚类处理,得到K个第一聚类中心;其中,所述K为正整数;

[0236] 根据各个第一聚类中心,确定各个第一局部区域;

[0237] 对所述多个第二候选区域进行聚类处理,得到K个第二聚类中心;

[0238] 根据各个第二聚类中心,确定各个第二局部区域。

[0239] 作为一种实施方式,所述处理器执行所述程序时实现:根据各个第一局部区域的特征,重建各个第一局部区域的图像;

[0240] 根据各个第二局部区域的特征,重建各个第二局部区域的图像;其中,所述第一局部区域与所述第二局部区域的区域大小相等;

[0241] 从源域图像中抠取各个第一局部区域位置处的图像,得到真实的各个第一局部区域的图像;

[0242] 从目标域图像中抠取各个第二局部区域位置处图像,得到真实的各个第二局部区域的图像;

[0243] 对重建的各个第一局部区域的图像、真实的各个第一局部区域的图像、重建的各个第二局部区域的图像以及真实的各个第二局部区域的图像进行真伪判别;

[0244] 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失以及真伪判别的损失确定所述损失值

[0245] 作为一种实施方式,所述处理器执行所述程序时实现:为各个第二局部区域赋予权重;其中,一个第二局部区域的权重表征该第二局部区域包含所述目标域图像中的目标的概率;

[0246] 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失以及真伪判别的损失确定所述损失值,包括:

[0247] 根据重建各个第一局部区域的图像的损失、重建各个第二局部区域的图像的损失、真伪判别的损失以及各个第二局部区域的权重,确定所述损失值。

[0248] 本申请实施例提供的神经网络的训练装置,能通过局部区域的对齐,促进物体检测算法的泛化能力,得到更好的物体检测效果。

[0249] 本申请实施例提供了一种目标检测装置,所述装置包括:

[0250] 获取模块,用于获取目标域图像;

[0251] 检测模块,用于利用目标检测网络对所述目标域图像进行处理,得到所述目标域图像中各个目标的位置和/或分类;

[0252] 其中,所述目标检测网络采用上文所述的神经网络的训练方法训练。

[0253] 实际应用中,上述获取模块和检测模块的具体结构均可对应于处理器。所述处理器具体的结构可以为CPU、MCU、DSP或PLC等具有处理功能的电子元器件或电子元器件的集合。其中,所述处理器包括可执行代码,所述可执行代码存储在存储介质中,所述处理器可以通过总线等通信接口与所述存储介质中相连,在执行具体的各单元的对应功能时,从所述存储介质中读取并运行所述可执行代码。所述存储介质用于存储所述可执行代码的部分优选为非瞬间存储介质。

[0254] 本申请实施例提供的目标检测装置,物体检测效果更强,适应领域更广泛,泛化能力强。

[0255] 本申请实施例还记载了一种目标检测装置,所述装置包括:存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现前述任意一个技术方案提供的目标检测方法。

[0256] 作为一种实施方式,所述处理器执行所述程序时实现:

[0257] 获取目标域图像;

[0258] 利用目标检测网络对所述目标域图像进行处理,得到所述目标域图像中各个目标的位置和/或分类;

[0259] 其中,所述目标检测网络采用上文所述的神经网络的训练方法训练。

[0260] 本申请实施例还记载了一种计算机存储介质,所述计算机存储介质中存储有计算机可执行指令,所述计算机可执行指令用于执行前述各个实施例所述的神经网络的训练方法。也就是说,所述计算机可执行指令被处理器执行之后,能够实现前述任意一个技术方案提供的神经网络的训练方法。

[0261] 本申请实施例还记载了一种计算机存储介质,所述计算机存储介质中存储有计算机可执行指令,所述计算机可执行指令用于执行前述各个实施例所述的目标检测方法。也就是说,所述计算机可执行指令被处理器执行之后,能够实现前述任意一个技术方案提供的目标检测方法。

[0262] 本领域技术人员应当理解,本实施例的计算机存储介质中各程序的功能,可参照前述各实施例所述的神经网络的训练方法的相关描述而理解。

[0263] 在本申请所提供的几个实施例中,应该理解到,所揭露的设备和方法,可以通过其它的方式实现。以上所描述的设备实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,如:多个单元或组件可以结合,或可以集成到另一个系统,或一些特征可以忽略,或不执行。另外,所显示或讨论的各组成部分相互之间的耦合、或直接耦合、或通信连接可以是通过一些接口,设备或单元的间接耦合或通信连接,可以是电性的、机械的或其它形式的。

[0264] 上述作为分离部件说明的单元可以是、或也可以不是物理上分开的,作为单元显示的部件可以是、或也可以不是物理单元;既可以位于一个地方,也可以分布到多个网络单元上;可以根据实际的需要选择其中的部分或全部单元来实现本实施例方案的目的。

[0265] 另外,在本申请各实施例中的各功能单元可以全部集成在一个处理单元中,也可以是各单元分别单独作为一个单元,也可以两个或两个以上单元集成在一个单元中;上述集成的单元既可以采用硬件的形式实现,也可以采用硬件加软件功能单元的形式实现。

[0266] 本领域普通技术人员可以理解:实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成,前述的程序可以存储于计算机可读取存储介质中,该程序在执行时,执行包括上述方法实施例的步骤;而前述的存储介质包括:移动存储设备、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0267] 或者,本申请上述集成的单元如果以软件功能模块的形式实现并作为独立的产品销售或使用,也可以存储在一个计算机可读取存储介质中。基于这样的理解,本申请实施

例的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机、服务器、或者网络设备等)执行本申请各个实施例所述方法的全部或部分。而前述的存储介质包括:移动存储设备、ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0268] 以上所述,仅为本申请的具体实施方式,但本申请的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本申请揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本申请的保护范围之内。因此,本申请的保护范围应以所述权利要求的保护范围为准。

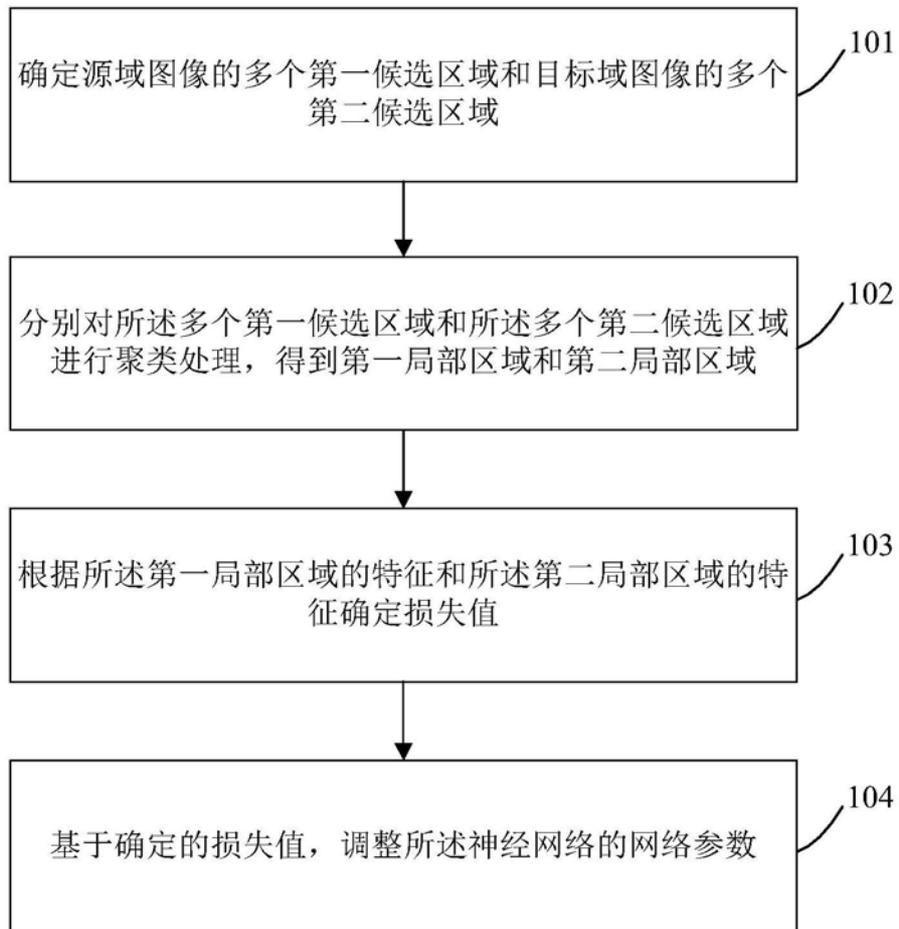


图1

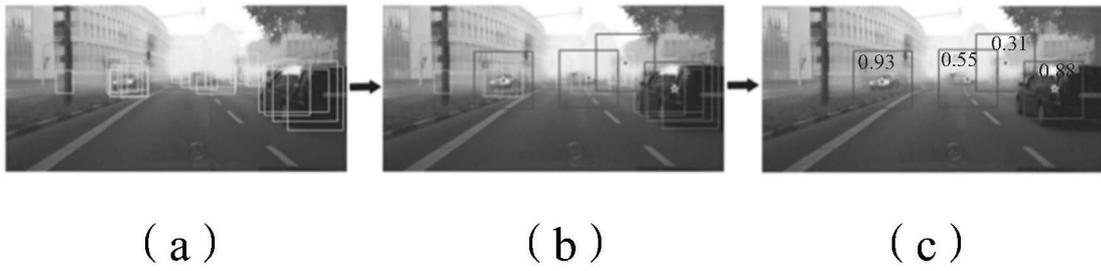


图3

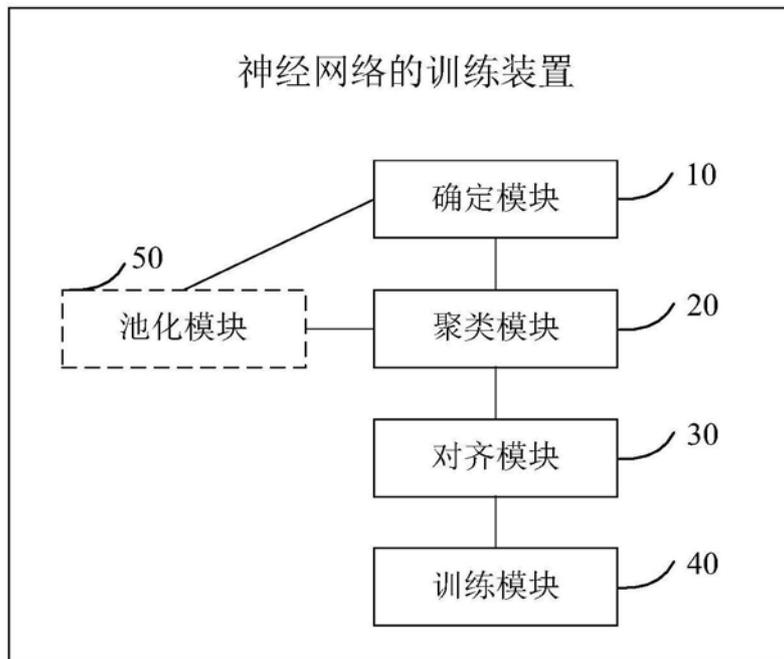


图4