图像处理方法及装置、电子设备和存储介质

本申请要求在2019年7月18日提交中国专利局、申请号为201910652025.2、发明名称为“图像处理方法及装置、电子设备和存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本公开涉及计算机技术领域，尤其涉及一种图像处理方法及装置、电子设备和存储介质。

背景技术

在深度学习网络中，可以对输入的特征图从某一固定的维度进行归一化计算，不仅能够加快模型的收敛速度，还能够缓解深层网络中的“梯度弥散”问题，从而更易于训练深度神经网络并得到更稳定的网络。

发明内容

本公开提出了一种图像处理技术方案。

根据本公开的一方面，提供了一种图像处理方法，包括：对待处理图像进行特征提取，得到所述待处理图像的第一特征图；根据所述第一特征图的维度信息及预设的拆分规则，将所述第一特征图拆分为多个第一子特征图，所述第一特征图的维度信息包括所述第一特征图的维度以及各个维度的尺寸；对所述多个第一子特征图分别进行归一化处理，得到多个第二子特征图；对所述多个第二子特征图进行拼接，得到所述待处理图像的第二特征图。

在一种可能的实现方式中，根据所述第一特征图的维度信息及预设的拆分规则，将所述第一特征图拆分为多个第一子特征图，包括：根据所述第一特征图的空间维度的尺寸及预设的拆分规则，在空间维度上对所述第一特征图进行拆分，得到多个第一子特征图。

在一种可能的实现方式中，对所述多个第一子特征图分别进行归一化处理，得到多个第二子特征图，包括：在通道维度上对每个第一子特征图进行分组，分别对所述第一子特征图的各组通道进行归一化处理，得到所述第一子特征图的第二子特征图。

在一种可能的实现方式中，对所述多个第二子特征图进行拼接，得到所述待处理图像的第二特征图，包括：根据所述多个第一子特征图在所述第一特征图中的位置，对所述多个第二子特征图进行拼接，得到所述待处理图像的第二特征图。

在一种可能的实现方式中，所述拆分规则包括特征图的待拆分维度、各待拆分维度的拆分位置、各待拆分维度的拆分数量、各待拆分维度的拆分尺寸、拆分后的子特征图的数量中的至少一种。

在一种可能的实现方式中，所述方法还包括：对所述待处理图像的至少一个第二特征图进行尺度缩小及多尺度融合，得到编码后的多个特征图，所述编码后的多个特征图中各个特征图的尺度不同；对所述编码后的多个特征图进行尺度放大及多尺度融合，得

到所述待处理图像的分类预测结果。

在一种可能的实现方式中，对所述待处理图像的至少一个第二特征图进行尺度缩小及多尺度融合处理，得到编码后的多个特征图，包括：对m个第二特征图进行尺度缩小，得到尺度缩小后的m个特征图，m为正整数；对所述尺度缩小后的m个特征图进行特征融合，得到第m+1个特征图，所述尺度缩小后的m个特征图的尺度等于所述第m+1个特征图的尺度；对所述m个第二特征图及所述第m+1个特征图分别进行特征优化及融合，得到编码后的m+1个特征图。

在一种可能的实现方式中，对所述编码后的多个特征图进行尺度放大及多尺度融合处理，得到所述待处理图像的分类预测结果，包括：对编码后的m+1个特征图进行融合及尺度放大，得到尺度放大后的m个特征图，m为正整数；对所述尺度放大后的m个特征图进行特征优化及融合，得到所述待处理图像的分类预测结果。

在一种可能的实现方式中，所述方法通过神经网络实现，所述神经网络包括特征提取网络、编码网络及解码网络，所述特征提取网络用于对待处理图像进行特征提取，所述编码网络用于对所述待处理图像的至少一个第二特征图进行尺度缩小及多尺度融合，所述解码网络用于对所述编码后的多个特征图进行尺度放大及多尺度融合。

在一种可能的实现方式中，所述方法还包括：根据预设的训练集，训练所述神经网络，所述训练集中包括已标注的多个样本图像。

根据本公开的另一方面，提供了一种图像处理装置，包括：特征提取模块，用于对待处理图像进行特征提取，得到所述待处理图像的第一特征图；拆分模块，用于根据所述第一特征图的维度信息及预设的拆分规则，将所述第一特征图拆分为多个第一子特征图，所述第一特征图的维度信息包括所述第一特征图的维度以及各个维度的尺寸；归一化模块，用于对所述多个第一子特征图分别进行归一化处理，得到多个第二子特征图；拼接模块，用于对所述多个第二子特征图进行拼接，得到所述待处理图像的第二特征图。

在一种可能的实现方式中，所述拆分模块包括：拆分子模块，用于根据所述第一特征图的空间维度的尺寸及预设的拆分规则，在空间维度上对所述第一特征图进行拆分，得到多个第一子特征图。

在一种可能的实现方式中，所述归一化模块包括：归一化子模块，用于在通道维度上对每个第一子特征图进行分组，分别对所述第一子特征图的各组通道进行归一化处理，得到所述第一子特征图的第二子特征图。

在一种可能的实现方式中，所述拼接模块包括：拼接子模块，用于根据所述多个第一子特征图在所述第一特征图中的位置，对所述多个第二子特征图进行拼接，得到所述待处理图像的第二特征图。

在一种可能的实现方式中，所述拆分规则包括特征图的待拆分维度、各待拆分维度的拆分位置、各待拆分维度的拆分数量、各待拆分维度的拆分尺寸、拆分后的子特征图的数量中的至少一种。

在一种可能的实现方式中，所述装置还包括：编码模块，用于对所述待处理图像的至少一个第二特征图进行尺度缩小及多尺度融合，得到编码后的多个特征图，所述编码后的多个特征图中各个特征图的尺度不同；解码模块，用于对所述编码后的多个特征图进行尺度放大及多尺度融合，得到所述待处理图像的分类预测结果。

在一种可能的实现方式中，所述编码模块包括：缩小子模块，用于对m个第二特征图进行尺度缩小，得到尺度缩小后的m个特征图，m为正整数；第一融合子模块，用于对所述尺度缩小后的m个特征图进行特征融合，得到第m+1个特征图，所述尺度缩小后的m个特征图的尺度等于所述第m+1个特征图的尺度；第二融合子模块，用于对所述m个第二特征图及所述第m+1个特征图分别进行特征优化及融合，得到编码后的m+1个特征图。

在一种可能的实现方式中，所述解码模块包括：放大子模块，用于对编码后的m+1个特征图进行融合及尺度放大，得到尺度放大后的m个特征图，m为正整数；第三融合子模块，用于对所述尺度放大后的m个特征图进行特征优化及融合，得到所述待处理图像的分类预测结果。

在一种可能的实现方式中，所述装置通过神经网络实现，所述神经网络包括特征提取网络、编码网络及解码网络，所述特征提取网络用于对待处理图像进行特征提取，所述编码网络用于对所述待处理图像的至少一个第二特征图进行尺度缩小及多尺度融合，所述解码网络用于对所述编码后的多个特征图进行尺度放大及多尺度融合。

在一种可能的实现方式中，所述装置还包括：训练模块，用于根据预设的训练集，训练所述神经网络，所述训练集中包括已标注的多个样本图像。

根据本公开的另一方面，提供了一种电子设备，包括：处理器；用于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。

根据本公开的另一方面，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上述方法。

根据本公开的另一方面，提供了一种计算机程序，所述计算机程序包括计算机可读代码，当所述计算机可读代码在电子设备中运行时，所述电子设备中的处理器执行上述方法。

在本公开实施例中，能够对特征图进行拆分并分别进行归一化处理，得到归一化后的多个子特征图，并将归一化后的多个子特征图拼接为完整的特征图，从而保留局部特征信息，减小完整特征图归一化时的统计误差，提高所提取特征的有效性。

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。根据下面参考附图对示例性实施例的详细说明，本公开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的技术方案。

图1示出根据本公开实施例的图像处理方法的流程图。

图2示出根据本公开实施例的图像处理方法的处理过程的示意图。

图3a、图3b及图3c示出根据本公开实施例的图像处理方法的多尺度融合过程的示意图。

图4示出根据本公开实施例的图像处理装置的框图。

图5示出根据本公开实施例的一种电子设备的框图。

图6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。

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，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，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，以便于凸显本公开的主旨。

图1示出根据本公开实施例的图像处理方法的流程图，如图1所示，所述图像处理方法包括：

在步骤S11中，对待处理图像进行特征提取，得到所述待处理图像的第一特征图；

在步骤S12中，根据所述第一特征图的维度信息及预设的拆分规则，将所述第一特征图拆分为多个第一子特征图，所述第一特征图的维度信息包括所述第一特征图的维度以及各个维度的尺寸；

在步骤S13中，对所述多个第一子特征图分别进行归一化处理，得到多个第二子特征图；

在步骤S14中，对所述多个第二子特征图进行拼接，得到所述待处理图像的第二特征图。

在一种可能的实现方式中，所述图像处理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。

在一种可能的实现方式中，待处理图像可以是图像采集设备（例如摄像头）拍摄的监控区域（例如路口、商场等区域）的图像，也可以是通过其他方式获取的图像（例如网络下载的图像）。待处理图像中可包括一定数量的目标（例如行人、车辆、顾客等）。本公开对待处理图像的类型、获取方式以及图像中目标的类型不作限制。

在一种可能的实现方式中，可在步骤S11中通过特征提取网络对待处理图像进行特征提取，得到待处理图像的第一特征图。该特征提取网络可例如包括卷积神经网络，本公开对特征提取网络的具体网络类型不作限制。

在一种可能的实现方式中，第一特征图可具有维度信息，该维度信息包括第一特征图的维度以及各个维度的尺寸。例如第一特征图包括高度H、宽度W以及通道C这三个维度，其中高度H和宽度W为空间维度，通道C为通道维度。各个维度的尺寸例如高度H和宽度W均为256，通道C为16（即16个通道），则可表示为H×W×C=256×256×16。本公开对第一特征图的维度数量以及各个维度的具体尺寸均不作限制。

在一种可能的实现方式中，如果待处理图像中各个区域之间存在差异，例如各个区域的目标（行人）数量不同，则对第一特征图直接进行归一化处理可能会导致统计误差。在该情况下，可以将第一特征图进行拆分，以便分区域进行归一化处理以降低误差。

在一种可能的实现方式中，可预先设定有针对特征图的拆分规则，以便根据该拆分规则拆分特征图，例如均匀拆分为多块、指定拆分为特定尺寸的块等。可对所有的特征图采用同样的拆分规则，也可对不同的特征图采用不同的拆分规则，本公开对此不作限制。

在一种可能的实现方式中，拆分规则可包括特征图的待拆分维度、各待拆分维度的拆分位置、各待拆分维度的拆分数量、各待拆分维度的拆分尺寸、拆分后的子特征图的数量中的至少一种。

在一种可能的实现方式中，待拆分维度用于指示特征图要进行拆分的维度，例如对空间维度中的高度H和宽度W进行拆分；各待拆分维度的拆分位置用于指示对特征图的各个待拆分维度进行拆分的拆分点的位置，例如高度H（尺寸为256）的拆分位置包括52、108及160等；各待拆分维度的拆分数量用于指示对特征图的各个待拆分维度进行拆分的数量，例如在高度H（尺寸为256）维度方向上均匀拆分为三块；各待拆分维度的拆分尺寸用于指示对特征图的各个待拆分维度进行拆分的尺寸，例如在高度H（尺寸为256）维度方向上拆分的尺寸为64；拆分后的子特征图的数量用于指示对特征图拆分后得到的子特征图的数量，例如子特征图的数量为9，则可在高度H和宽度W维度方向上分别拆分为3块。

应当理解，本领域技术人员可根据实际情况设定具体的拆分规则及其内容，本公开对此不作限制。

在一种可能的实现方式中，可在步骤S12中根据第一特征图的维度信息及预设的拆分规则，将第一特征图拆分为多个第一子特征图。例如拆分规则指示待拆分维度为高度H和宽度W，在高度H和宽度W维度方向上分别拆分为2块，则可将第一特征图拆分为4个第一子特征图（128×128×16）。

在一种可能的实现方式中，可在步骤S13中对多个第一子特征图分别进行归一化处理，得到多个第二子特征图。本公开对归一化的具体方式不作限制。

在一种可能的实现方式中，可在步骤S14中对多个第二子特征图进行拼接，得到待处理图像的第二特征图。也即，根据归一化后的各个子特征图的位置，对多个子特征图进行拼接，得到完整的特征图（第二特征图），以便进行后续的处理。

根据本公开的实施例，能够对特征图进行拆分并分别进行归一化处理，得到归一化后的多个子特征图，并将归一化后的多个子特征图拼接为完整的特征图，从而保留局部特征信息，减小完整特征图归一化时的统计误差，提高所提取特征的有效性。

在一种可能的实现方式中，步骤S12可包括：

根据所述第一特征图的空间维度的尺寸及预设的拆分规则，在空间维度上对所述第一特征图进行拆分，得到多个第一子特征图。

举例来说，预设的拆分规则可设定为在空间维度（H和W）上对第一特征图进行拆分，例如在各维度方向上分别均匀拆分为4块。在该情况下，根据第一特征图的空间维度（H和W）的尺寸（256×256），可在空间维度（H和W）上将第一特征图拆分为尺寸为64×64的16个第一子特征图。各个第一子特征图的其它维度与第一特征图的其它维度的尺度相同（例如通道数量同为16）。

通过这种方式，可在空间维度上对特征图进行区域拆分，将特征图拆分为各个空间区域的子特征图，以便对特征图的各个空间区域分别进行归一化处理，从而减小完整特征图归一化时的统计误差。

在一种可能的实现方式中，步骤S13可包括：

在通道维度上对每个第一子特征图进行分组，分别对所述第一子特征图的各组通道进行归一化处理，得到所述第一子特征图的第二子特征图。

举例来说，在相关技术中，可通过批归一化（Batch Normalization，BN）对每一批的数据进行归一化。然而在训练时无法使用大的批尺寸（batch size）的任务上，例如物体检测、语义分割、人群密度估计，批归一化的效果较差。在该情况下，可采用组归一化（Group Normalization，GN）的方式对特征图进行归一化处理。

在一种可能的实现方式中，在得到多个子特征图后，可在通道（channel）维度C上对每个第一子特征图进行分组，例如将第一子特征图的16个通道分为2组，每组包括8个通道。然后在各组通道上分别进行归一化处理。也即，分别统计第一子特征图的各组通道的平均值和方差，再对第一子特征图的各组通道的每一个位置的值进行归一化计算，得到归一化后的结果（第一子特征图的第二子特征图）。这样，可得到与多个第一子特征图对应的多个第二子特征图。本公开对通道的分组数量及每组通道所包括的通道数量不作限制。

通过这种方式，能够对子特征图的各组通道分别进行归一化，从而进一步减归一化时的统计误差，提高所提取特征的有效性。

在一种可能的实现方式中，步骤S14可包括：

根据所述多个第一子特征图在所述第一特征图中的位置，对所述多个第二子特征图进行拼接，得到所述待处理图像的第二特征图。

举例来说，在得到归一化后的多个第二子特征图后，可将各个第一子特征图在第一特征图中的位置，确定为对应的各个第二子特征图的位置，也即，使得拼接的顺序与拆分时相同。根据各个第二子特征图的位置对各个第二子特征图进行拼接，得到拼接后的第二特征图。

通过这种方式，可保证第二特征图与第一特征图的局部特征信息的分布保持一致。

图2示出根据本公开实施例的图像处理方法的处理过程的示意图。如图2所示，第一

特征图21可包括高度H、宽度W以及通道C三个维度；预设的拆分规则可为在空间维度H上拆分为2块，在空间维度W上拆分为2块；根据第一特征图21的维度信息以及该拆分规则，可将第一特征图21拆分为4个第一子特征图22；可对4个第一子特征图22分别进行组归一化处理（通道C分组），并将得到的结果（4个第二子特征图）按拆分的顺序进行拼接，得到完整的第二特征图23，以便进行后续操作。

在一种可能的实现方式中，所述方法还可包括：

对所述待处理图像的至少一个第二特征图进行尺度缩小及多尺度融合，得到编码后的多个特征图，所述编码后的多个特征图中各个特征图的尺度不同；

对所述编码后的多个特征图进行尺度放大及多尺度融合，得到所述待处理图像的分类预测结果。

举例来说，可对待处理图像进行特征提取得到不同尺度的多个第一特征图，经步骤S12-S14的拆分、归一化及拼接处理后，可得到多个第二特征图。在后续处理中，可通过神经网络的编码网络对待处理图像的至少一个第二特征图进行尺度缩小及多尺度融合，得到编码后的多个特征图，编码后的多个特征图中各个特征图的尺度不同。这样，可在每个尺度上将全局和局部的信息进行融合，提高所提取的特征的有效性。

在一种可能的实现方式中，编码网络可例如包括卷积层、残差层、上采样层、融合层等。可通过编码网络的第一卷积层（步长>1）对第二特征图进行尺度缩小，得到尺度缩小后的特征图；通过第二卷积层（步长=1）和/或残差层对第二特征图及尺度缩小后的特征图进行特征优化，得到特征优化后的多个特征图；再通过编码网络的上采样层、卷积层（步长>1）和/或融合层等对特征优化后的多个特征图进行融合，得到编码后的多个特征图。

在一种可能的实现方式中，在得到编码后的多个特征图后，可通过解码网络对编码后的多个特征图进行尺度放大及多尺度融合处理，得到待处理图像的分类预测结果。

在一种可能的实现方式中，解码网络可例如包括融合层、反卷积层、卷积层、残差层、上采样层等。可通过解码网络的融合层对编码后的多个特征图进行融合，得到融合后的多个特征图；再通过反卷积层对融合后的多个特征图进行尺度放大，得到尺度放大后的多个特征图；通过融合层、卷积层（步长=1）和/或残差层等分别对多个特征图进行融合及优化，得到解码后的特征图（分类预测结果）。本公开对编码网络及解码网络的具体网络结构不作限制。

在一种可能的实现方式中，可在编码网络及解码网络的任意网络层（融合层、反卷积层、卷积层、残差层、上采样层等）之后进行步骤S12-S14的拆分、归一化及拼接处理，以便对各网络层的操作结果进行归一化，提高网络层的操作结果的鲁棒性。

通过这种方式，能够通过编码网络对图像的特征图进行尺度缩小及多尺度融合，并通过解码网络对编码后的多个特征图进行尺度放大及多尺度融合，从而在编码及解码过程中多次融合多尺度的全局信息和局部信息，保留了更有效的多尺度信息，提高了预测结果的质量及鲁棒性。

在一种可能的实现方式中，对所述待处理图像的至少一个第二特征图进行尺度缩小

及多尺度融合处理，得到编码后的多个特征图的步骤可包括：

对m个第二特征图进行尺度缩小，得到尺度缩小后的m个特征图，m为正整数；

对所述尺度缩小后的m个特征图进行特征融合，得到第m+1个特征图，所述尺度缩小后的m个特征图的尺度等于所述第m+1个特征图的尺度；

对所述m个第二特征图及所述第m+1个特征图分别进行特征优化及融合，得到编码后的m+1个特征图。

举例来说，可设定待处理的第二特征图为m个，m为任意正整数。可通过编码网络的m个卷积子网络（每个卷积子网络包括至少一个第一卷积层）对m个第二特征图分别进行尺度缩小，得到尺度缩小后的m个特征图，该尺度缩小后的m个特征图的尺度相同且尺度小于第m个第二特征图的尺度（等于第m+1个特征图的尺度）；通过融合层对该尺度缩小后的m个特征图进行特征融合，得到第m+1个特征图。

在一种可能的实现方式中，每个卷积子网络包括至少一个第一卷积层，第一卷积层的卷积核尺寸为3×3，步长为2，用于对特征图进行尺度缩小。卷积子网络的第一卷积层数量与对应的特征图的尺度相关联，例如，编码后 的第一个第二特征图的尺度为4x（宽和高分别为待处理图像的1/4），而待生成的m个特征图的尺度为16x（宽和高分别为待处理图像的1/16），则第一个卷积子网络包括两个第一卷积层。应当理解，本领域技术人员可根据实际情况设定卷积子网络第一卷积层的数量、卷积核尺寸及步长，本公开对此不作限制。

在一种可能的实现方式中，可通过编码网络的融合层对编码的m个第二特征图进行多尺度融合，得到融合后的m个特征图；通过m+1个特征优化子网络（每个特征优化子网络包括第二卷积层和/或残差层）分别对融合后的m个特征图和第m+1个特征图进行特征优化，得到特征优化后的m+1个特征图；然后通过m+1个融合子网络分别对特征优化后的m+1个特征图进行多尺度融合，得到编码后的m+1个特征图。

在一种可能的实现方式中，可以对多尺度融合后的m+1个特征图再次进行特征优化及多尺度融合，以便进一步提高所提取的多尺度特征的有效性。本公开对特征优化及多尺度融合的次数不作限制。

在一种可能的实现方式中，可直接通过第二卷积层对特征图进行优化，第二卷积层的卷积核尺寸为3×3，步长为1；也可通过由第二卷积层及残差层组成基本块（basic block）对特征图进行优化。该基本块可作为优化的基本单元，每个基本块可包括两个连续的第二卷积层，然后通过残差层将输入的特征图与卷积得到的特征图相加作为结果输出。本公开对特征优化的具体方式不作限制。

在一种可能的实现方式中，每个特征优化子网络可包括至少一个基本块。可通过各个特征优化子网络的基本块分别对m个第二特征图和第m+1个特征图进行特征优化，得到特征优化后的m+1个特征图。应当理解，本领域技术人员可根据实际情况设定第二卷积层的数量及卷积核尺寸，本公开对此不作限制。

通过这种方式，可进一步提高提取的多尺度特征的有效性。

在一种可能的实现方式中，编码网络的m+1个融合子网络可分别对特征优化后的m+1个特征图分别进行融合。对于m+1个融合子网络的第k个融合子网络（k为整数且1≤k≤

m+1），改第k个融合子网络首先可将m+1个特征图的尺度调整为特征优化后的第k个特征图的尺度。在1<k<m+1的情况下，在特征优化后的第k个特征图之前的k-1个特征图的尺度均大于特征优化后的第k个特征图，例如第k个特征图的尺度为16x（宽和高分别为待处理图像的1/16），第k个特征图之前的特征图的尺度为4x和8x。在该情况下，可通过至少一个第一卷积层对尺度大于特征优化后的第k个特征图的k-1个特征图进行尺度缩小，得到尺度缩小后的k-1个特征图。也即，将尺度为4x和8x的特征图均缩小为16x的特征图，可通过两个第一卷积层对4x的特征图进行尺度缩小，可通过一个第一卷积层对8x的特征图进行尺度缩小。这样，可以得到尺度缩小后的k-1个特征图。

在一种可能的实现方式中，在1<k<m+1的情况下，在特征优化后的第k个特征图之后的m+1-k个特征图的尺度均小于特征优化后的第k个特征图，例如第k个特征图的尺度为16x（宽和高分别为待处理图像的1/16），第k个特征图之后的m+1-k个特征图为32x。在该情况下，可通过上采样层对32x的特征图进行尺度放大，并通过第三卷积层（卷积核尺寸为1×1）对尺度放大后的特征图进行通道调整，使得尺度放大后的特征图的通道数与第k个特征图的通道数相同，从而得到尺度为16x的特征图。这样，可以得到尺度放大后的m+1-k个特征图。

在一种可能的实现方式中，在k=1的情况下，特征优化后的第1个特征图之后的m个特征图的尺度均小于特征优化后的第1个特征图，则可对后m个特征图均进行尺度放大及通道调整，得到尺度放大后的后m个特征图；在k=m+1的情况下，特征优化后的第m+1个特征图之前的m个特征图的尺度均大于特征优化后的第m+1个特征图，则可对前m个特征图均进行尺度缩小，得到尺度缩小后的前m个特征图。

在一种可能的实现方式中，第k个融合子网络可对尺度调整后的m+1个特征图进行融合。在1<k<m+1的情况下，尺度调整后的m+1个特征图包括尺度缩小后的k-1个特征图、特征优化后的第k个特征图及所述尺度放大后的m+1-k个特征图，可以对尺度缩小后的k-1个特征图、特征优化后的第k个特征图及所述尺度放大后的m+1-k个特征图这三者进行融合（相加），得到编码后的第k个特征图。

在一种可能的实现方式中，在k=1的情况下，尺度调整后的m+1个特征图包括特征优化后的第1个特征图和尺度放大后的m个特征图，可对特征优化后的第1个特征图和尺度放大后的m个特征图这两者进行融合（相加），得到编码后的第1个特征图。

在一种可能的实现方式中，在k=m+1的情况下，尺度调整后的m+1个特征图包括尺度缩小后的m个特征图和特征优化后的第m+1个特征图，可对尺度缩小后的m个特征图和特征优化后的第m+1个特征图这两者进行融合（相加），得到编码后的第m+1个特征图。

图3a、图3b及图3c示出根据本公开实施例的图像处理方法的多尺度融合过程的示意图。在图3a、图3b及图3c中，以待融合的特征图为三个为例进行说明。

如图3a所示，在k=1的情况下，可对第2个和第3个特征图分别进行尺度放大（上采样）及通道调整（1×1卷积），得到与第1个特征图的尺度及通道数相同的两个特征图，再将这三个特征图相加得到融合后的特征图。

如图3b所示，在k=2的情况下，可对第1个特征图进行尺度缩小（卷积核尺寸为3×3，步长为2的卷积）；对第3个特征图进行尺度放大（上采样）及通道调整（1×1卷积），从而得到与第2个特征图的尺度及通道数相同的两个特征图，再将这三个特征图相加得到融

合后的特征图。

如图3c所示，在k=3的情况下，可对第1个和第2个特征图进行尺度缩小（卷积核尺寸为3×3，步长为2的卷积）。由于第1个特征图与第3个特征图之间的尺度差为4倍，因此可进行两次卷积（卷积核尺寸为3×3，步长为2）。经尺度缩小后，可得到与第3个特征图的尺度及通道数相同的两个特征图，再将这三个特征图相加得到融合后的特征图。

通过这种方式，可以实现尺度不同的多个特征图之间的多尺度融合，在每个尺度上将全局和局部的信息进行融合，提取更加有效的多尺度特征。

在一种可能的实现方式中，对所述编码后的多个特征图进行尺度放大及多尺度融合处理，得到所述待处理图像的分类预测结果的步骤可包括：

对编码后的m+1个特征图进行融合及尺度放大，得到尺度放大后的m个特征图，m为正整数；

对所述尺度放大后的m个特征图进行特征优化及融合，得到所述待处理图像的分类预测结果。

举例来说，可先对编码后的m+1个特征图进行融合，在融合多尺度信息的同时减小特征图的数量。可设置有m个第一融合子网络，该m个第一融合子网络与编码后的m+1个特征图中的前m个特征图相对应。例如待融合的特征图包括尺度为4x、8x、16x及32x的四个特征图，则可设置有三个第一融合子网络，以便融合得到尺度为4x、8x及16x的三个特征图。

在一种可能的实现方式中，解码网络的m个第一融合子网络的网络结构可与编码网络的融合子网络的网络结构类似。例如，对于第q个第一融合子网络（1≤q≤m），第q个第一融合子网络可首先将m+1个特征图的尺度调整为解码后的第q个特征图的尺度，再对尺度调整后的m+1个特征图进行融合，得到融合后的第q个特征图。这样，可得到融合后的m个特征图。此处对尺度调整及融合的具体过程不再重复描述。

在一种可能的实现方式中，可通过解码网络的反卷积子网络对融合后的m个特征图分别进行尺度放大，例如将尺度为4x、8x及16x的三个融合后的特征图放大为2x、4x及8x的三个特征图。经放大后，得到尺度放大后的m个特征图。

在一种可能的实现方式中，在得到尺度放大后的m个特征图后，可通过m个第二融合子网络分别对该m个特征图进行尺度调整及融合，得到融合的m个特征图。此处对尺度调整及融合的具体过程不再重复描述。

在一种可能的实现方式中，可通过解码网络的特征优化子网络对融合的m个特征图分别进行优化，各个特征优化子网络均可包括至少一个基本块。经特征优化后，可得到解码的m个特征图。此处对特征优化的具体过程不再重复描述。

在一种可能的实现方式中，解码网络的多尺度融合及特征优化的过程可重复多次，以便进一步融合不同尺度的全局和局部特征。本公开对多尺度融合及特征优化的次数不作限制。

在一种可能的实现方式中，解码网络的融合及尺度放大的过程可重复多次，以便得到尺度与待处理图像一致的目标特征图；再对目标特征图进行优化，得到所述待处理图像的预测密度图。

在一种可能的实现方式中，可将该预测密度图直接作为待处理图像的预测结果；也可以对该预测密度图进行进一步的处理（例如通过softmax层等处理），得到待处理图像的分类预测结果。

通过这种方式，解码网络在尺度放大过程中多次融合全局信息和局部信息，提高了预测结果的质量。

在一种可能的实现方式中，根据本公开实施例的图像处理方法可通过神经网络实现，该神经网络包括特征提取网络、编码网络及解码网络，所述特征提取网络用于对待处理图像进行特征提取，所述编码网络用于对所述待处理图像的至少一个第二特征图进行尺度缩小及多尺度融合，所述解码网络用于对所述编码后的多个特征图进行尺度放大及多尺度融合。其中，特征提取网络、编码网络及解码网络的处理过程已在前文中说明，此处不再重复描述。

在一种可能的实现方式中，在应用本公开的神经网络之前，可对该神经网络进行训练。根据本公开实施例的图像处理方法还包括：

根据预设的训练集，训练所述神经网络，所述训练集中包括已标注的多个样本图像。

举例来说，可预先设置有多个样本图像，每个样本图像具有标注信息，例如样本图像中行人的位置、数量等信息。可将具有标注信息的多个样本图像组成训练集，训练所述神经网络。

在一种可能的实现方式中，可将样本图像输入特征提取网络，经由特征提取网络、编码网络及解码网络处理，输出样本图像的预测结果；根据样本图像的预测结果和标注信息，确定神经网络的网络损失；根据网络损失调整神经网络的网络参数；在满足预设的训练条件时，可得到训练后的神经网络。本公开对具体的训练方式不作限制。

通过这种方式，可得到高精度的神经网络。

根据本公开实施例的归一化方法，能够在空间维度上对特征图进行区域拆分，对各空间区域分别归一化，从而保留特征图局部的差异性，减小完整特征图归一化时的统计误差；根据本公开的实施例，在训练时使用小批尺寸也能保证网络的性能，能够应用于训练时只能使用小批尺寸的任务（例如人群密度估计、语义分割等），消除例如人群密度估计任务训练时不使用归一化层导致的梯度消失／爆炸等问题。

根据本公开实施例的图像处理方法，能够通过带步长的卷积操作来获取小尺度的特征图，在网络结构中不断进行全局和局部信息的融合来提取更有效的多尺度信息，并且通过其他尺度的信息来促进当前尺度信息的提取，增强网络对于多尺度目标（例如行人）识别的鲁棒性；能够在解码网络中放大特征图的同时进行多尺度信息的融合，保留多尺度信息，提高生成密度图的质量，从而提高模型预测的准确率。

根据本公开实施例的图像处理方法，能够应用于智能视频分析、安防监控等应用场景中，对场景中的目标（例如行人、车辆等）进行识别，预测场景中目标的数量、分布情况等，从而分析当前场景人群的行为。

可以理解，本公开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，本公开还提供了图像处理装置、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提供的任一种图像处理方法，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4示出根据本公开实施例的图像处理装置的框图，如图4所示，所述图像处理装置包括：

特征提取模块41，用于对待处理图像进行特征提取，得到所述待处理图像的第一特征图；拆分模块42，用于根据所述第一特征图的维度信息及预设的拆分规则，将所述第一特征图拆分为多个第一子特征图，所述第一特征图的维度信息包括所述第一特征图的维度以及各个维度的尺寸；归一化模块43，用于对所述多个第一子特征图分别进行归一化处理，得到多个第二子特征图；拼接模块44，用于对所述多个第二子特征图进行拼接，得到所述待处理图像的第二特征图。

在一种可能的实现方式中，所述拆分模块包括：拆分子模块，用于根据所述第一特征图的空间维度的尺寸及预设的拆分规则，在空间维度上对所述第一特征图进行拆分，得到多个第一子特征图。

在一种可能的实现方式中，所述归一化模块包括：归一化子模块，用于在通道维度上对每个第一子特征图进行分组，分别对所述第一子特征图的各组通道进行归一化处理，得到所述第一子特征图的第二子特征图。

在一种可能的实现方式中，所述拼接模块包括：拼接子模块，用于根据所述多个第一子特征图在所述第一特征图中的位置，对所述多个第二子特征图进行拼接，得到所述待处理图像的第二特征图。

在一种可能的实现方式中，所述拆分规则包括特征图的待拆分维度、各待拆分维度的拆分位置、各待拆分维度的拆分数量、各待拆分维度的拆分尺寸、拆分后的子特征图的数量中的至少一种。

在一种可能的实现方式中，所述装置还包括：编码模块，用于对所述待处理图像的至少一个第二特征图进行尺度缩小及多尺度融合，得到编码后的多个特征图，所述编码后的多个特征图中各个特征图的尺度不同；解码模块，用于对所述编码后的多个特征图进行尺度放大及多尺度融合，得到所述待处理图像的分类预测结果。

在一种可能的实现方式中，所述编码模块包括：缩小子模块，用于对m个第二特征图进行尺度缩小，得到尺度缩小后的m个特征图，m为正整数；第一融合子模块，用于对所述尺度缩小后的m个特征图进行特征融合，得到第m+1个特征图，所述尺度缩小后的m个特征图的尺度等于所述第m+1个特征图的尺度；第二融合子模块，用于对所述m个第二特征图及所述第m+1个特征图分别进行特征优化及融合，得到编码后的m+1个特征图。

在一种可能的实现方式中，所述解码模块包括：放大子模块，用于对编码后的m+1个特征图进行融合及尺度放大，得到尺度放大后的m个特征图，m为正整数；第三融合子模块，用于对所述尺度放大后的m个特征图进行特征优化及融合，得到所述待处理图像的分类预测结果。

在一种可能的实现方式中，所述装置通过神经网络实现，所述神经网络包括特征提取网络、编码网络及解码网络，所述特征提取网络用于对待处理图像进行特征提取，所述编码网络用于对所述待处理图像的至少一个第二特征图进行尺度缩小及多尺度融合，所述解码网络用于对所述编码后的多个特征图进行尺度放大及多尺度融合。

在一种可能的实现方式中，所述装置还包括：训练模块，用于根据预设的训练集，训练所述神经网络，所述训练集中包括已标注的多个样本图像。

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模块可以用于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。

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或易失性计算机可读存储介质。

本公开实施例还提出一种电子设备，包括：处理器；用于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。

本公开实施例还提出一种计算机程序，所述计算机程序包括计算机可读代码，当所述计算机可读代码在电子设备中运行时，所述电子设备中的处理器执行上述方法。

电子设备可以被提供为终端、服务器或其它形态的设备。

图5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。

参照图5，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CMOS或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WiFi，2G或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。

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6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6，电子设备1900包括处理组件1922，其进一步包括一个或多个处理器，以及由存储器1932所代表的存储器资源，用于存储可由处理组件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）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Windows ServerTM，Mac OS XTM，UnixTM, LinuxTM，FreeBSDTM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执行以完成上述方法。

本公开可以是系统、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的各个方面的计算机可读程序指令。

计算机可读存储介质可以是可以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。

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。也要注意的是，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。

在不违背逻辑的情况下，本公开不同实施例之间可以相互结合，不同实施例描述有所侧重，为侧重描述的部分可以参见其他实施例的记载。

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。