



(12) 发明专利

(10) 授权公告号 CN 110084309 B

(45) 授权公告日 2022.06.21

(21) 申请号 201910364870.X

G06V 10/774 (2022.01)

(22) 申请日 2019.04.30

G06N 3/04 (2006.01)

(65) 同一申请的已公布的文献号  
申请公布号 CN 110084309 A

G06T 5/50 (2006.01)

G06T 11/00 (2006.01)

(43) 申请公布日 2019.08.02

(56) 对比文件

CN 109635882 A, 2019.04.16

(73) 专利权人 北京市商汤科技开发有限公司  
地址 100080 北京市海淀区北四环西路58号11层1101-1117室

CN 109598768 A, 2019.04.09

US 9262843 B1, 2016.02.16

Jiaqi Wang et al..CARAFE: Content-Aware ReAssembly of FEatures.《arXiv》.2019,

(72) 发明人 王佳琦 陈恺 徐瑞 刘子纬  
吕健勤 林达华

毕鹏程 等.轻量化卷积神经网络技术研究.《计算机工程与应用》.2019,

(74) 专利代理机构 北京博思佳知识产权代理有限公司 11415

审查员 傅志良

专利代理师 林祥

(51) Int. Cl.

G06K 9/62 (2022.01)

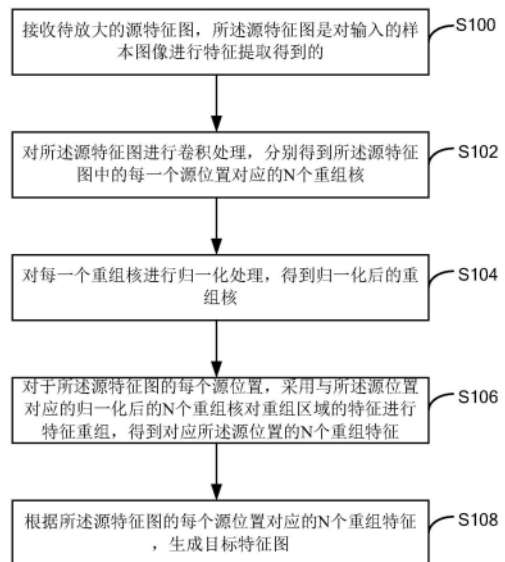
权利要求书5页 说明书17页 附图10页

(54) 发明名称

特征图放大方法、装置和设备及计算机可读存储介质

(57) 摘要

本公开实施例提供一种特征图放大方法、装置和设备及计算机可读存储介质,其中,所述方法包括:接收待放大的源特征图,所述源特征图是对输入的样本图像进行特征提取得到的;对源特征图进行卷积处理分别得到源特征图中的每一个源位置对应的N个重组核,N是源特征图的放大倍数的平方;对每一个重组核进行归一化处理;对于源特征图的每个源位置,采用与源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组,得到源位置的N个重组特征;根据源特征图的每个源位置对应的N个重组特征,生成目标特征图。本公开提升了特征图放大的性能。



1. 一种特征图放大方法,其特征在于,所述方法包括:

接收待放大的源特征图,所述源特征图是对输入的样本图像进行特征提取得到的;

对所述源特征图进行卷积处理,分别得到所述源特征图中的每一个源位置对应的N个重组核,所述N是源特征图的放大倍数的平方;所述N个重组核是通过所述源位置的内容进行卷积处理得到,所述源位置是源特征图上的一个像素;

对每一个重组核进行归一化处理,得到归一化后的重组核;

对于所述源特征图的每个源位置,采用与所述源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组,得到对应所述源位置的N个重组特征;所述重组区域根据所述源位置确定;所述重组区域是以所述源位置为中心的一个区域;

根据所述源特征图的每个源位置对应的N个重组特征,生成目标特征图。

2. 根据权利要求1所述的方法,其特征在于,

所述对所述源特征图进行卷积处理,分别得到所述源特征图中的每一个源位置对应的N个重组核,包括:

以 $C_{up}$ 个过滤器对所述源特征图进行卷积处理,得到 $C_{up}$ 通道的重组核图;所述 $C_{up}$ 是根据所述放大倍数和重组核的尺寸确定;

将所述重组核图中,对应于所述源特征图同一源位置的 $C_{up}$ 通道的卷积输出特征,均分成N份,每一份作为一个所述重组核,一个所述重组核包括 $C_{up}/N$ 通道的卷积输出特征。

3. 根据权利要求2所述的方法,其特征在于,所述采用与所述源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组之前,所述方法还包括:

对于与所述源位置对应的每一个重组核,将所述重组核中包括的 $C_{up}/N$ 通道的卷积输出特征平铺,得到二维的所述重组核;

获取以所述源位置为中心的所述重组区域,所述重组区域的尺寸与所述重组核相同。

4. 根据权利要求2所述的方法,其特征在于,所述对每一个重组核进行归一化处理,得到归一化后的重组核,包括:

对每一个重组核中的 $C_{up}/N$ 通道的卷积输出特征进行归一化,得到归一化后的重组核。

5. 根据权利要求1-4任一所述的方法,其特征在于,采用与一个源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组,得到对应所述源位置的N个重组特征,包括:

对于每一个所述重组核,将所述重组核分别与源特征图的各个通道的所述重组区域进行特征重组,得到与所述重组核对应的所述各个通道的通道特征;

将所述重组核对应的各个通道的所述通道特征组合,得到与所述重组核对应的重组特征。

6. 根据权利要求5所述的方法,其特征在于,所述将所述重组核分别与源特征图的各个通道的所述重组区域进行特征重组,得到与所述重组核对应的所述各个通道的通道特征,包括:

对于其中一个通道,将所述重组核与所述通道的重组区域中的对应位置的特征进行加权求和,所述加权求和的结果作为所述通道对应的通道特征。

7. 根据权利要求1所述的方法,其特征在于,所述根据所述源特征图的每个源位置对应的N个重组特征,生成目标特征图,包括:

根据所述源位置和放大倍数,得到目标特征图中与所述源位置对应的目标位置;

将所述N个重组特征,填充入目标特征图中的所述目标位置。

8. 根据权利要求1所述的方法,其特征在于,所述源特征图是待训练的神经网络对输入的样本图像进行特征提取得到的;所述对源特征图进行卷积处理,具体包括:通过所述待训练的神经网络中的内容编码卷积层对所述源特征图进行卷积处理;

所述方法还包括:

在生成目标特征图之后,所述待训练的神经网络根据所述目标特征图,得到所述样本图像的预测结果;

基于所述样本图像的标注结果和预测结果之间的差别,调整所述待训练的神经网络中的网络参数,所述网络参数包括所述内容编码卷积层的参数。

9. 根据权利要求8所述的方法,其特征在于,所述通过所述待训练的神经网络中的内容编码卷积层对所述源特征图进行卷积处理之前,所述方法还包括:

通过所述待训练的神经网络中的通道压缩卷积层对所述源特征图进行通道压缩;

通过所述内容编码卷积层对所述源特征图进行卷积处理,具体包括:通过所述内容编码卷积层对通道压缩后的源特征图进行卷积处理;

基于所述样本图像的标注结果和预测结果之间的差别,调整所述待训练的神经网络中的网络参数包括:基于所述样本图像的标注结果和预测结果之间的差别,调整所述通道压缩卷积层的参数。

10. 根据权利要求1所述的方法,其特征在于,所述源特征图是训练完成的神经网络对输入的样本图像进行特征提取得到的;所述对源特征图进行卷积处理,具体为:通过训练完成的神经网络中的内容编码卷积层对所述源特征图进行卷积处理;

所述方法还包括:在生成目标特征图之后,所述训练完成的神经网络根据所述目标特征图,得到输入图像的处理结果。

11. 根据权利要求10所述的方法,其特征在于,通过所述内容编码卷积层对所述源特征图进行卷积处理之前,所述方法还包括:

通过所述训练完成的神经网络中的通道压缩卷积层对所述源特征图进行通道压缩;

通过所述内容编码卷积层对所述源特征图进行卷积处理,具体包括:通过所述内容编码卷积层对通道压缩后的源特征图进行卷积处理。

12. 根据权利要求10或11所述的方法,其特征在于,所述输入图像中包括待检测的至少一个目标物体;所述训练完成的神经网络根据所述目标特征图,得到所述输入图像的处理结果,包括:

所述训练完成的神经网络根据所述目标特征图,输出所述输入图像中的每个目标物体的检测框、以及所述目标物体的所属类别。

13. 根据权利要求10或11所述的方法,其特征在于,所述输入图像包括残缺部分;所述训练完成的神经网络根据所述目标特征图,得到所述输入图像的处理结果,包括:

所述训练完成的神经网络根据所述目标特征图,输出所述输入图像的残缺部分的像素。

14. 根据权利要求10或11所述的方法,其特征在于,所述训练完成的神经网络根据所述目标特征图,得到所述输入图像的处理结果,包括:

所述训练完成的神经网络根据所述目标特征图,输出所述输入图像上各个像素的所属

类别。

15. 根据权利要求10或11所述的方法,其特征在于,所述训练完成的神经网络根据所述目标特征图,得到所述输入图像的处理结果,包括:

所述训练完成的神经网络根据所述目标特征图,输出所述输入图像中的每个实例、以及各个实例内的各像素的所属类别。

16. 一种特征图放大装置,其特征在于,所述装置包括:

源图接收模块,用于接收待放大的源特征图,所述源特征图是对输入的样本图像进行特征提取得到的;

核生成模块,用于对所述源特征图进行卷积处理,分别得到所述源特征图中的每一个源位置对应的N个重组核,所述N是源特征图的放大倍数的平方;所述N个重组核是通过对所源位置的内容进行卷积处理得到,所述源位置是源特征图上的一个像素;

归一处理模块,用于对每一个重组核进行归一化处理,得到归一化后的重组核;

特征重组模块,用于对于所述源特征图的每个源位置,采用与所述源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组,得到对应所述源位置的N个重组特征;所述重组区域根据所述源位置确定;所述重组区域是以所述源位置为中心的一个区域;

目标图生成模块,用于根据所述源特征图的每个源位置对应的N个重组特征,生成目标特征图。

17. 根据权利要求16所述的装置,其特征在于,

以 $C_{up}$ 个过滤器对所述源特征图进行卷积处理,得到 $C_{up}$ 通道的重组核图;所述 $C_{up}$ 是根据所述放大倍数和重组核的尺寸确定;

将所述重组核图中,对应于所述源特征图同一源位置的 $C_{up}$ 通道的卷积输出特征,均分成N份,每一份作为一个所述重组核,一个所述重组核包括 $C_{up}/N$ 通道的卷积输出特征。

18. 根据权利要求17所述的装置,其特征在于,

所述特征重组模块,还用于:在采用与所述源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组之前,对于与所述源位置对应的每一个重组核,将所述重组核中包括的 $C_{up}/N$ 通道的卷积输出特征平铺,得到二维的所述重组核;获取以所述源位置为中心的所述重组区域,所述重组区域的尺寸与所述重组核相同。

19. 根据权利要求17所述的装置,其特征在于,所述归一处理模块具体用于:

对每一个重组核中的 $C_{up}/N$ 通道的卷积输出特征进行归一化,得到归一化后的重组核。

20. 根据权利要求16至19任一所述的装置,其特征在于,

所述特征重组模块,用于在采用与一个源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组,得到对应所述源位置的N个重组特征时,包括:对于每一个所述重组核,将所述重组核分别与源特征图的各个通道的所述重组区域进行特征重组,得到与所述重组核对应的所述各个通道的通道特征;将所述重组核对应的各个通道的所述通道特征组合,得到与所述重组核对应的重组特征。

21. 根据权利要求20所述的装置,其特征在于,

所述特征重组模块,用于在将所述重组核分别与源特征图的各个通道的所述重组区域进行特征重组,得到与所述重组核对应的所述各个通道的通道特征时,包括:对于其中一个通道,将所述重组核与所述通道的重组区域中的对应位置的特征进行加权求和,所述加权

求和的结果作为所述通道对应的通道特征。

22. 根据权利要求16所述的装置,其特征在於,

所述目标图生成模块,具体用于:根据所述源位置和放大倍数,得到目标特征图中与所述源位置对应的目标位置;将所述N个重组特征,填充入目标特征图中的所述目标位置。

23. 根据权利要求16所述的装置,其特征在於,所述特征图放大装置包含于待训练的神经网络中;所述源图接收模块接收到的源特征图是由所述待训练的神经网络对输入的样本图像进行特征提取得到的;所述核生成模块用于在对所述源特征图进行卷积处理时,具体包括:通过所述待训练的神经网络中的内容编码卷积层对所述源特征图进行卷积处理;

所述装置还包括:预测处理模块和参数调整模块;

所述预测处理模块,用于在目标图生成模块生成目标特征图之后,根据所述目标特征图,得到所述样本图像的预测结果;

所述参数调整模块,用于基于所述样本图像的标注结果和预测结果之间的差别,调整所述待训练的神经网络中的网络参数,所述网络参数包括所述内容编码卷积层的参数。

24. 根据权利要求23所述的装置,其特征在於,所述装置还包括:通道压缩模块,用于通过所述待训练的神经网络中的通道压缩卷积层对所述源特征图进行通道压缩;

所述核生成模块,用于在进行卷积处理时,具体包括通过内容编码卷积层对通道压缩后的源特征图进行卷积处理;

所述参数调整模块,还用于基于所述样本图像的标注结果和预测结果之间的差别,调整所述通道压缩卷积层的参数。

25. 根据权利要求16所述的装置,其特征在於,所述特征图放大装置包含于训练完成的神经网络;所述源图接收模块接收到的源特征图是由所述训练完成的神经网络对输入的样本图像进行特征提取得到的;所述核生成模块用于在对所述源特征图进行卷积处理时,具体包括通过所述训练完成的神经网络中的内容编码卷积层对所述源特征图进行卷积处理;

所述装置还包括:结果输出模块,用于在目标图生成模块生成目标特征图之后,根据所述目标特征图,得到输入图像的处理结果。

26. 根据权利要求25所述的装置,其特征在於,所述装置还包括:

通道压缩模块,用于通过所述训练完成的神经网络中的通道压缩卷积层对所述源特征图进行通道压缩;

所述核生成模块,用于在进行卷积处理时,具体包括通过所述内容编码卷积层对通道压缩后的源特征图进行卷积处理。

27. 根据权利要求25或26所述的装置,其特征在於,所述输入图像中包括待检测的至少一个目标物体;

所述结果输出模块,具体用于:根据所述目标特征图,输出所述输入图像中的每个目标物体的检测框、以及所述目标物体的所属类别。

28. 根据权利要求25或26所述的装置,其特征在於,所述输入图像包括残缺部分;

所述结果输出模块,具体用于:根据所述目标特征图,输出所述输入图像的残缺部分的像素。

29. 根据权利要求25或26所述的装置,其特征在於,

所述结果输出模块,具体用于:根据所述目标特征图,输出所述输入图像中的每个实

例、以及各个实例内的各像素的所属类别。

30. 根据权利要求25或26所述的装置,其特征在于,

所述结果输出模块,具体用于:根据所述目标特征图,输出所述输入图像上各个像素的所属类别。

31. 一种特征图放大设备,其特征在于,所述设备包括存储器、处理器,所述存储器用于存储可在处理器上运行的计算机指令,所述处理器用于在执行所述计算机指令时实现权利要求1至15任一所述的方法。

32. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,所述程序被处理器执行时实现权利要求1至15任一所述的方法。

## 特征图放大方法、装置和设备及计算机可读存储介质

### 技术领域

[0001] 本公开涉及计算机视觉技术,具体涉及一种特征图放大方法、装置和设备及计算机可读存储介质。

### 背景技术

[0002] 计算机视觉任务中,特征图放大是最基本的操作之一。例如,在需要进行密集预测的任务中,如,物体检测、图像语义分割、图像修复等,需要将神经网络输出的低分辨率特征图放大,以此来得到高分辨率的预测结果,进行神经网络的学习和测试。又例如,在神经网络设计中,将低分辨率的特征图放大后,与高分辨率的特征图进行融合,也是一种常见的操作。

[0003] 实现特征图放大操作的方法,可以包括最近邻插值、双线性插值和反卷积等操作。其中,最近邻插值和双线性插值根据空间距离来放大特征图,放弃了利用特征图包含的丰富语义特征。反卷积尝试利用特征图的语义特征来放大特征图,但是,放大的特征图的性能有待改进。

### 发明内容

[0004] 有鉴于此,本公开至少提供一种特征图放大方法、装置和设备及计算机可读存储介质。

[0005] 第一方面,提供一种特征图放大的方法,所述方法包括:

[0006] 接收待放大的源特征图,所述源特征图是对输入的样本图像进行特征提取得到的;

[0007] 对所述源特征图进行卷积处理,分别得到所述源特征图中的每一个源位置对应的N个重组核,所述N是源特征图的放大倍数的平方;

[0008] 对每一个重组核进行归一化处理,得到归一化后的重组核;

[0009] 对于所述源特征图的每个源位置,采用与所述源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组,得到对应所述源位置的N个重组特征;所述重组区域根据所述源位置确定;

[0010] 根据所述源特征图的每个源位置对应的N个重组特征,生成目标特征图。

[0011] 结合本公开提供的任一实施方式,所述对所述源特征图进行卷积处理,分别得到所述源特征图中的每一个源位置对应的N个重组核,包括:

[0012] 以 $C_{up}$ 个过滤器对所述源特征图进行卷积处理,得到 $C_{up}$ 通道的重组核图;所述 $C_{up}$ 是根据所述放大倍数和重组核的尺寸确定;

[0013] 将所述重组核图中,对应于所述源特征图同一源位置的 $C_{up}$ 通道的卷积输出特征,均分成N份,每一份作为一个所述重组核,一个所述重组核包括 $C_{up}/N$ 通道的卷积输出特征。

[0014] 结合本公开提供的任一实施方式,所述采用与源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组之前,所述方法还包括:

[0015] 对于与所述源位置对应的每一个重组核,将所述重组核中包括的 $C_{up}/N$ 通道的卷积输出特征平铺,得到二维的所述重组核;

[0016] 获取以所述源位置为中心的所述重组区域,所述重组区域的尺寸与所述重组核相同。

[0017] 结合本公开提供的任一实施方式,所述对每一个重组核进行归一化处理,得到归一化后的重组核,包括:对每一个重组核中的 $C_{up}/N$ 通道的卷积输出特征进行归一化,得到归一化后的重组核。

[0018] 结合本公开提供的任一实施方式,采用与一个源位置对应的归一化后的 $N$ 个重组核对重组区域的特征进行特征重组,得到对应所述源位置的 $N$ 个重组特征,包括:

[0019] 对于每一个所述重组核,将所述重组核分别与源特征图的各个通道的所述重组区域进行特征重组,得到与所述重组核对应的所述各个通道的通道特征;

[0020] 将所述重组核对应的各个通道的所述通道特征组合,得到与所述重组核对应的重组特征。

[0021] 结合本公开提供的任一实施方式,所述将所述重组核分别与源特征图的各个通道的所述重组区域进行特征重组,得到与所述重组核对应的所述各个通道的通道特征,包括:

[0022] 对于其中一个通道,将所述重组核与所述通道的重组区域中的对应位置的特征进行加权求和,所述加权求和的结果作为所述通道对应的通道特征。

[0023] 结合本公开提供的任一实施方式,所述根据所述源特征图的每个源位置对应的 $N$ 个重组特征,生成目标特征图,包括:

[0024] 根据所述源位置和放大倍数,得到目标特征图中与所述源位置对应的目标位置;将所述 $N$ 个重组特征,填充入目标特征图中的所述目标位置。

[0025] 结合本公开提供的任一实施方式,所述源特征图是待训练的神经网络对输入的样本图像进行特征提取得到的;所述对源特征图进行卷积处理,具体包括:通过所述待训练的神经网络中的内容编码卷积层对所述源特征图进行卷积处理;

[0026] 所述方法还包括:

[0027] 在生成目标特征图之后,所述待训练的神经网络根据所述目标特征图,得到所述样本图像的预测结果;

[0028] 基于所述样本图像的标注结果和预测结果之间的差别,调整所述待训练的神经网络中的网络参数,所述网络参数包括所述内容编码卷积层的参数。

[0029] 结合本公开提供的任一实施方式,所述通过所述待训练的神经网络中的内容编码卷积层对所述源特征图进行卷积处理之前,所述方法还包括:

[0030] 通过所述待训练的神经网络中的通道压缩卷积层对所述源特征图进行通道压缩;

[0031] 通过所述内容编码卷积层对所述源特征图进行卷积处理,具体包括:通过所述内容编码卷积层对通道压缩后的源特征图进行卷积处理;

[0032] 基于所述样本图像的标注结果和预测结果之间的差别,调整所述待训练的神经网络中的网络参数包括:基于所述样本图像的标注结果和预测结果之间的差别,调整所述通道压缩卷积层的参数。

[0033] 结合本公开提供的任一实施方式,所述源特征图是训练完成的神经网络对输入的样本图像进行特征提取得到的;所述对源特征图进行卷积处理,具体为:通过训练完成的神



神经网络中的内容编码卷积层对所述源特征图进行卷积处理；

[0034] 所述方法还包括：在生成目标特征图之后，所述训练完成的神经网络根据所述目标特征图，得到所述输入图像的处理结果。

[0035] 结合本公开提供的任一实施方式，通过所述内容编码卷积层对所述源特征图进行卷积处理之前，所述方法还包括：

[0036] 通过所述训练完成的神经网络中的通道压缩卷积层对所述源特征图进行通道压缩；通过所述内容编码卷积层对所述源特征图进行卷积处理，具体包括：通过所述内容编码卷积层对通道压缩后的源特征图进行卷积处理。

[0037] 结合本公开提供的任一实施方式，所述输入图像中包括待检测的至少一个目标物体；所述训练完成的神经网络根据所述目标特征图，得到所述输入图像的处理结果，包括：

[0038] 所述训练完成的神经网络根据所述目标特征图，输出所述输入图像中的每个目标物体的检测框、以及所述目标物体的所属类别。

[0039] 结合本公开提供的任一实施方式，所述输入图像包括残缺部分；所述训练完成的神经网络根据所述目标特征图，得到所述输入图像的处理结果，包括：

[0040] 所述训练完成的神经网络根据所述目标特征图，输出所述输入图像的残缺部分的像素。

[0041] 结合本公开提供的任一实施方式，所述训练完成的神经网络根据所述目标特征图，得到所述输入图像的处理结果，包括：

[0042] 所述训练完成的神经网络根据所述目标特征图，输出所述输入图像上各个像素的所属类别。

[0043] 结合本公开提供的任一实施方式，所述训练完成的神经网络根据所述目标特征图，得到所述输入图像的处理结果，包括：

[0044] 所述训练完成的神经网络根据所述目标特征图，输出所述输入图像中的每个实例、以及各个实例内的各像素的所属类别。

[0045] 第二方面，提供一种特征图放大装置，所述装置包括：

[0046] 源图接收模块，用于接收待放大的源特征图，所述源特征图是对输入的样本图像进行特征提取得到的；

[0047] 核生成模块，用于对所述源特征图进行卷积处理，分别得到所述源特征图中的每一个源位置对应的N个重组核，所述N是源特征图的放大倍数的平方；

[0048] 归一处理模块，用于对每一个重组核进行归一化处理，得到归一化后的重组核；

[0049] 特征重组模块，用于对于所述源特征图的每个源位置，采用与所述源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组，得到对应所述源位置的N个重组特征；所述重组区域根据所述源位置确定；

[0050] 目标图生成模块，用于根据所述源特征图的每个源位置对应的N个重组特征，生成目标特征图。

[0051] 结合本公开提供的任一实施方式，所述核生成模块，具体用于：

[0052] 以 $C_{up}$ 个过滤器对所述源特征图进行卷积处理，得到 $C_{up}$ 通道的重组核图；所述 $C_{up}$ 是根据所述放大倍数和重组核的尺寸确定；

[0053] 将所述重组核图中，对应于所述源特征图同一源位置的 $C_{up}$ 通道的卷积输出特征，

均分成N份,每一份作为一个所述重组核,一个所述重组核包括 $C_{up}/N$ 通道的卷积输出特征。

[0054] 结合本公开提供的任一实施方式,所述特征重组模块,还用于:在采用与所述源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组之前,对于与所述源位置对应的每一个重组核,将所述重组核中包括的 $C_{up}/N$ 通道的卷积输出特征平铺,得到二维的所述重组核;获取以所述源位置为中心的所述重组区域,所述重组区域的尺寸与所述重组核相同。

[0055] 结合本公开提供的任一实施方式,所述归一化处理模块具体用于:对每一个重组核中的 $C_{up}/N$ 通道的卷积输出特征进行归一化,得到归一化后的重组核。

[0056] 结合本公开提供的任一实施方式,所述特征重组模块,用于在采用与一个源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组,得到对应所述源位置的N个重组特征时,包括:对于每一个所述重组核,将所述重组核分别与源特征图的各个通道的所述重组区域进行特征重组,得到与所述重组核对应的所述各个通道的通道特征;将所述重组核对应的各个通道的所述通道特征组合,得到与所述重组核对应的重组特征。

[0057] 结合本公开提供的任一实施方式,所述特征重组模块,用于在将所述重组核分别与源特征图的各个通道的所述重组区域进行特征重组,得到与所述重组核对应的所述各个通道的通道特征时,包括:对于其中一个通道,将所述重组核与所述通道的重组区域中的对应位置的特征进行加权求和,所述加权求和的结果作为所述通道对应的通道特征。

[0058] 结合本公开提供的任一实施方式,所述目标图生成模块,具体用于:根据所述源位置和放大倍数,得到目标特征图中与所述源位置对应的目标位置;将所述N个重组特征,填充入目标特征图中的所述目标位置。

[0059] 结合本公开提供的任一实施方式,所述特征图放大装置包含于待训练的神经网络中;所述源图接收模块接收到的源特征图是由所述待训练的神经网络对输入的样本图像进行特征提取得到的;所述核生成模块用于在对所述源特征图进行卷积处理时,具体包括:通过所述待训练的神经网络中的内容编码卷积层对所述源特征图进行卷积处理;

[0060] 所述装置还包括:预测处理模块和参数调整模块;

[0061] 所述预测处理模块,用于在目标图生成模块生成目标特征图之后,根据所述目标特征图,得到所述样本图像的预测结果;

[0062] 所述参数调整模块,用于基于所述样本图像的标注结果和预测结果之间的差别,调整所述待训练的神经网络中的网络参数,所述网络参数包括所述内容编码卷积层的参数。

[0063] 结合本公开提供的任一实施方式,所述装置还包括:通道压缩模块,用于通过所述待训练的神经网络中的通道压缩卷积层对所述源特征图进行通道压缩;

[0064] 所述核处理模块,用于在进行卷积处理时,具体包括通过内容编码卷积层对通道压缩后的源特征图进行卷积处理;

[0065] 所述参数调整模块,还用于基于所述样本图像的标注结果和预测结果之间的差别,调整所述通道压缩卷积层的参数。

[0066] 结合本公开提供的任一实施方式,所述特征图放大装置包含于训练完成的神经网络;所述源图接收模块接收到的源特征图是由所述训练完成的神经网络对输入的样本图像进行特征提取得到的;所述核生成模块用于在对所述源特征图进行卷积处理时,具体包括

通过所述训练完成的神经网络中的内容编码卷积层对所述源特征图进行卷积处理；

[0067] 所述装置还包括：结果输出模块，用于在目标图生成模块生成目标特征图之后，根据所述目标特征图，得到所述输入图像的处理结果。

[0068] 结合本公开提供的任一实施方式，所述装置还包括：

[0069] 通道压缩模块，用于通过所述训练完成的神经网络中的通道压缩卷积层对所述源特征图进行通道压缩；

[0070] 所述核处理模块，用于在进行卷积处理时，具体包括通过所述内容编码卷积层对通道压缩后的源特征图进行卷积处理。

[0071] 结合本公开提供的任一实施方式，所述输入图像中包括待检测的至少一个目标物体；所述结果输出模块，具体用于：根据所述目标特征图，输出所述输入图像中的每个目标物体的检测框、以及所述目标物体的所属类别。

[0072] 结合本公开提供的任一实施方式，所述输入图像包括残缺部分；

[0073] 所述结果输出模块，具体用于：根据所述目标特征图，输出所述输入图像的残缺部分的像素。

[0074] 结合本公开提供的任一实施方式，所述结果输出模块，具体用于：根据所述目标特征图，输出所述输入图像中的每个实例、以及各个实例内的各像素的所属类别。

[0075] 结合本公开提供的任一实施方式，所述结果输出模块，具体用于：根据所述目标特征图，输出所述输入图像上各个像素的所属类别。

[0076] 第三方面，提供一种特征图放大设备，所述设备包括存储器、处理器，所述存储器用于存储可在处理器上运行的计算机指令，所述处理器用于在执行所述计算机指令时实现本公开任一实施例的特征图放大方法。

[0077] 第四方面，提供一种计算机可读存储介质，其上存储有计算机程序，所述程序被处理器执行时实现本公开任一实施例的特征图放大方法。

[0078] 本公开一个或多个实施例的特征图放大方法、装置和设备及计算机可读存储介质，通过根据源特征图的不同位置的内容进行卷积，得到对应该位置内容的重组核，使得即使是同一个源特征图，不同的源特征图位置获得的是不同的重组核，本实施例的重组核是与源特征图中的不同位置对应，并且是感知所述位置的源特征图内容得到，因此在利用重组核对特征图进行特征重组时充分利用了特征图中的内容，从而使得利用该重组核对源特征图放大时，特征图放大的准确率更高，提升了特征图放大的性能。

## 附图说明

[0079] 为了更清楚地说明本公开一个或多个实施例或相关技术中的技术方案，下面将对实施例或相关技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本公开一个或多个实施例中记载的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

[0080] 图1为本公开至少一个实施例提供的一种特征图放大的方法；

[0081] 图2为本公开至少一个实施例提供的一种特征图放大网络；

[0082] 图3为本公开至少一个实施例提供的一种应用于待训练的神经网络的特征图放大方法；

- [0083] 图4为本公开至少一个实施例提供了一种源特征图放大原理图；
- [0084] 图5为本公开至少一个实施例提供的另一种特征图放大方法；
- [0085] 图6为本公开至少一个实施例提供的重组核图的生成原理示意图；
- [0086] 图7为本公开至少一个实施例提供了一种重组核图的变换示意图；
- [0087] 图8为本公开至少一个实施例提供了一种特征重组的流程图；
- [0088] 图9为本公开至少一个实施例提供的重组特征的生成示意图；
- [0089] 图10为本公开至少一个实施例提供了一种应用于训练完成的神经网络的特征图放大方法；
- [0090] 图11为本公开至少一个实施例提供了一种FPN的网络结构示意图；
- [0091] 图12为本公开至少一个实施例提供了一种特征图放大装置；
- [0092] 图13为本公开至少一个实施例提供的另一种特征图放大装置；
- [0093] 图14为本公开至少一个实施例提供的又一种特征图放大装置。

### 具体实施方式

[0094] 为了使本技术领域的人员更好地理解本公开一个或多个实施例中的技术方案，下面将结合本公开一个或多个实施例中的附图，对本公开一个或多个实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本公开一部分实施例，而不是全部的实施例。基于本公开一个或多个实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都应当属于本公开保护的范围。

[0095] 特征图放大操作在物体检测、图像语义分割等多种任务中应用广泛，本公开提供了一种特征图放大方法，该方法不同于传统的特征图放大的地方在于，能够根据待放大的特征图的不同位置的内容，得到与该位置内容对应的重组核。其中，如果将待放大的特征图称为源特征图，放大后的特征图称为目标特征图，那么重组核的作用是用于根据源特征图上的特征获得目标特征图上的特征。

[0096] 由于该方法是基于特征图不同位置的内容得到重组核，可以称为“基于内容感知的特征图放大方法”、或者“基于内容感知的特征重组算法” (Content-Aware ReAssembly of Features, 简称:CARAFE)。

[0097] 图1示例了本公开的一种特征图放大的方法，如图1所示，该方法包括：

[0098] 在步骤S100中，接收待放大的源特征图，所述源特征图是对输入的样本图像进行特征提取得到的。

[0099] 在步骤S102中，对所述源特征图进行卷积处理，分别得到所述源特征图中的每一个源位置对应的N个重组核。

[0100] 本步骤中，所述的一个源位置可以是源特征图上的一个像素。所述的N是源特征图的放大倍数的平方。所述的卷积处理可以通过卷积层对源特征图进行卷积，并且每一个源位置对应得到N个重组核。

[0101] 在步骤S104中，对每一个重组核进行归一化处理，得到归一化后的重组核。

[0102] 例如，可以使用softmax对重组核进行归一化。实际实施中不局限于此。

[0103] 在步骤S106中，对于所述源特征图的每个源位置，采用与所述源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组，得到对应所述源位置的N个重组特征；

所述重组区域根据所述源位置确定。

[0104] 本步骤中,每一个重组核与源特征图的重组区域进行特征重组,得到对应的一个重组特征;N个重组核共可以得到N个重组特征。所述的重组区域例如是以源位置为中心的一个区域,且该重组区域的尺寸与重组核的尺寸相同。

[0105] 在步骤S108中,根据所述源特征图的每个源位置对应的N个重组特征,生成目标特征图。

[0106] 例如,可以将每个源位置对应的N个重组特征,填入目标特征图中对应源位置的目标位置,最终得到将源特征图放大的目标特征图。

[0107] 本实施例的特征图放大方法,通过根据源特征图的不同位置的内容进行卷积,得到对应该位置内容的重组核,使得即使是同一个源特征图,不同的源特征图位置获得的是不同的重组核,这区别于传统方式中对同一源特征图的所有位置都采用同一个核的方式,本实施例的重组核是与源特征图中的不同位置对应,并且是感知所述位置的源特征图内容得到,因此,在利用重组核对特征图进行特征重组(也就是放大时)时充分利用了特征图中的内容,从而使得利用该重组核对源特征图放大时,特征图放大的准确率更高,提升了特征图放大的性能。

[0108] 图2示例了一种特征图放大网络,该特征图放大网络可以执行图1所示的特征图放大方法。如图2的示意,该特征图放大网络可以包括:核预测模块11、以及重组模块12。当待放大的源特征图13输入该特征图放大网络时,可以先通过核预测模块11根据该源特征图13的内容,获得重组核14。重组模块12再使用该重组核14,对源特征图13上的特征进行重组,得到的特征可以称为重组特征。将重组特征填充到目标特征图15对应的位置,从而完成对源特征图的放大。

[0109] 上述的特征图放大网络,要经过训练再应用到实际任务中。如下将分别描述该特征图放大网络的训练阶段和应用阶段。其中,训练时涉及到特征图放大网络中的核预测模块的参数调整,重组模块可以不涉及参数调整。

[0110] 网络训练

[0111] 图3示例了本公开的一种特征图放大方法,该方法可以是待训练的神经网络在训练过程中对输入的样本图像提取得到源特征图,并进行了特征图放大。并且,在该神经网络的训练过程中,该神经网络中包括的特征图放大网络也一并进行了训练,对其中的参数进行了调整。结合图3和图2,该方法可以包括如下处理:

[0112] 在步骤200中,接收待放大的源特征图,所述源特征图是由神经网络对输入的样本图像进行特征提取得到。

[0113] 本步骤中,特征图放大网络的训练阶段,可以是将特征图放大网络集成在一个待训练的神经网络中使用。例如,可以是将该特征图放大网络集成在用于物体检测的FPN (Feature Pyramid Network) 网络中,或者,将该特征图放大网络集成在用于图像修复的网络中。即,可以将该特征图放大网络作为上述的神经网络的一部分,参与该神经网络的训练。

[0114] 该神经网络可以提取输入的样本图像的特征,得到所述的源特征图。再通过该特征图放大网络,将源特征图放大到目标特征图后,神经网络还可以继续根据所述的目标特征图继续处理,得到样本图像对应的预测结果。

[0115] 例如,请结合图2,源特征图的形状参数可以是 $C*H*W$ ,其中, $C$ 是源特征图的通道数, $H$ 是高度, $W$ 是宽度。 $X$ 可以是该源特征图的标识。

[0116] 在步骤202中,通过内容编码卷积层对所述源特征图进行卷积处理,分别得到所述源特征图中的每一源位置对应的 $N$ 个重组核。

[0117] 其中,内容编码卷积层是特征图放大网络的一部分,也就是待训练的神经网络的一部分。

[0118] 本步骤中,特征图放大网络的核预测模块11,可以包括“内容编码卷积层”,该内容编码卷积层可以通过卷积核对源特征图进行卷积处理。通过卷积处理,可以得到源特征图中的每一个源位置对应的 $N$ 个重组核。所述 $N$ 是源特征图的放大倍数的平方。

[0119] 例如,假设特征图放大网络要将源特征图放大两倍,那么放大倍数是2。源特征图中的每一个源位置都可以得到 $2*2=4$ 个重组核。所述的“一个源位置”可以是源特征图上的一个像素。

[0120] 由于重组核是通过内容编码卷积层对源特征图的不同位置的内容进行卷积处理得到,所以,对于源特征图的不同位置,内容发生变化,重组核也随之变化,要根据更新后的源特征图内容进行卷积处理获得。

[0121] 在步骤204中,对每一个重组核进行归一化处理,得到归一化后的重组核。

[0122] 本步骤中,每一个重组核中可以包括多个特征,可以对属于同一个重组核的多个特征进行归一化处理,以保证每一个重组核的特征和为1。由于对重组核进行了归一化,使得在对源特征图放大时,不会改变源特征图的特征平均值。

[0123] 在归一化之后,可以将重组核应用于重组模块12的特征重组处理。

[0124] 在步骤206中,对于所述源特征图的每个源位置,采用与所述源位置对应的归一化后的 $N$ 个重组核对重组区域的特征进行特征重组,得到对应所述源位置的 $N$ 个重组特征。

[0125] 经过上述步骤的处理,源特征图的每一个源位置都可以对应 $N$ 个重组核,且每一个重组核都进行了归一化处理。

[0126] 本步骤中,在进行特征重组处理时,对于源特征图的任一源位置,执行的处理可以包括如下:

[0127] 首先,可以获取以源位置为中心的重组区域,该重组区域的尺寸与重组核相同。例如,假设重组核是一个 $5*5$ 的尺寸,那么重组区域也是 $5*5$ 的尺寸。

[0128] 接着,分别通过源位置对应的归一化后的 $N$ 个重组核,对根据源位置确定的重组区域进行特征重组,一个重组核可以对应得到一个重组特征,共可以得到 $N$ 个重组特征。

[0129] 在步骤208中,根据所述源特征图的每个源位置对应的 $N$ 个重组特征,生成目标特征图。

[0130] 本步骤中,可以将步骤206中得到的所述 $N$ 个重组特征,填充入目标特征图中对应所述源位置的目标位置。所述的目标位置比源位置的面积要大。

[0131] 举例来说,以将源特征图放大两倍为例,源特征图中的一个像素可以对应目标特征图上的四个像素。请参见图4的示意,其中,放大后的目标位置的高度和宽度,都是源位置的高度和宽度的两倍。该源位置的一个像素可以对应四个重组核(图4仅是示意,实际重组核的尺寸可以大于源位置的尺寸),分别为重组核a、重组核b、重组核c和重组核d,这四个重组核分别与源位置对应的重组区域进行特征重组处理,共得到四个不同的重组特征。例如,

重组核a与重组区域进行特征重组处理后得到一个重组特征,重组核b与重组区域进行特征重组处理后得到另一个重组特征。该四个重组特征填充入目标位置的A、B、C和D四个区域,完成了将源位置放大到目标位置。

[0132] 源特征图的每一个源位置都进行上述的步骤206和步骤208的处理,就可以实现将源特征图放大到目标特征图。其中,当源位置位于源特征图的边缘时,获取重组区域时,对于重组区域超出源特征图的部分,可以填充为0。

[0133] 例如,请结合图2,目标特征图的形状参数可以是 $C \times \sigma H \times \sigma W$ ,其中,C是目标特征图的通道数, $\sigma H$ 是高度, $\sigma W$ 是宽度, $\sigma$ 是源特征图到目标特征图的放大倍数。 $X^2$ 可以是该目标特征图的标识。

[0134] 在步骤210中,根据所述目标特征图,得到所述样本图像的预测结果。

[0135] 本步骤中,特征图放大网络所在的神经网络,可以根据目标特征图,预测得到样本图像对应的预测结果。神经网络的具体场景不同,预测结果也可以不同。例如,预测结果可以是物体的类别,或者也可以是样本图像上的某个像素所属的类别,等。

[0136] 在步骤212中,基于所述样本图像的标注结果和预测结果之间的差别,调整神经网络的网络参数,该网络参数包括所述内容编码卷积层的参数。

[0137] 本步骤中,特征图放大网络所在的神经网络,可以根据样本图像的标注结果和预测结果之间的差别,反向传播调整神经网络的参数。其中,特征图放大网络是神经网络的一部分,该特征图放大网络中的核预测模块中包括的内容编码卷积层的参数,可以在训练过程中进行调整。

[0138] 本实施例的特征图放大网络的训练方法,通过根据源特征图的不同位置的内容进行卷积,得到对应该位置内容的重组核,使得即使是同一个源特征图,不同的源特征图位置获得的是不同的重组核,这区别于传统方式中对同一源特征图的所有位置都采用同一个核的方式,本实施例的重组核是与源特征图中的不同位置对应,并且是感知所述位置的源特征图内容得到,因此,在利用重组核对特征图进行特征重组(也就是放大时)时充分利用了特征图中的内容,从而使得利用该重组核对源特征图放大时,特征图放大的准确率更高,提升了特征图放大的性能。

[0139] 图5更详细的显示了通过特征图放大网络将源特征图放大到目标特征图的过程,其中,该图5相对于图2实施例的区别在于,在图2的特征图放大基础上,可以增加一些使得特征图放大操作性能更优的处理,以及对特征图放大操作的各个处理步骤分别描述了示例性的实践处理方式。图5中的核预测模块和重组模块都是特征图放大网络的组成部分。

[0140] 其中,源特征图和目标特征图与图2相同。

[0141] **【核预测模块的处理】**

[0142] 请参见图5,核预测模块(Kernel Prediction Module)对输入的源特征图可以先进行通道压缩(Channel Compressor),具体可以通过通道压缩卷积层对源特征图进行通道压缩。该通道压缩卷积层可以是一个 $1 \times 1$ 的卷积层。

[0143] 输入的源特征图的通道数量是C,经过通道压缩卷积层的处理后,输出通道数可以是 $C_m$ ,其中, $C_m < C$ 。通过对源特征图进行通道压缩,可以提高本公开的特征图放大算法的效率,并且,效率的提高可以允许算法感知更大范围的特征图内容,使得本CARAFE算法能够聚合较大的感受野范围内的源特征图内容信息,从而进一步提升算法性能。实验表明,通道压

缩不会损害算法性能。

[0144] 此外,核预测模块在增加了通道压缩卷积层之后,训练过程中在调整待训练的神经网络的网络参数时,该网络参数包括了前述的内容编码卷积层的参数和通道压缩卷积层的参数。

[0145] 请继续参见图5,压缩通道后的源特征图,继续通过核预测模块的内容编码卷积层进行卷积处理(Content Encoder)。

[0146] 如图6的示意,所述内容编码卷积层包括 $\sigma^2 * k_{up}^2 = C_{up}$ 个过滤器(Filter),其中, $\sigma^2$ 是源特征图的放大倍数的平方, $k_{up}^2$ 是一个重组核的尺寸。每一个过滤器包括 $C_m$ 个卷积核,每一个卷积核的尺寸是 $k_{encoder} * k_{encoder}$ 。在进行卷积处理时,每一个过滤器与源特征图进行卷积,得到一个通道的卷积输出特征图,该卷积输出特征图的尺寸是 $H * W$ 。内容编码卷积层总共可以得到 $C_{up}$ 个通道的卷积输出特征图,如图6所示,通过卷积处理,得到了 $H * W * \sigma^2 * k_{up}^2$ 的重组核图,也可以称为 $C_{up}$ 通道的重组核图。

[0147] 在得到重组核图后,继续获取源特征图每一个位置对应的 $N$ 个重组核:

[0148] 首先,一个卷积输出特征图可以包括很多像素,其中每一个像素的特征可以称为一个卷积输出特征。例如,请参见图7,示例了卷积输出特征图中的一个卷积输出特征61(对应图7中的一个小方格),另一个卷积输出特征62。

[0149] 接着,继续参见图7,重组核图的深度是 $\sigma^2 * k_{up}^2$ ,即源特征图的每一个位置(例如,每一个像素)对应着 $\sigma^2 * k_{up}^2$ 通道的卷积输出特征。当源特征图的放大倍数是 $\sigma$ 时,目标特征图的高度是 $\sigma H$ ,目标特征图的宽度是 $\sigma W$ ,那么可以将重组核图的深度变更为 $k_{up}^2$ 。具体的,可以将 $\sigma^2 * k_{up}^2$ 通道的重组核图在深度方向上进行切分,均分成 $N = \sigma^2$ 份。图7是以放大倍数是2为例,那就是均分为4份,每一份包括的通道数是 $k_{up}^2$ 。每一份可以称为一个重组核,该重组核包括 $k_{up}^2$ 通道的卷积输出特征。

[0150] 此外,图7中是以将重组核图在深度方向上均分后,相邻的通道合并在一起组成一个重组核,实际实施中不局限于此,也可以将深度方向上的各个通道自由组合,只要最终在通道数量上组合为均分的 $N$ 份即可。但是,需要保证源特征图的不同像素对应的通道采用相同的组合方式,比如,假设有四个通道0至3,如果把通道0和通道2作为一组,通道1和通道3作为一组,那么所有像素都要这样分组。

[0151] 通过上述的内容编码卷积层的处理,源特征图的每一个源位置都可以分别得到对应的 $N$ 个重组核。此外,实验表明,当卷积核的尺寸是 $K_{encoder} * K_{encoder}$ ,所述重组核的尺寸是 $K_{up} * K_{up}$ 时,设置所述 $K_{encoder}$ 等于 $K_{up}$ 减2,可以较好的达到“感知更大的感受野”与“提高算法的效率”之间的平衡。当然,本实施例不限制实际实施中 $K_{encoder}$ 与 $K_{up}$ 的取值,不过,增大 $K_{encoder}$ ,可以提升卷积核的感受野,从而使得能够感知更大区域范围的源特征图内容,对应获得的重组核用于特征图放大的效果也更好。

[0152] 请继续参见图5,在得到源特征图的每一个源位置对应的重组核后,还可以继续对每一个重组核进行归一化处理(Kernel Normalizer)。如上所述,每一个重组核可以包括通道数是 $k_{up}^2$ 的卷积输出特征,可以对这些卷积输出特征进行归一化。本实施例不限制归一化的具体方式,示例性的,可以通过softmax函数进行归一化。

[0153] 在获得源特征图的每一个源位置的重组核并进行归一化之后,接着可以通过重组模块,利用重组核进行源特征图的放大处理。



[0154] 【重组模块的处理】

[0155] 请继续参见图5,重组模块(Reassembly Modules)在对源特征图放大时,可以对源特征图中的每一个源位置都进行图8所示的处理。

[0156] 在步骤700中,获取与所述源位置对应的N个重组核,对于任一个重组核,都执行步骤702至708。

[0157] 本步骤中,同一个源位置对应N个重组核。例如,当放大倍数是2时,N=4,一个源位置对应4个重组核。每一个重组核都对重组区域的特征进行特征重组处理,得到对应的重组特征。

[0158] 在步骤702中,将所述重组核中包括的各个通道的卷积输出特征平铺,得到二维的重组核。

[0159] 请参见图5所示,如上面提到的,一个重组核包括 $k_{up}^2$ 通道的卷积输出特征,可以将这些卷积输出特征平铺成一个二维的重组核,重组核的尺寸是 $K_{up} * K_{up}$ 。

[0160] 在步骤704中,获取以所述源位置为中心的重组区域,所述重组区域的尺寸与所述重组核相同。

[0161] 本步骤中,重组区域是以源位置为中心,且该重组区域的尺寸同样也是 $K_{up} * K_{up}$ 。图5中的 $N(X_1, k_{up})$ ,其中,N表示一个特征重组的函数,例如,可以是加权和,后续描述。 $X_1$ 表示以源位置 $l(i, j)$ 为中心的重组区域, $k_{up}$ 表示该重组区域的尺寸也是 $K_{up} * K_{up}$ 。

[0162] 在步骤706中,将所述重组核分别与源特征图的各个通道的所述重组区域进行特征重组,得到分别对应每一个通道的通道特征。

[0163] 本步骤中,源特征图的通道数是C,需要注意的是,重组模块的处理是针对未进行通道压缩的源特征图,而上面提到的通道压缩是在核预测模块进行卷积处理之前进行。

[0164] 一个重组核分别与源特征图的各个通道的所述重组区域进行特征重组,每一个通道的重组区域在进行特征重组后可以得到一个对应的通道特征。如图9所示,示例了一个通道特征81,该通道特征81的面积对应于目标位置的 $1/N$ ,比如,放大倍数是2时,通道特征81的面积相当于目标位置面积的 $1/4$ 。各个通道的通道特征组合起来就可以形成图9所示的立方体的形状,共C个通道。

[0165] 本步骤中,一个重组核与一个通道的重组区域在进行特征重组时,使用的特征重组方式包括但不限于如下示例的两种:

[0166] 例如,可以将重组核与所述重组区域中的对应位置的特征进行加权求和,将重组核中的各个卷积输出特征作为权重,并将加权求和的结果作为所述通道特征。

$$[0167] \quad N(X_l, k_{up}) = \sum_{n=-r}^r \sum_{m=-r}^r W_{l'(n,m)} \cdot X_{(i+n,j+m)} \dots\dots\dots (1)$$

[0168] 其中, $r = \lfloor k_{up}/2 \rfloor$ ;  $w_{l'(n,m)}$ 表示重组核上的任一位置,例如可以是图5中的重组核 $w_{l'}$ 中的任一个小方格。

[0169] 如上所述,本公开实施例的特征图放大方法与反卷积相比,反卷积的卷积核在训练时确定,在测试过程中不能改变,比如,不论输入的是哪个图像,都采用该确定的卷积核

进行特征图放大。而本公开的特征图放大方法的重组核在测试时会根据图像的内容预测得到,因为本方法的重组核的生成过程是通过根据源特征图的内容卷积处理得到,不同的特征图内容就会得到不同的卷积结果即重组核,所以该方法在特征图放大过程中更能根据图像的内容做出不同的重组操作。实践证明,本公开的特征图放大方法的效果显著好于反卷积。

[0170] 又例如,还可以是由重组核的多个卷积输出特征中,确定数值最大的卷积输出特征,并找到与该最大值的卷积输出特征位置对应的重组区域中的特征,作为所述通道特征。

[0171] 在步骤708中,将各个通道对应的所述通道特征组合,得到所述源位置对应的一个重组特征。例如,图9所示即为一个重组特征,该重组特征包括C个通道的通道特征。

[0172] 每一个重组核都可以得到一个重组特征,N个重组核共得到N个重组特征。可以将一个源位置对应的N个重组特征,填充入目标特征图中对应所述源位置的目标位置。

[0173] 例如,类似图4所示,四个重组特征可以填充入目标位置的A、B、C和D四个区域。本实施例不限制所述的四个重组特征在这四个区域的填充顺序,只要保证每个像素使用一样的填充顺序就可以。可以根据源位置与目标位置之间的对应关系,确定目标位置:根据所述源位置和放大倍数,得到目标特征图上的所述目标位置。

[0174] 例如,对于目标特征图上的每一个位置 $l'(i', j')$ 它对应于源特征图上的位置 $l(i, j)$ ,其中 $i = \lfloor \frac{i'}{\sigma} \rfloor$ ,  $j = \lfloor \frac{j'}{\sigma} \rfloor$ ,  $\sigma$ 是放大倍数。类似于图4所示,源特征图上的一个源位置对应目标特征图上的一个目标位置,并且该目标位置的区域面积是源位置区域面积的 $\sigma$ 的平方倍。

[0175] 对于源特征图的每一个源位置都进行上述的操作,就可以实现将源特征图放大到目标特征图。

[0176] 当特征图放大网络训练完成后,就可以应用于特征图放大的任务。特征图放大网络在实际使用时,任何涉及到特征放大的操作都可以使用该特征图放大网络,例如,可以将该特征图放大网络集成于某个已有的神经网络(例如,FPN)使用,只要把该神经网络中使用到特征放大的操作替换为本实施例的特征图放大网络就可以。

[0177] 当特征图放大网络集成于不同的神经网络,或者用于完成不同的任务时,可以分别通过训练该神经网络从而训练该神经网络中集成的特征图放大网络。

[0178] 网络应用

[0179] 可以使用上述本公开的任一实施例描述的特征图放大网络,进行特征图的放大。例如,图10描述了一个例子的特征图放大方法,该方法可以是训练完成的神经网络对输入图像提取得到源特征图,并进行特征图放大后的相关处理。如图10所示,该方法可以包括:

[0180] 在步骤900中,训练完成的神经网络从输入图像中提取源特征图。

[0181] 其中,所述训练完成的神经网络中包含采用本公开任一实施例所述的方法训练的内容编码卷积层。

[0182] 所述的源特征图可以由神经网络对输入图像直接提取特征得到,或者还可以是,神经网络中包括多个特征图放大网络,从而对提取的源特征图进行多次放大。

[0183] 例如,假设要将源特征图放大4倍,而训练的特征图放大网络可以将源特征图放大2倍,则可以连续使用两次该特征图放大网络。对于神经网络提取特征得到的特征图,可以

先通过特征图放大网络将特征图放大2倍,在此基础上,再使用特征图放大网络继续放大2倍。

[0184] 在步骤902中,通过内容编码卷积层对所述源特征图进行卷积处理,分别得到所述源特征图中的每一源位置对应的N个重组核。

[0185] 本步骤可以利用训练完成的神经网络中的内容编码卷积层对源特征图进行卷积处理。在另一个例子中,在卷积处理之前,还可以通过训练完成的神经网络中的通道压缩卷积层对所述源特征图进行通道压缩,而内容编码卷积层可以是对通道压缩后的源特征图进行卷积处理。

[0186] 训练完成的神经网络的网络参数已经确定,因此,训练完成的神经网络中集成的特征图放大网络中的网络参数也已经确定,特征图放大网络中的内容编码卷积层的参数和通道压缩卷积层的参数也已经确定。

[0187] 在步骤904中,对每一个重组核进行归一化处理,得到归一化后的重组核。

[0188] 在步骤906中,对于所述源特征图的每个源位置,采用与所述源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组,得到对应所述源位置的N个重组特征。

[0189] 步骤908中,根据所述源特征图的每个源位置对应的N个重组特征,生成目标特征图。

[0190] 步骤910中,根据所述目标特征图,得到所述输入图像的处理结果。

[0191] 如下示例几种神经网络根据目标特征图得到处理结果的例子:

[0192] 例如,在物体检测的神经网络中,神经网络的输入图像可以是包括待检测的至少一个目标物体,神经网络可以首先提取输入图像的特征,得到源特征图。该神经网络中包括上述的CARAFE对应的特征图放大网络,可以通过该特征图放大网络得到放大后的目标特征图。训练完成的神经网络还可以根据该目标特征图继续处理,输出输入图像中的每个目标物体的检测框、以及所述目标物体的所属类别概率。

[0193] 又例如,在图像修复的神经网络中,神经网络的输入图像可以包括残缺部分,神经网络可以通过其主干网络提取输入图像的特征,得到源特征图。该神经网络中包括上述的CARAFE对应的特征图放大网络,可以通过该特征图放大网络得到放大后的目标特征图。神经网络还可以根据该目标特征图继续处理,输出所述输入图像的残缺部分的像素。

[0194] 再例如,在语义预测的神经网络中,神经网络可以通过包括的特征图放大网络得到放大后的目标特征图,并可以根据该目标特征图继续处理,输出所述输入图像上各个像素的所属类别。

[0195] 还例如,在实例分割的神经网络中,神经网络可以通过包括的特征图放大网络得到放大后的目标特征图,并可以根据该目标特征图继续处理,输出所述输入图像中的每个实例、以及各个实例内的各像素的所属类别。

[0196] 上述本说明书实施例提供的特征图放大算法可以称为CARAFE算法。该CARAFE的一种应用场景是FPN(feature pyramid network)(特征金字塔网络),这种网络的结构如图11所示。在FPN中,首先获得来输入图像的多层特征(c2,c3,c4,c5),其中,从c2开始每一层特征的大小依次缩小2倍。

[0197] FPN通过下采样c5(例如max pool(最大值池化),average pool(平均值池化))得到p6,然后将各层特征依次向上一级特征传递,传递的过程是将较小的特征图放大2倍,然

后与较大的特征图相加。

[0198] 在标准的FPN中,上述的放大操作使用最近邻插值实现。CARAFE在FPN中的使用,是通过将最近邻插值直接用CARAFE替换,并在训练过程中同时将CARAFE的参数和FPN的参数进行优化。

[0199] FPN的一种主要使用场景是物体检测任务中著名的Faster RCNN算法,我们将CARAFE使用在FPN中,并和原本的FPN在Faster RCNN上进行对比。在MS COCO 2017数据集上,使用MS COCO官方的评价标准,即从IoU 0.5到0.95的mAP平均值。(注:intersection-over-union (IoU):检测框和目标框之间的面积交集/面积并集,mean average precision (mAP):每一类平均精度 (average precision) 在各类上的平均 (mean),满分100,最少0。对于每一个目标框,当至少有一个检测框成功检测该目标框的类别,同时与该目标框的iou大于特定阈值,则认为该物体被检测到)。在测试集 (test-dev) 上,CARAFE可以将算法的精度从36.9提高到38.1。该实验表明将CARAFE使用在FPN结构上可以显著提高Faster RCNN的算法精度。

[0200] 本说明书不再列举其他的CARAFE的应用实例,实际试验中,该CARAFE算法计算的资源开销较小,运行速度快,可以很容易的集成应用于物体检测任务、图像修复任务、语义分割任务和实例分割任务等多种任务的网络结构中。并且通过对在多种任务中的CARAFE算法的表现进行评估,该算法使得各个任务都获得了较为一致和稳固的性能提升,效果很好。

[0201] 图12提供了一种特征图放大装置,如图12所示,该装置可以包括:源图接收模块1201、核生成模块1202、归一处理模块1203、特征重组模块1204和目标图生成模块1205。

[0202] 源图接收模块1201,用于接收待放大的源特征图,所述源特征图是对输入的样本图像进行特征提取得到的;

[0203] 核生成模块1202,用于对所述源特征图进行卷积处理,分别得到源特征图中的每一个源位置对应的N个重组核,所述N是源特征图的放大倍数的平方;

[0204] 归一处理模块1203,用于对每一个重组核进行归一化处理,得到归一化后的重组核;

[0205] 特征重组模块1204,用于对于所述源特征图的每个源位置,采用与所述源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组,得到对应所述源位置的N个重组特征;所述重组区域根据所述源位置确定;

[0206] 目标图生成模块1205,用于根据所述源特征图的每个源位置对应的N个重组特征,生成目标特征图。

[0207] 在一个例子中,核生成模块1202,具体用于:以 $C_{up}$ 个过滤器对所述源特征图进行卷积处理,得到 $C_{up}$ 通道的重组核图;所述 $C_{up}$ 是根据所述放大倍数和重组核的尺寸确定;将所述重组核图中,对应于所述源特征图同一源位置的 $C_{up}$ 通道的卷积输出特征,均分成N份,每一份作为一个所述重组核,一个所述重组核包括 $C_{up}/N$ 通道的卷积输出特征。

[0208] 在一个例子中,特征重组模块1204,还用于:在采用与所述源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组之前,对于与所述源位置对应的每一个重组核,将所述重组核中包括的 $C_{up}/N$ 通道的卷积输出特征平铺,得到二维的所述重组核;获取以所述源位置为中心的所述重组区域,所述重组区域的尺寸与所述重组核相同。

[0209] 在一个例子中,所述归一处理模块1203具体用于:对每一个重组核中的 $C_{up}/N$ 通道

的卷积输出特征进行归一化,得到归一化后的重组核。

[0210] 在一个例子中,特征重组模块1204,用于在采用与一个源位置对应的归一化后的N个重组核对重组区域的特征进行特征重组,得到对应所述源位置的N个重组特征时,包括:对于每一个所述重组核,将所述重组核分别与源特征图的各个通道的所述重组区域进行特征重组,得到与所述重组核对应的所述各个通道的通道特征;将所述重组核对应的各个通道的所述通道特征组合,得到与所述重组核对应的重组特征。

[0211] 在一个例子中,特征重组模块1204,用于在将所述重组核分别与源特征图的各个通道的所述重组区域进行特征重组,得到与所述重组核对应的所述各个通道的通道特征时,包括:对于其中一个通道,将所述重组核与所述通道的重组区域中的对应位置的特征进行加权求和,所述加权求和的结果作为所述通道对应的通道特征。

[0212] 在一个例子中,目标图生成模块1205,具体用于:根据所述源位置和放大倍数,得到目标特征图中与所述源位置对应的所述目标位置;将所述N个重组特征,填充入目标特征图中的所述目标位置。

[0213] 在一个例子中,如图13所示,所述特征图放大装置包含于待训练的神经网络中,该装置还包括预测处理模块1206和参数调整模块1207。

[0214] 源图接收模块1201,接收到的源特征图是由所述待训练的神经网络对输入的样本图像进行特征提取得到的。

[0215] 核生成模块1202用于在对所述源特征图进行卷积处理时,具体包括:通过所述待训练的神经网络中的内容编码卷积层对所述源特征图进行卷积处理。

[0216] 预测处理模块1206,用于在目标图生成模块生成目标特征图之后,根据所述目标特征图,得到所述样本图像的预测结果;

[0217] 所述参数调整模块1207,用于基于所述样本图像的标注结果和预测结果之间的差别,调整所述待训练的神经网络中的网络参数,所述网络参数包括所述内容编码卷积层的参数。

[0218] 在一个例子中,如图13所示,该装置还包括:通道压缩模块1208,用于通过所述待训练的神经网络中的通道压缩卷积层对所述源特征图进行通道压缩;

[0219] 所述核处理模块,用于在进行卷积处理时,具体包括:通过内容编码卷积层对通道压缩后的源特征图进行卷积处理;

[0220] 所述参数调整模块,还用于基于所述样本图像的标注结果和预测结果之间的差别,调整所述通道压缩卷积层的参数。

[0221] 在一个例子中,如图14所示,所述特征图放大装置包含于训练完成的神经网络,所述源图接收模块1201接收到的源特征图是由所述训练完成的所述神经网络对输入的样本图像进行特征提取得到的;所述核生成模块1202用于在对所述源特征图进行卷积处理时,具体包括:通过所述训练完成的神经网络中的内容编码卷积层对所述源特征图进行卷积处理。该装置还包括:结果输出模块1206,用于在目标图生成模块生成目标特征图之后,根据所述目标特征图,得到输入图像的处理结果。

[0222] 在一个例子中,如图14所示,该装置还可以包括:通道压缩模块1207,用于通过所述训练完成的神经网络中的通道压缩卷积层对所述源特征图进行通道压缩。所述核处理模块1202,用于在进行卷积处理时,具体包括:通过所述内容编码卷积层对通道压缩后的源特

征图进行卷积处理。

[0223] 在一个例子中,所述输入图像中包括待检测的至少一个目标物体;所述结果输出模块1206,具体用于:根据所述目标特征图,输出所述输入图像中的每个目标物体的检测框、以及所述目标物体的所属类别。

[0224] 在一个例子中,所述输入图像包括残缺部分;结果输出模块1206,具体用于:根据所述目标特征图,输出所述输入图像的残缺部分的像素分。

[0225] 在一个例子中,结果输出模块1206,具体用于:根据所述目标特征图,输出所述输入图像中的每个实例、以及各个实例内的各像素的所属类别。

[0226] 在一个例子中,结果输出模块1206,具体用于:根据所述目标特征图,输出所述输入图像上各个像素的所属类别。

[0227] 本公开还提供了一种特征图放大设备,所述设备包括存储器、处理器,所述存储器用于存储可在处理器上运行的计算机指令,所述处理器用于在执行所述计算机指令时实现本公开任一实施例的特征图放大方法。

[0228] 本公开还提供了一种计算机可读存储介质,其上存储有计算机程序,所述程序被处理器执行时实现本公开任一实施例的特征图放大方法。

[0229] 本领域技术人员应明白,本公开一个或多个实施例可提供为方法、系统或计算机程序产品。因此,本公开一个或多个实施例可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且,本公开一个或多个实施例可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0230] 本公开实施例还提供一种计算机可读存储介质,该存储介质上可以存储有计算机程序,所述程序被处理器执行时实现本公开任一实施例描述的用于文字识别的神经网络的训练方法的步骤,和/或,实现本公开任一实施例描述的文字识别方法的步骤。其中,所述的“和/或”表示至少具有两者中的其中一个,例如,“N和/或B”包括三种方案:N、B、以及“N和B”。

[0231] 本公开中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于数据处理设备实施例而言,由于其基本相似于方法实施例,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0232] 上述对本公开特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一些情况下,在权利要求书中记载的行为或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外,在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中,多任务处理和并行处理也是可以的或者可能是有利的。

[0233] 本公开中描述的主题及功能操作的实施例可以在以下中实现:数字电子电路、有形体现的计算机软件或固件、包括本公开中公开的结构及其结构性等同物的计算机硬件、或者它们中的一个或多个的组合。本公开中描述的主题的实施例可以实现为一个或多个计算机程序,即编码在有形非暂时性程序载体上以被数据处理装置执行或控制数据处理装置的操作的计算机程序指令中的一个或多个模块。可替代地或附加地,程序指令可以被编码

在人工生成的传播信号上,例如机器生成的电、光或电磁信号,该信号被生成以将信息编码并传输到合适的接收机装置以由数据处理装置执行。计算机存储介质可以是机器可读存储设备、机器可读存储基板、随机或串行存取存储器设备、或它们中的一个或多个的组合。

[0234] 本公开中描述的处理及逻辑流程可以由执行一个或多个计算机程序的一个或多个可编程计算机执行,以通过根据输入数据进行操作并生成输出来执行相应的功能。所述处理及逻辑流程还可以由专用逻辑电路—例如FPGN(现场可编程门阵列)或NSIC(专用集成电路)来执行,并且装置也可以实现为专用逻辑电路。

[0235] 适合用于执行计算机程序的计算机包括,例如通用和/或专用微处理器,或任何其他类型的中央处理单元。通常,中央处理单元将从只读存储器和/或随机存取存储器接收指令和数据。计算机的基本组件包括用于实施或执行指令的中央处理单元以及用于存储指令和数据的一个或多个存储器设备。通常,计算机还将包括用于存储数据的一个或多个大容量存储设备,例如磁盘、磁光盘或光盘等,或者计算机将可操作地与此大容量存储设备耦接以从其接收数据或向其传送数据,抑或两种情况兼而有之。然而,计算机不是必须具有这样的设备。此外,计算机可以嵌入在另一设备中,例如移动电话、个人数字助理(PDN)、移动音频或视频播放器、游戏操纵台、全球定位系统(GPS)接收机、或例如通用串行总线(USB)闪存驱动器的便携式存储设备,仅举几例。

[0236] 适合于存储计算机程序指令和数据的计算机可读介质包括所有形式的非易失性存储器、媒介和存储器设备,例如包括半导体存储器设备(例如EPROM、EEPROM和闪存设备)、磁盘(例如内部硬盘或可移动盘)、磁光盘以及CD ROM和DVD-ROM盘。处理器和存储器可由专用逻辑电路补充或并入专用逻辑电路中。

[0237] 虽然本公开包含许多具体实施细节,但是这些不应被解释为限制任何公开的范围或所要求保护的公开的范围,而是主要用于描述特定公开的具体实施例的特征。本公开内在多个实施例中描述的某些特征也可以在单个实施例中被组合实施。另一方面,在单个实施例中描述的各种特征也可以在多个实施例中分开实施或以任何合适的子组合来实施。此外,虽然特征可以如上所述在某些组合中起作用并且甚至最初如此要求保护,但是来自所要求保护的组合中的一个或多个特征在一些情况下可以从该组合中去除,并且所要求保护的组合可以指向子组合或子组合的变型。

[0238] 类似地,虽然在附图中以特定顺序描绘了操作,但是这不应被理解为要求这些操作以所示的特定顺序执行或顺次执行、或者要求所有例示的操作被执行,以实现期望的结果。在某些情况下,多任务和并行处理可能是有利的。此外,上述实施例中的各种系统模块和组件的分离不应被理解为在所有实施例中均需要这样的分离,并且应当理解,所描述的程序组件和系统通常可以一起集成在单个软件产品中,或者封装成多个软件产品。

[0239] 由此,主题的特定实施例已被描述。其他实施例在所附权利要求书的范围以内。在某些情况下,权利要求书中记载的动作可以以不同的顺序执行并且仍实现期望的结果。此外,附图中描绘的处理并非必需所示的特定顺序或顺次顺序,以实现期望的结果。在某些实现中,多任务和并行处理可能是有利的。

[0240] 以上所述仅为本公开一个或多个实施例的较佳实施例而已,并不用以限制本公开一个或多个实施例,凡在本公开一个或多个实施例的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本公开一个或多个实施例保护的公开范围之内。

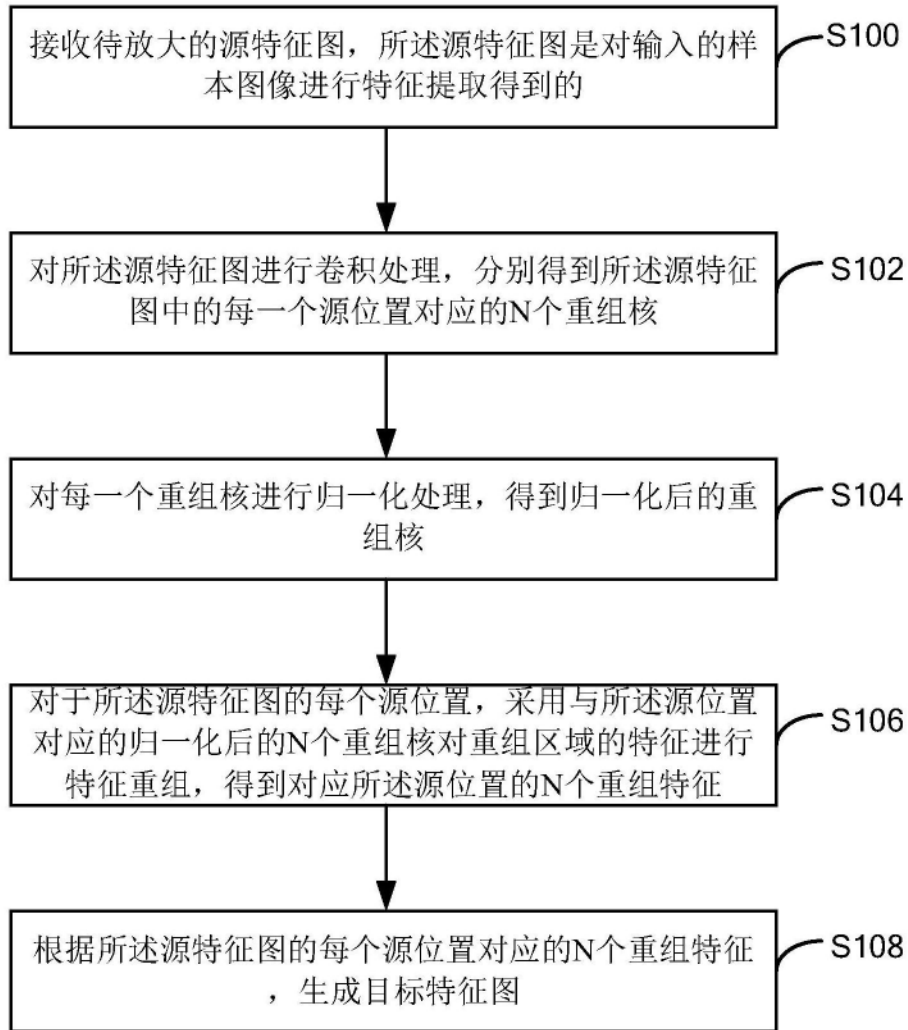


图1



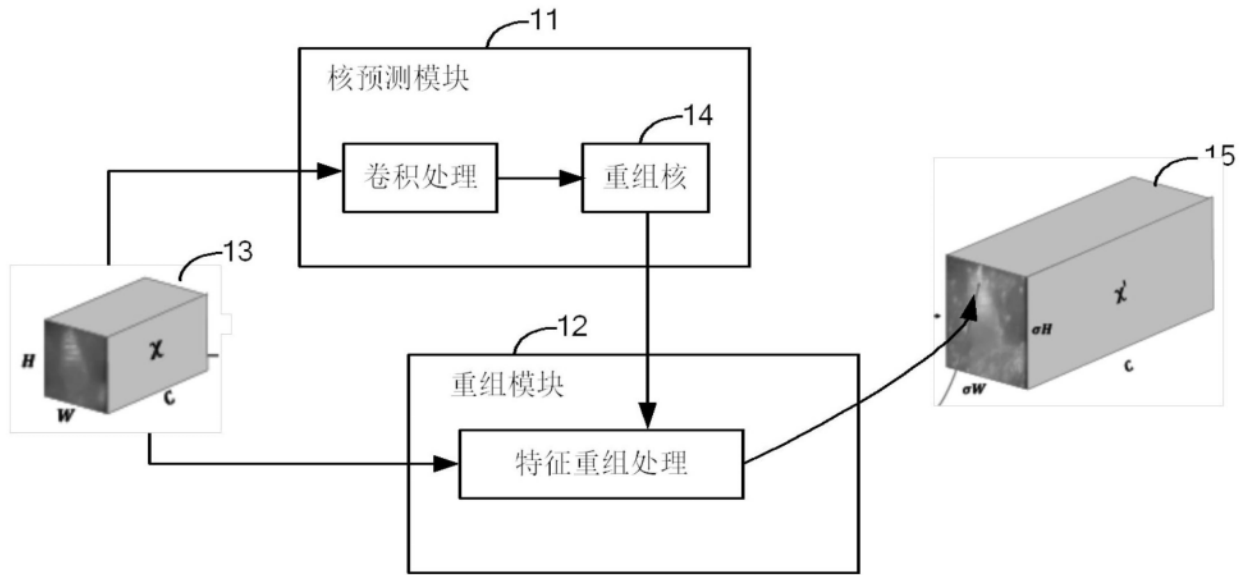


图2

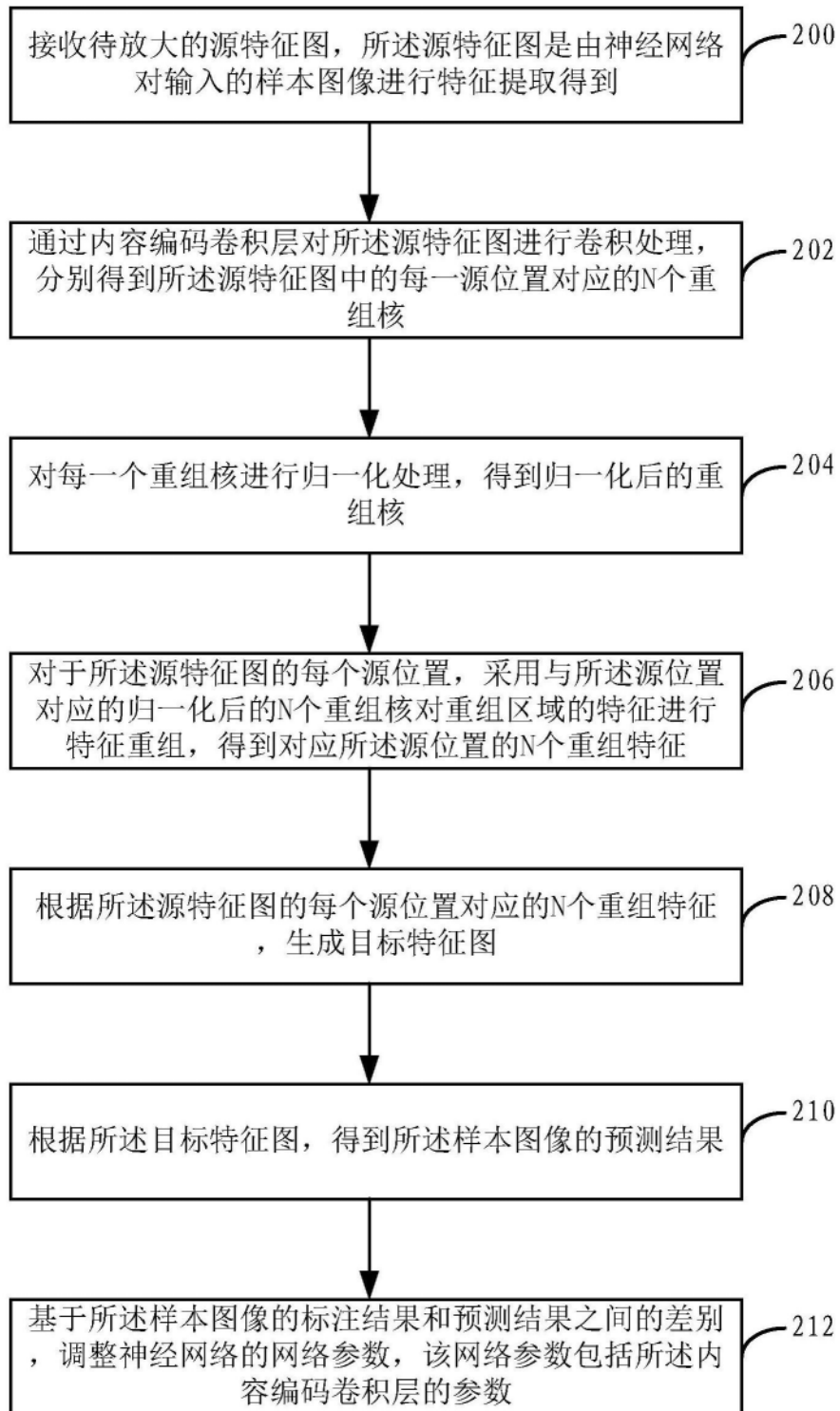


图3

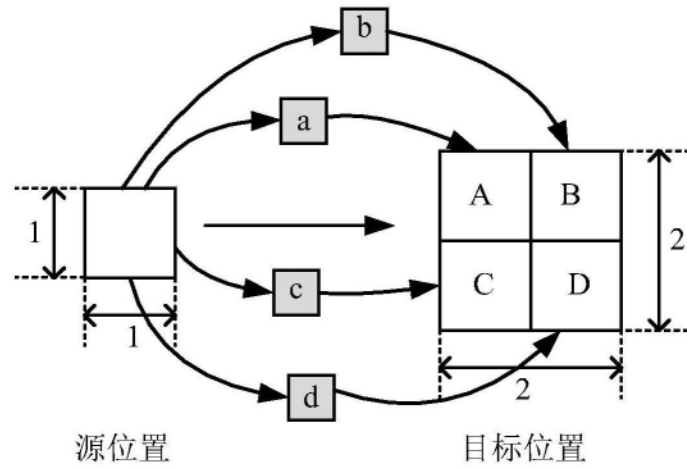


图4

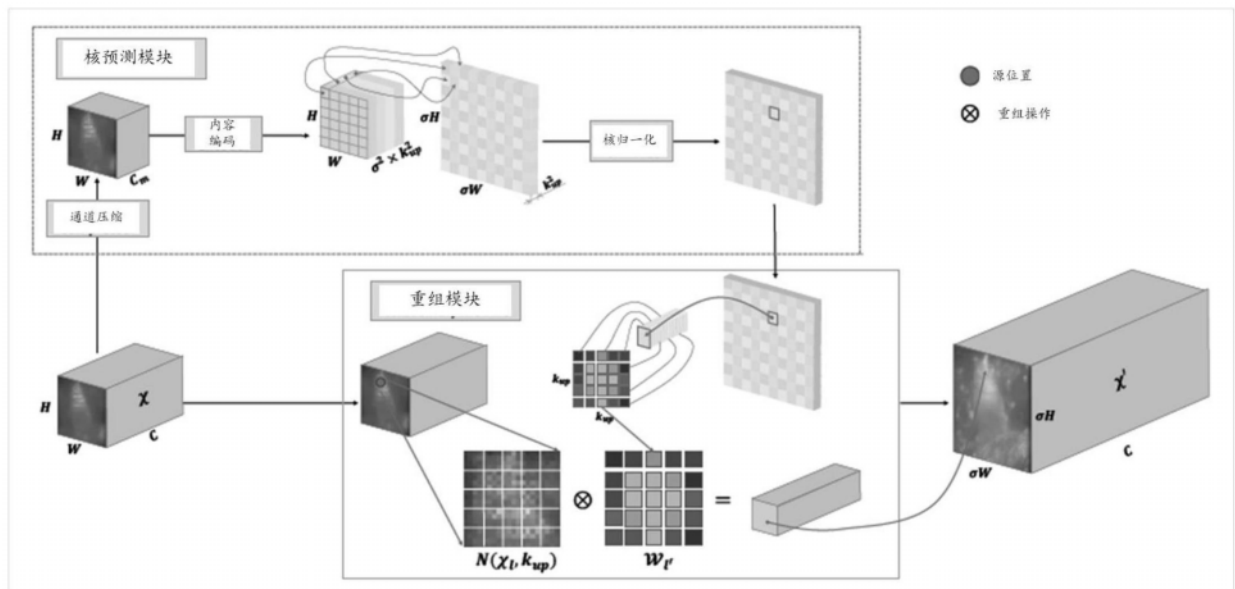


图5

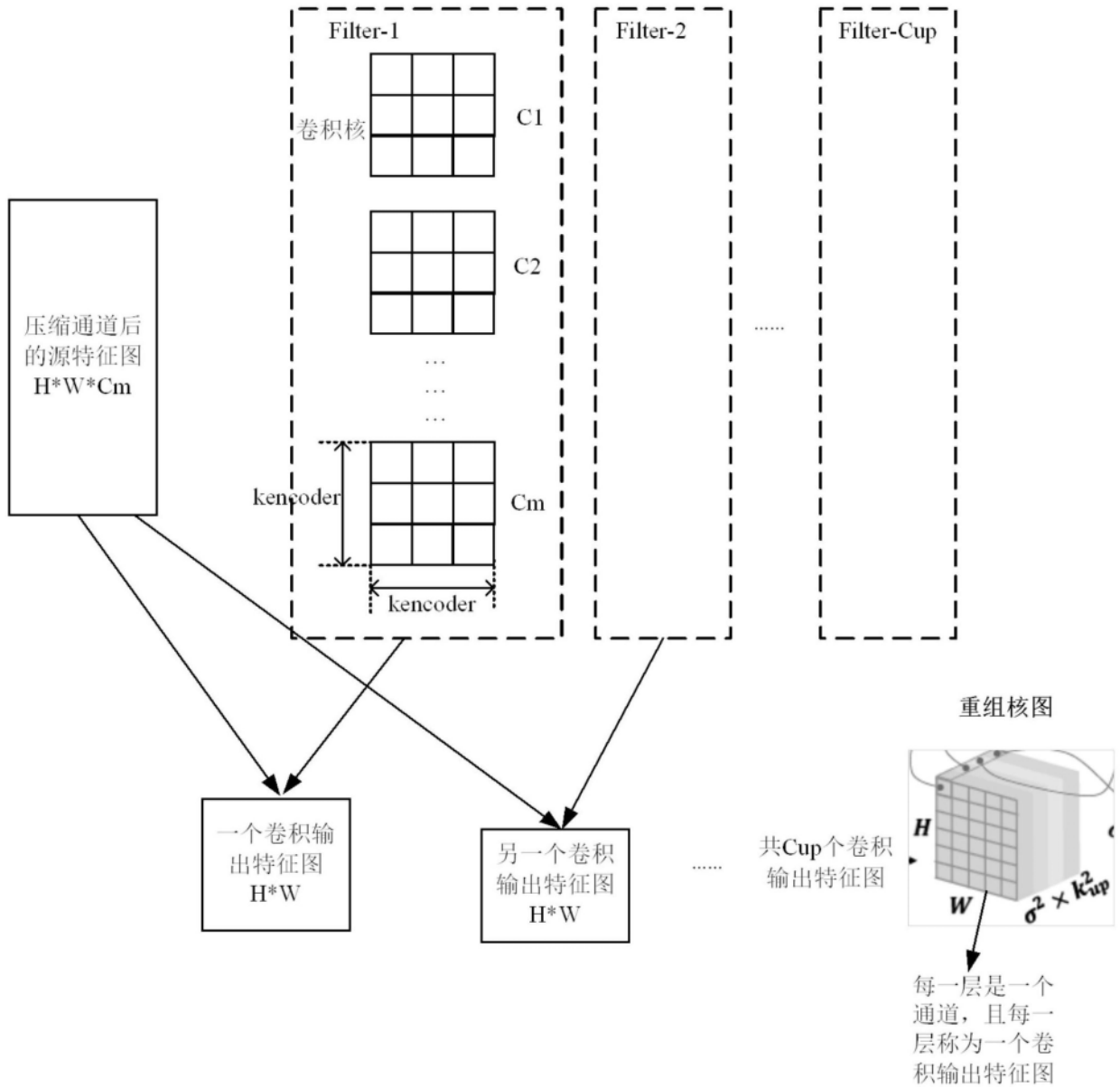


图6

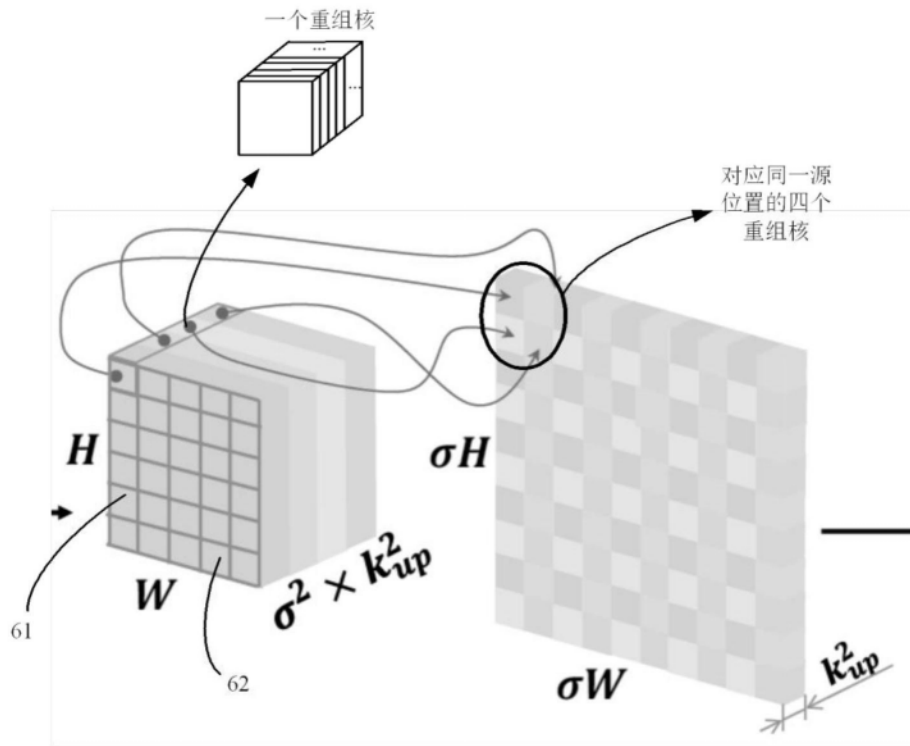


图7

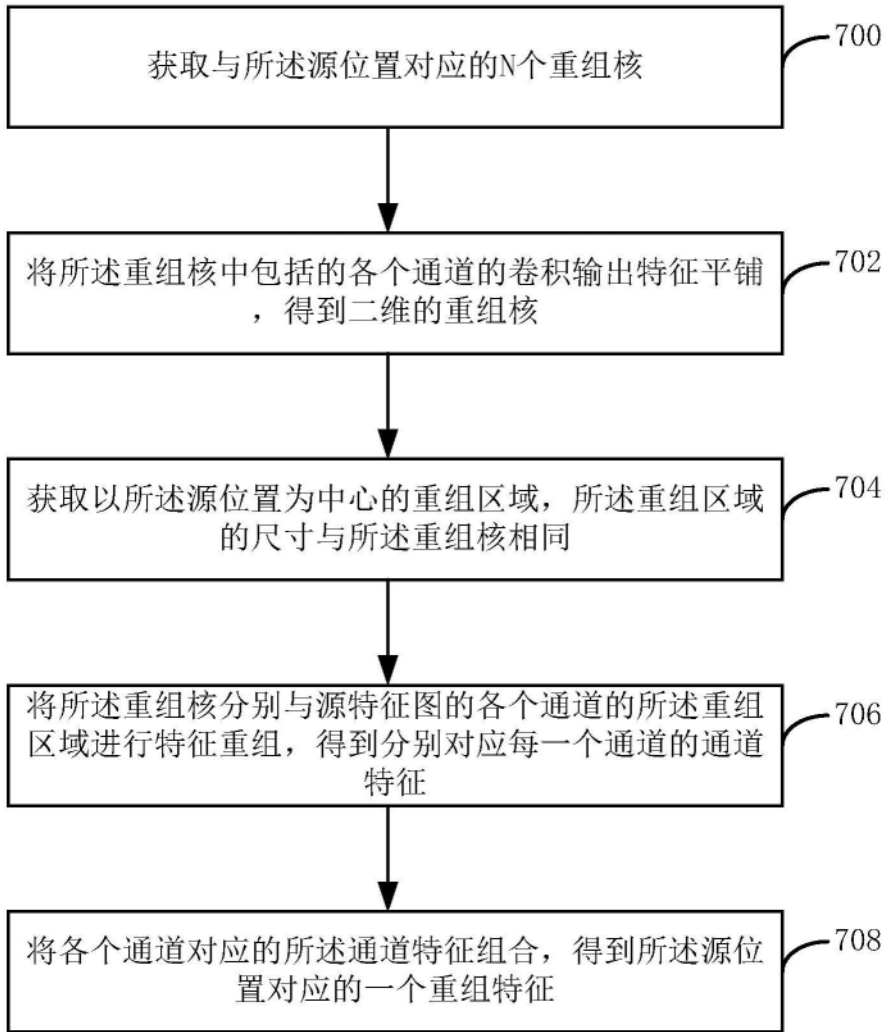


图8

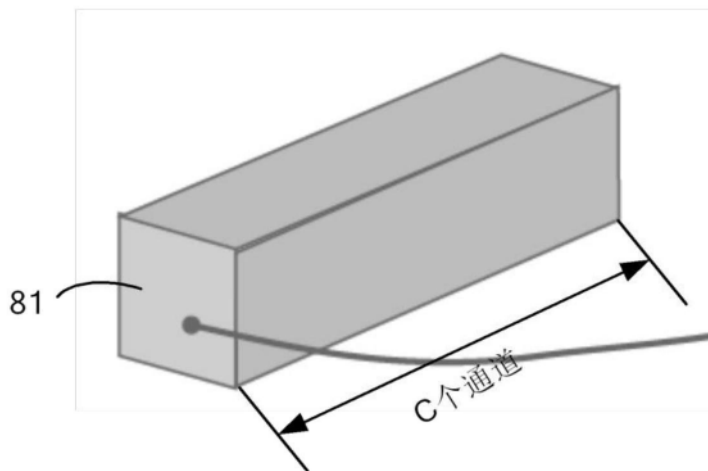


图9

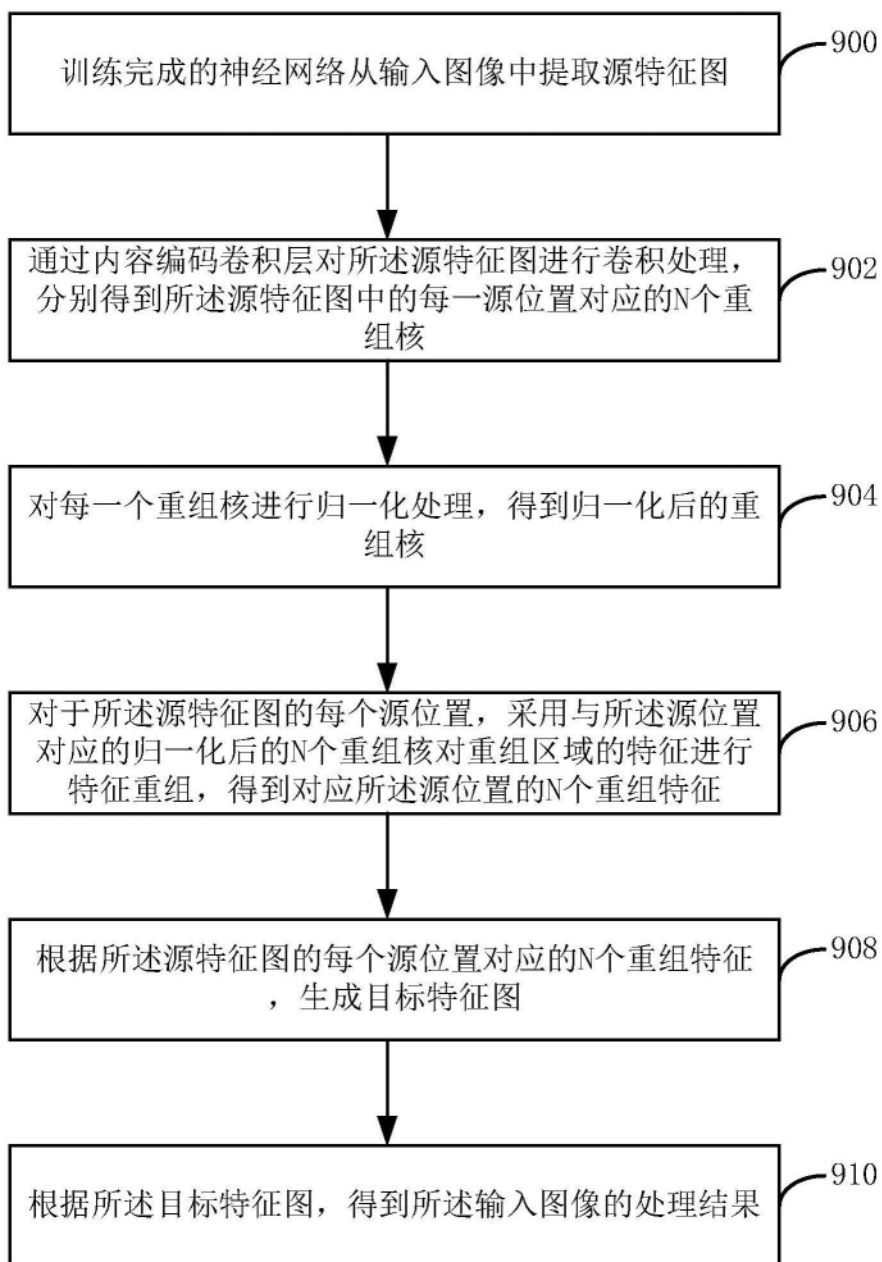


图10

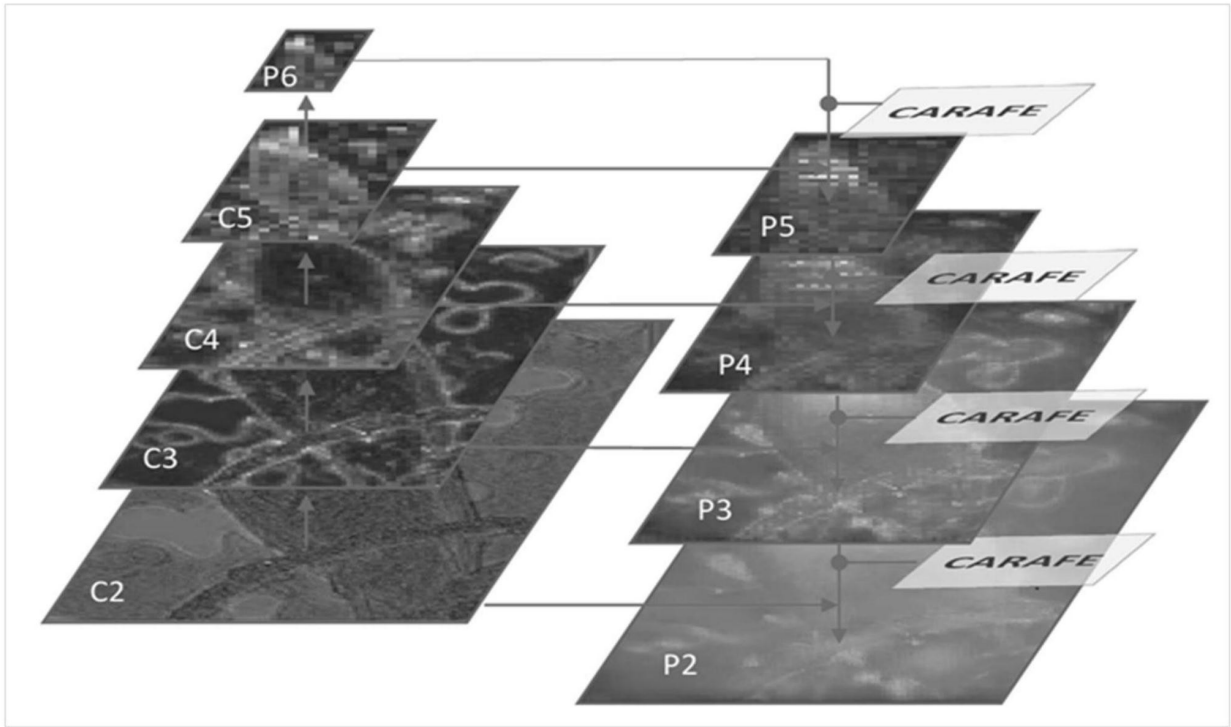


图11

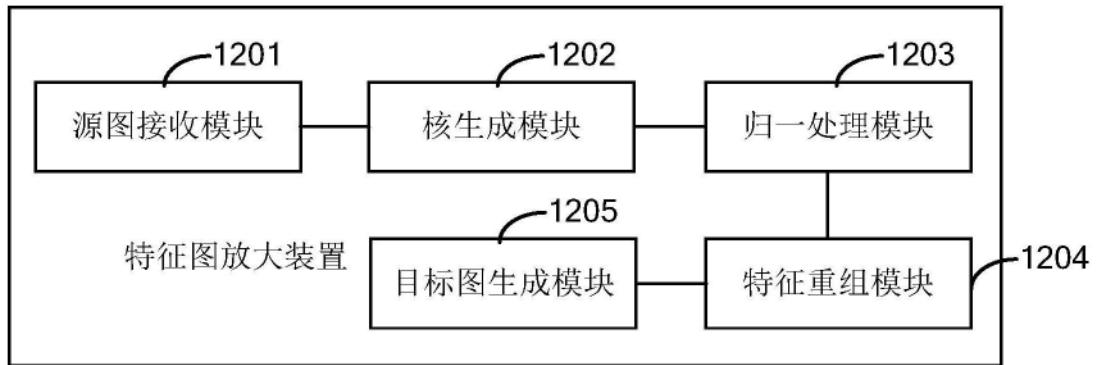


图12



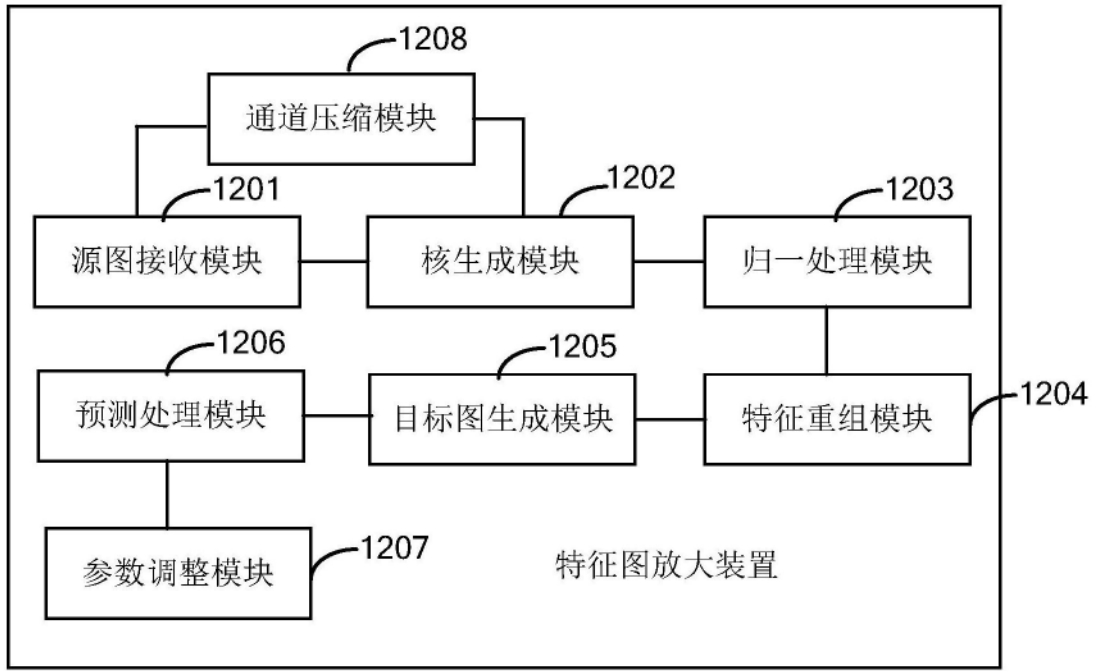


图13

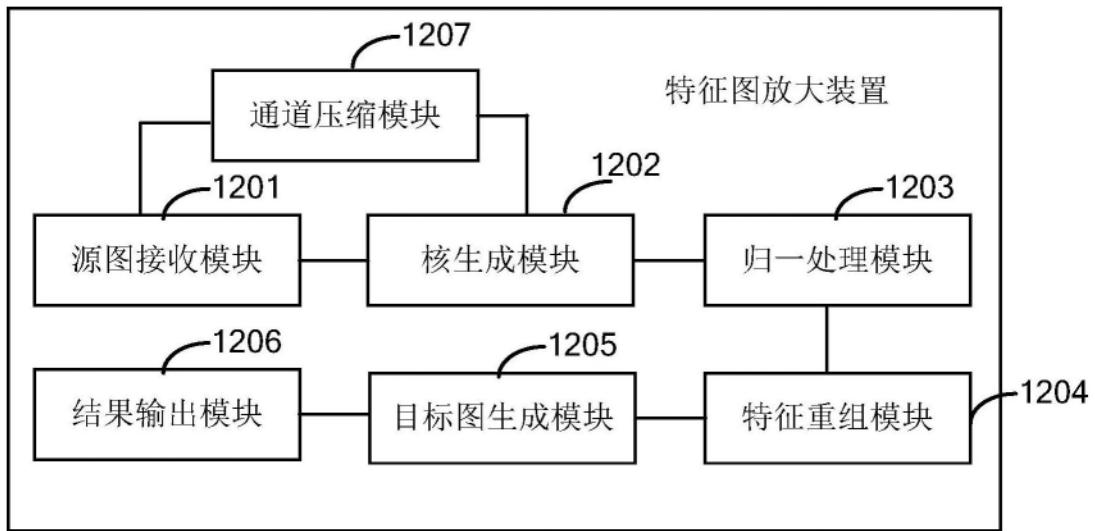


图14