



(12) 发明专利

(10) 授权公告号 CN 112767539 B

(45) 授权公告日 2023. 08. 08

(21) 申请号 202110035802.6
(22) 申请日 2021.01.12
(65) 同一申请的已公布的文献号
申请公布号 CN 112767539 A

(43) 申请公布日 2021.05.07

(73) 专利权人 杭州师范大学
地址 311121 浙江省杭州市余杭区仓前街
道海曙路58号

(72) 发明人 何鑫睿 李秀梅 孙军梅

(74) 专利代理机构 杭州君度专利代理事务所
(特殊普通合伙) 33240
专利代理师 朱亚冠

(51) Int. Cl.
G06T 17/00 (2006.01)
G06N 3/045 (2023.01)

G06N 3/0464 (2023.01)
G06N 3/048 (2023.01)
G06N 3/047 (2023.01)
G06N 3/084 (2023.01)
G06N 3/082 (2023.01)

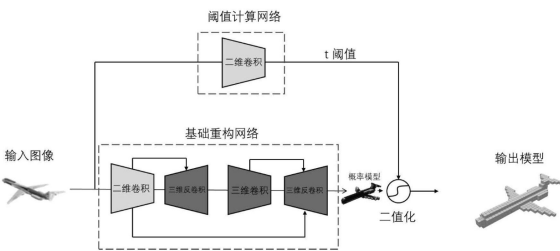
(56) 对比文件
CN 101930624 A, 2010.12.29
CN 108830157 A, 2018.11.16
CN 110544297 A, 2019.12.06
US 6049793 A, 2000.04.11
刘杨圣彦;潘翔;刘复昌;张三元.面向三维
模型视图特征提取的残差卷积网络优化.计算机
辅助设计与图形学学报.2019, (第06期), 第70-
76页.

审查员 马晓晓

权利要求书3页 说明书8页 附图2页

(54) 发明名称
一种基于深度学习的图像三维重构方法及
系统

(57) 摘要
本发明公开一种基于深度学习的图像三维
重构方法及系统。本发明将利用Pix2vox网络作
为基本重构网络,在此基础上添加多通道连接并
融合通道注意力机制,获取多尺度图像特征。同
时结合阈值计算网络构成完整重构,网络阈值计
算网络可以为不同类别目标提供相适应二值化
阈值。并且采用融合型损失函数,弥补重构数据
不平衡及类间差异性。使用所述方法有效提高了
单图像多类别三维模型重构的准确率。



1. 一种基于深度学习网络的图像三维重构方法,其特征在于该方法包括以下步骤:

步骤一、获取二维图像与对应三维模型数据,构建数据集;

步骤二、对数据集中的二维图像数据与三维模型数据进行预处理;

步骤三、神经网络搭建和训练;

3-1. 搭建完整重构网络

完整重构网络为卷积神经网络,包括基础重构网络、阈值计算网络;其中基础重构网络为改进Pix2vox-A网络,用于根据单张二维图像输出三维概率模型,三维概率模型结合阈值计算网络获得的二值化阈值进行二值化处理,输出重构三维模型;所述阈值计算网络,用于识别基础重构网络的输入图像类别,根据图像类别获得对应三维模型体素的二值化阈值;

3-2. 模型训练

将步骤二预处理后图像数据输入到步骤3-1搭建的完整重构网络中进行训练;训练将通过融合型损失函数计算中完整重构网络输出的重构三维模型与目标模型的损失值;通过损失值反向传播,网络迭代参数不断减小损失值以完成训练;

步骤四、根据步骤三所搭建并训练的网络模型,输入待重构的单张二维图像,最终输出重构三维模型。

2. 根据权利要求1所述的一种基于深度学习网络的图像三维重构方法,其特征在于步骤3-1中,基础重构网络包括生成器,优化器;生成器用于获得初步重建模型,优化器用于优化重建模型,生成器与优化器之间增加两条中浅尺度特征传输通道连接,以保留全局特征信息。

3. 根据权利要求2所述的一种基于深度学习网络的图像三维重构方法,其特征在于所述生成器包括特征提取模块、深层连接模块、中层连接模块、浅层连接模块、第一三维反卷积模块、第一注意力机制层、第二三维反卷积模块、第二注意力机制层、第三三维反卷积模块;

所述特征提取模块由现有网络的VGG16前8个串联的二维卷积单元构成;第一个二维卷积单元作为所述特征提取模块的输入端,接预处理后二维图像数据;第二个二维卷积单元的输出端接浅层连接模块的输入端;第五个二维卷积单元的输出端接中层连接模块的输入端;第八个二维卷积单元的输出端接深层连接模块的输入端;

所述深层连接模块的输出端接第一三维反卷积模块的输入端;浅层连接模块的第一输出端接第二注意力机制层的第一输入端,第二输出端接优化器的第一输入端构成浅尺度特征传输通道;中层连接模块的第一输出端接第一注意力机制层的第一输入端,第二输出端接优化器的第二输入端构成中尺度特征传输通道;

所述第一三维反卷积模块由两个串联的第一、二三维反卷积单元构成;第一三维反卷积单元的输入端作为第一三维反卷积模块的输入端,第二三维反卷积单元的输出端接第一注意力机制层的第二输入端;

第二三维反卷积模块由三个串联的第三至五三维反卷积单元构成;第三三维反卷积单元的输入端作为第二三维反卷积模块的输入端,接第一注意力机制层的输出端;第五三维反卷积单元的输出端接第二注意力机制层的第二输入端;

第三三维反卷积模块由两个串联的第六、七三维反卷积单元构成;第六三维反卷积单元的输入端作为第三三维反卷积模块的输入端,接第二注意力机制层的输出端;第七三维

反卷积单元的输出端输出初步模型数据。

4. 根据权利要求3所述的一种基于深度学习网络的图像三维重构方法,其特征之处在于所述优化器包括第一三维卷积模块,全连接层模块,第四三维反卷积模块;其中第一三维卷积模块由三个串联的第一至三三维卷积单元构成;全连接层模块由两个串联的第一、二全连接单元构成;第四三维反卷积模块由三个串联的第八至十三维卷积单元构成;

第一三维卷积单元的输入端作为优化器网络的第三输入端,接生成器网络输出的初步模型数据;第一三维卷积单元的第一输出端接第二三维卷积单元的输入端,第二输出端接第十三维反卷积单元的第二输入端;第二三维卷积单元的第一输出端接第三三维卷积单元的第一输入端,第二输出端接第九三维反卷积单元的第二输入端;第三三维卷积单元的第一输出端接第一全连接单元的输入端,第二输出端接第八三维反卷积单元的第二输入端;第一全连接单元的输入端接第二全连接层的输入端;第二全连接层的输出端接第八三维反卷积单元的第一输入端;第八三维反卷积单元输出端接第九三维反卷积单元第一输入端;第九三维反卷积单元输出端接第十三维反卷积单元的第一输入端,第三输入端作为优化器网络的第二输入端,接生成器的中层连接模块第二输出端;第十三维反卷积单元的第三输入端作为优化器网络的第一输入端,接生成器的浅层连接模块第二输出端,输出端输出数据与第七三维反卷积单元输出端输出数据相加作为优化后的三维模型。

5. 根据权利要求3或4所述的一种基于深度学习网络的图像三维重构方法,其特征之处在于所述深层连接模块、中层连接模块、浅层连模块结构相同,均由第一至三二维卷积单元,及一个重排列层依次串联构成;第一、三二维卷积单元由二维卷积层、批量归一化层、ELU激活函数层依次串联构成;第二二维卷积单元由二维卷积层、批量归一化层、ELU激活函数层、最大池化层依次串联构成。

6. 根据权利要求4所述的一种基于深度学习网络的图像三维重构方法,其特征之处在于所述第一至六三维反卷积单元由三维反卷积层、批量归一化层、Relu激活函数层依次串联构成;所述第七三维反卷积单元由三维反卷积层、批量归一化层、Sigmoid激活函数层依次串联构成;所述第八至九三维反卷积单元由三维反卷积层、批量归一化层、Relu激活函数层依次串联构成;所述第十三维反卷积单元由三维反卷积层、批量归一化层、Sigmoid激活函数层依次串联构成;

所述第一至三三维卷积单元由三维卷积层、批量归一化层、LeakyRelu激活函数层、最大池化层依次串联构成;

所述第一至二全连接单元由全连接层、Relu激活函数层依次串联构成。

7. 根据权利要求3所述的一种基于深度学习网络的图像三维重构方法,其特征之处在于所述注意力机制为通道注意力机制。

8. 根据权利要求2所述的一种基于深度学习网络的图像三维重构方法,其特征之处在于步骤3-1中,所述阈值计算网络包括VGG13网络、全连接层、Softmax激活函数层;VGG13网络的输入端接基础重构网络输入图像,输出端接全连接层的输入端;全连接层的输出端接Softmax激活函数层的输入端;Softmax激活函数层的输出端输出类别概率向量;最后根据最大概率的类别赋予二值化阈值。

9. 根据权利要求1所述的一种基于深度学习网络的图像三维重构方法,其特征之处在于步骤3-2模型训练中所述融合型损失函数见式(4),融合损失为内容损失D_loss与类别损失C_

loss相加构成;

$$D_loss = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (2)$$

其中X为完整重构网络输出的重构三维模型,Y为目标模型;

$$C_loss = Bce(D(X), y) \quad (3)$$

其中D为模型类别预测网络,输出结果为预测类别,y为真实类别,Bce为二元交叉熵函数;

$$Total_loss = D_loss + C_loss \quad (4)。$$

10.一种实现权利要求1-9任一项所述方法的基于深度学习的图像三维重构系统,其特征在于包括:

数据获取模块,用于获取二维图像与对应三维模型,构建训练集;

数据预处理模块,用于对二维图像与对应三维模型调整格式与大小;

完整重构网络模块,包括基础重构网络、阈值计算网络;基础维重构网络用于根据单张二维图像输出三维概率模型,三维概率模型结合阈值计算网络给出的相应二值化阈值进行二值化处理,输出重构三维模型;阈值计算网络,用于识别基础维重构网络输入图像类别,根据图像类别获得对应三维模型体素的二值化阈值;

基础重构网络包括生成器、优化器,生成器用于获得初步重建模型,优化器用于优化重建模型,生成器与优化器之间增加两条中浅尺度特征传输通道连接。

一种基于深度学习的图像三维重构方法及系统

技术领域

[0001] 本发明涉及深度学习及图像三维重构领域,具体涉及一种基于深度学习的图像三维重构方法及系统。

背景技术

[0002] 如今,游戏动漫、三维打印、工业设计等领域对三维模型的需求日益增加,而手动建模耗时耗力,通过三维扫描设备建模成本较高,因此基于图像进行三维重构由于其自动化、低成本的优势成为计算机视觉领域一大研究热点。

[0003] 图像的三维重建是从有限信息的二维图像中逆向恢复原模型,重构对象具有非常不同的表现形式,从纯粹的关系信息、稀疏或密集度量的深度信息到单个对象甚至场景。所以一种有效、兼容、稳定的图像三维重构方法设计,已成为研究者们共同的研究目标。

[0004] 传统图像重构方法可以分为基于模型的重建方法与基于几何外形恢复的重建方法。基于模型的重建方法,如形变模型,通过预设参数模型不断优化参数拟合目标模型,完成重构,所依赖的光照模型存在参数较多、模型复杂、优化计算过程耗时等问题。基于几何外形恢复的重建方法主要通过物理几何计算图像的光影、纹理、运动等求解三维模型,光影模型需要对投影物体提出约束条件,运动模型重构要求具备多视图图像,难以通过学习从单个视图的原始像素重建可变形对象。

[0005] 近年来,随着卷积神经网络的广泛应用,图像重构技术方法也从传统的物理几何计算,发展到深度学习的应用,从多图像恢复发展到单图像重构。深度学习方法,利用计算机建立神经网络模型,模拟人类视觉感知学习三维空间的过程。神经网络模型可以学习重构模型的先验分布,通过迭代训练建立图像与对应三维模型的映射关系。使用深度学习方法进行图像的三维重构,可利用端到端网络计算有效地降低计算复杂度,基于单图像完成图像重构任务。然而现有深度学习方法仍然存在由于不同类型三维模型类间差异,数据的不平衡性,导致重构精度不高的问题。

发明内容

[0006] 本发明的目的是在于提出一种基于深度学习的图像三维重构方法及系统,弥补类间差异与不平衡数据,提升图像数据利用效率,最终提高基于深度学习的图像三维重构精度。

[0007] 一种基于深度学习的图像三维重构方法,具体如下:

[0008] 步骤一、获取二维图像与对应三维模型数据。数据来源为Shapenet公开数据集。

[0009] 步骤二、对数据集中的图像数据与三维模型数据进行预处理。预处理方法包括二维图像缩放、裁剪、归一化,三维模型体素化。

[0010] 步骤三、神经网络搭建和训练。

[0011] 3-1.搭建完整重构网络

[0012] 完整重构网络为卷积神经网络,包括基础重构网络、阈值计算网络。所述的基础维

重构网络为主干网络,该网络为改进Pix2vox-A(Pixel to Voxel)网络,用于输入图像输出三维概率模型,概率模型结合阈值计算网络给出的二值化阈值进行二值化处理,输出重构三维模型。所述阈值计算网络,用于识别基础重构网络输入图像类别,根据图像类别获得对应三维模型体素的二值化阈值。本发明提出使用阈值计算网络自适应输出二值化阈值方法替代传统设定固定二值化阈值的方法,以提升重构网络对输入不同类别图像的适应性。

[0013] 3-2. 模型训练

[0014] 将所重构的图像数据输入到步骤3-1记载的深度学习网络中进行训练。训练将通过融合型损失函数计算中网络输出的重构模型与目标模型的损失值。通过损失值反向传播,网络迭代参数不断减小损失值以完成训练。本发明改进在于使用所述融合型损失函数替换二元交叉熵损失函数,使得损失值计算更加全面。

[0015] 步骤四、根据步骤三所搭建并训练的网络模型,输入单张待重构二维图像,最终输出重构三维模型。

[0016] 作为优选,步骤3-1中,基础重构网络网络结构包括模型的生成器、优化器;生成器用于获得初步重建模型,优化器用于优化重建模型,生成器与优化器之间主要由一条模型数据传输通道连接。本发明所述网络改进为生成器与优化器之间增加两条中浅尺度特征传输通道连接,以保留全局特征信息;优化器中增加中浅尺度连接并添加连接模块,以保留多尺度图像信息;优化器中增加注意力机制,以自适应不同尺度通道的特征学习。

[0017] 所述生成器网络包括特征提取模块、深层连接模块、中层连接模块、浅层连接模块、第一三维反卷积模块、第一注意力机制层、第二三维反卷积模块、第二注意力机制层、第三三维反卷积模块;

[0018] 所述特征提取模块由现有网络的VGG16前8个串联的二维卷积单元构成;第一个二维卷积单元作为所述特征提取模块的输入端,接预处理后二维图像数据;第二个二维卷积单元的输出端接浅层连接模块的输入端;第五个二维卷积单元的输出端接中层连接模块的输入端;第八个二维卷积单元的输出端接深层连接模块的输入端;

[0019] 所述深层连接模块的输出端接第一三维反卷积模块的输入端;浅层连接模块的第一输出端接第二注意力机制层的第一输入端,第二输出端接优化器的第一输入端构成浅尺度特征传输通道;中层连接模块的第一输出端接第一注意力机制层的第一输入端,第二输出端接优化器的第二输入端构成中尺度特征传输通道;

[0020] 所述第一三维反卷积模块由两个串联的第一、二三维反卷积单元构成;第一三维反卷积单元的输入端作为第一三维反卷积模块的输入端,第二三维反卷积单元的输出端接第一注意力机制层的第二输入端;

[0021] 第二三维反卷积模块由三个串联的第三至五三维反卷积单元构成;第三三维反卷积单元的输入端作为第二三维反卷积模块的输入端,接第一注意力机制层的输出端;第五三维反卷积单元的输出端接第二注意力机制层的第二输入端;

[0022] 第三三维反卷积模块由两个串联的第六、七三维反卷积单元构成;第六三维反卷积单元的输入端作为第三三维反卷积模块的输入端,接第二注意力机制层的输出端;第七三维反卷积单元的输出端输出初步模型数据;

[0023] 所述优化器网络包括第一三维卷积模块,全连接层模块,第四三维反卷积模块;其中第一三维卷积模块由三个串联的第一至三三维卷积单元构成;全连接层模块由两个串联

的第一、二全连接单元构成；第四三维反卷积模块由三个串联的第八至十三维卷积单元构成；

[0024] 第一三维卷积单元的输入端作为优化器网络的第三输入端，接生成器网络输出的初步模型数据；第一三维卷积单元的第一输出端接第二三维卷积单元的输入端，第二输出端接第十三维反卷积单元的第二输入端；第二三维卷积单元的第一输出端接第三三维卷积单元的第一输入端，第二输出端接第九三维反卷积单元的第二输入端；第三三维卷积单元的第一输出端接第一全连接单元的输入端，第二输出端接第八三维反卷积单元的第二输入端；第一全连接单元的输入端接第二全连接层的输入端；第二全连接层的输出端接第八三维反卷积单元的第一输入端；第八三维反卷积单元输出端接第九三维反卷积单元第一输入端；第九三维反卷积单元输出端接第十三维反卷积单元的第一输入端，第三输入端作为优化器网络的第二输入端，接生成器的中层连接模块第二输出端；第十三维反卷积单元的第三输入端作为优化器网络的第一输入端，接生成器的浅层连接模块第二输出端，输出端输出数据与第七三维反卷积单元输出端输出数据相加作为优化后的三维模型；

[0025] 作为优选，所述深层连接模块、中层连接模块、浅层连模块结构相同，均由第一至三二维卷积单元，及一个重排列层依次串联构成；第一、三二维卷积单元由二维卷积层、批量归一化层、ELU激活函数层依次串联构成；第二二维卷积单元由二维卷积层、批量归一化层、ELU激活函数层、最大池化层依次串联构成；

[0026] 作为优选，所述第一至六三维反卷积单元由三维反卷积层、批量归一化层、Relu激活函数层依次串联构成；所述第七三维反卷积单元由三维反卷积层、批量归一化层、Sigmoid激活函数层依次串联构成；所述第八至九三维反卷积单元由三维反卷积层、批量归一化层、Relu激活函数层依次串联构成；所述第十三维反卷积单元由三维反卷积层、批量归一化层、Sigmoid激活函数层依次串联构成；

[0027] 作为优选，所述注意力机制为通道注意力机制；

[0028] 作为优选，所述第一至三三维卷积单元由三维卷积层、批量归一化层、LeakyRelu激活函数层、最大池化层依次串联构成；

[0029] 作为优选，所述第一至二全连接单元由全连接层、Relu激活函数层依次串联构成；

[0030] 所述基础重构网络工作流程如下：所述基础重构网络中生成器网络输入224*224大小图像，输出32*32*32大小的三维矩阵，矩阵代表32³尺寸模型空间，每个矩阵元素值表示32³空间内单位体积属于目标模型体素的概率，取值范围为0至1；然后将所得三维矩阵输入优化器网络，输出优化后的32*32*32三维矩阵；最后将三维矩阵每个元素经阈值分配网络所得的二值化阈值进行如式(1)二值化处理，输出重构三维模型；

[0031]
$$\text{Model} = \text{voxel}_{(i,j,k)} > t \quad (1)$$

[0032] 式中voxel表示三维矩阵元素表示的体素概率值，(i,j,k)表示32³空间内单位体素坐标，t表示二值化阈值，Model为图像三维重构网络最终输出的三维模型；

[0033] 作为优选，步骤3-1中，所述阈值计算网络包括VGG13网络、一个全连接层、一个Softmax激活函数层。其中VGG13由十个二维卷积单元与三个全连接层组成，用于图像类别特征提取。全连接层用于特征降维。Sigmoid激活函数层用于特征概率化。VGG13网络的输入端接基础重构网络输入图像，输出端接全连接层的输入端；全连接层的输出端接Softmax激活函数层的输入端；Softmax激活函数层的输出端输出类别概率向量；最后该网络根据最大

概率的类别赋予特定二值化阈值。

[0034] 作为优选,步骤3-2模型训练中融合损失函数见式(4),融合损失为内容损失D_loss与类别损失C_loss相加构成。

$$[0035] \quad D_loss = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (2)$$

[0036] 其中X定义为重构模型,Y定义为目标模型,基于二者相交区域与相并区域比值,含义为生成模型与目标模型间相似程度,损失值越小即生成模型越接近于目标模型。

$$[0037] \quad C_loss = Bce(D(X), y) \quad (3)$$

[0038] 其中D为模型类别预测网络,输出结果为预测类别,y为真实类别,Bce为二元交叉熵函数。

[0039] 所述模型类别预测网络结构由四个三维卷积单元,三个全连接单元依次串联实现;四个三维卷积单元均由三维卷积层、批量归一化层、LeakyRelu激活函数层依次串联构成;三个全连接单元中前两个均由全连接层、Relu激活函数层依次串联构成,最后一个由全连接层、Softmax激活函数层串联依次构成;其工作流程是获取重构网络所输出的模型,通过四个三维卷积单元进行下采样并展开采样特征,然后将三维模型的采样特征输入全连接单元,全连接单元输出模型类别的概率向量,从而预测输入模型类别。

$$[0040] \quad Total_loss = D_loss + C_loss \quad (4)$$

[0041] 本发明的另一个目的是提供一种基于深度学习的图像三维重构系统,系统包括:

[0042] 数据获取模块,用于获取二维图像与对应三维模型,构建训练集;

[0043] 数据预处理模块,用于对二维图像与对应三维模型调整格式与大小;

[0044] 完整重构网络模块,包括基础重构网络、阈值计算网络;基础维重构网络用于根据单张二维图像输出三维概率模型,三维概率模型结合阈值计算网络给出的相应二值化阈值进行二值化处理,输出重构三维模型;阈值计算网络,用于识别基础维重构网络输入图像类别,根据图像类别获得对应三维模型体素的二值化阈值;本发明提出使用阈值计算网络自适应输出二值化阈值方法替代传统设定固定二值化阈值的方法,以提升重构网络对输入不同类别图像的适应性。

[0045] 基础重构网络包括生成器、优化器,生成器用于获得初步重建模型,优化器用于优化重建模型,生成器与优化器之间增加两条中浅尺度特征传输通道连接。

[0046] 本发明的有益效果如下:

[0047] 1、本发明面向单图像多类别目标重构任务,引进图像识别模型作为阈值计算网络,基于类别的阈值分配相比于固定阈值设定更加具有类别适应性。

[0048] 2、本发明针对不平衡数据问题,采用基于Dice损失的内容损失加上类别损失,构建融合型损失函数,更好的对抗不平衡数据提升重构精度,其中类别损失由类别损失计算网络输出。

[0049] 3、本发明网络增加生成器网络中浅层网络连接并融合注意力机制,以获取多尺度图像特征,强化重要通道特征学习,并将生成器网络多尺度图像特征同时输入优化器网络,使二维特征与优化器原先获得的三维特征共同优化生成模型,利用充分图像信息,获得更好的重构效果。

附图说明

- [0050] 图1是本发明中完整重构网络的模型结构图。
[0051] 图2是本发明中基础重构网络的模型结构图。
[0052] 图3是本发明中阈值计算网络的模型结构图。
[0053] 图4是本发明中模型类别预测网络的模型结构图

具体实施方式

- [0054] 下面结合具体实施例对本发明做进一步分析。
- [0055] 一种基于深度学习的图像三维重构系统,包括以下内容:
- [0056] 数据获取模块,用于获取二维图像与对应三维模型,构建训练集;
- [0057] 数据预处理模块,用于对二维图像与对应三维模型调整格式与大小;
- [0058] 完整重构网络模块,包括基础重构网络、阈值计算网络;基础维重构网络用于根据单张二维图像输出三维概率模型,三维概率模型结合阈值计算网络给出的相应二值化阈值进行二值化处理,输出重构三维模型;阈值计算网络,用于识别基础维重构网络输入图像类别,根据图像类别获得对应三维模型体素的二值化阈值;本发明提出使用阈值计算网络自适应输出二值化阈值方法替代传统设定固定二值化阈值的方法,以提升重构网络对输入不同类别图像的适应性。
- [0059] 基础重构网络包括生成器、优化器,生成器用于获得初步重建模型,优化器用于优化重建模型,生成器与优化器之间增加两条中浅尺度特征传输通道连接。
- [0060] 一种基于深度学习的图像三维重构方法,包括以下内容:
- [0061] 步骤一、获取二维图像与对应三维模型数据。数据来源为Shapene公开数据集。
- [0062] 步骤二、对数据集中的图像数据与三维模型数据进行预处理。预处理方法包括二维图像缩放、裁剪、归一化,三维模型体素化。
- [0063] 步骤三、神经网络搭建和训练。
- [0064] 3-1.搭建完整重构网络
- [0065] 如图1完整重构网络为卷积神经网络,包括基础重构网络、阈值计算网络。所述的基础维重构网络为主干网络,该网络为改进Pix2vox-A(Pixel to Voxel)网络,用于输入图像输出三维概率模型,概率模型结合阈值计算网络给出的二值化阈值进行二值化处理,输出重构三维模型。所述阈值计算网络,用于识别基础重构网络输入图像类别,根据图像类别获得对应三维模型体素的二值化阈值。本发明提出使用阈值计算网络自适应输出二值化阈值方法替代传统设定固定二值化阈值的方法,以提升重构网络对输入不同类别图像的适应性。
- [0066] 3-2.模型训练
- [0067] 将所重构的图像数据输入到步骤3-1记载的深度学习网络中进行训练。训练将通过融合型损失函数计算中网络输出的重构模型与目标模型的损失值。通过损失值反向传播,网络迭代参数不断减小损失值以完成训练。本发明改进在于使用所述融合型损失函数替换二元交叉熵损失函数,使得损失值计算更加全面。
- [0068] 步骤四、根据步骤三所搭建并训练的网络模型,输入单张待重构二维图像,最终输出重构三维模型。

[0069] 步骤3-1中,基础重构网络网络结构如图2包括模型的生成器、优化器;生成器用于获得初步重建模型,优化器用于优化重建模型,生成器与优化器之间主要由一条模型数据传输通道连接。本发明所述网络改进为生成器与优化器之间增加两条中浅尺度特征传输通道连接,以保留全局特征信息;优化器中增加中浅尺度连接并添加连接模块,以保留多尺度图像信息;优化器中增加注意力机制,以自适应不同尺度通道的特征学习。

[0070] 所述生成器网络包括特征提取模块、深层连接模块、中层连接模块、浅层连接模块、第一三维反卷积模块、第一注意力机制层、第二三维反卷积模块、第二注意力机制层、第三三维反卷积模块;

[0071] 所述特征提取模块由现有网络的VGG16前8个串联的二维卷积单元构成;第一个二维卷积单元作为所述特征提取模块的输入端,接预处理后二维图像数据;第二个二维卷积单元的输出端接浅层连接模块的输入端;第五个二维卷积单元的输出端接中层连接模块的输入端;第八个二维卷积单元的输出端接深层连接模块的输入端;

[0072] 所述深层连接模块的输出端接第一三维反卷积模块的输入端;浅层连接模块的第一输出端接第二注意力机制层的第一输入端,第二输出端接优化器的第一输入端构成浅尺度特征传输通道;中层连接模块的第一输出端接第一注意力机制层的第一输入端,第二输出端接优化器的第二输入端构成中尺度特征传输通道;

[0073] 所述第一三维反卷积模块由两个串联的第一、二三维反卷积单元构成;第一三维反卷积单元的输入端作为第一三维反卷积模块的输入端,第二三维反卷积单元的输出端接第一注意力机制层的第二输入端;

[0074] 第二三维反卷积模块由三个串联的第三至五三维反卷积单元构成;第三三维反卷积单元的输入端作为第二三维反卷积模块的输入端,接第一注意力机制层的输出端;第五三维反卷积单元的输出端接第二注意力机制层的第二输入端;

[0075] 第三三维反卷积模块由两个串联的第六、七三维反卷积单元构成;第六三维反卷积单元的输入端作为第三三维反卷积模块的输入端,接第二注意力机制层的输出端;第七三维反卷积单元的输出端输出初步模型数据;

[0076] 所述优化器网络包括第一三维卷积模块,全连接层模块,第四三维反卷积模块;其中第一三维卷积模块由三个串联的第一至三三维卷积单元构成;全连接层模块由两个串联的第一、二全连接单元构成;第四三维反卷积模块由三个串联的第八至十三维卷积单元构成;

[0077] 第一三维卷积单元的输入端作为优化器网络的第三输入端,接生成器网络输出的初步模型数据;第一三维卷积单元的第一输出端接第二三维卷积单元的输入端,第二输出端接第十三维反卷积单元的第二输入端;第二三维卷积单元的第一输出端接第三三维卷积单元的第一输入端,第二输出端接第九三维反卷积单元的第二输入端;第三三维卷积单元的第一输出端接第一全连接单元的输入端,第二输出端接第八三维反卷积单元的第二输入端;第一全连接单元的输入端接第二全连接层的输入端;第二全连接层的输出端接第八三维反卷积单元的第一输入端;第八三维反卷积单元输出端接第九三维反卷积单元第一输入端;第九三维反卷积单元输出端接第十三维反卷积单元的第一输入端,第三输入端作为优化器网络的第二输入端,接生成器的中层连接模块第二输出端;第十三维反卷积单元的第三输入端作为优化器网络的第一输入端,接生成器的浅层连接模块第二输出端,输出端输

出数据与第七三维反卷积单元输出端输出数据相加作为优化后的三维模型；

[0078] 作为优选，所述深层连接模块、中层连接模块、浅层连接模块结构相同，均由第一至三二维卷积单元，及一个重排列层依次串联构成；第一、三二维卷积单元由二维卷积层、批量归一化层、ELU激活函数层依次串联构成；第二二维卷积单元由二维卷积层、批量归一化层、ELU激活函数层、最大池化层依次串联构成；

[0079] 作为优选，所述第一至六三维反卷积单元由三维反卷积层、批量归一化层、Relu激活函数层依次串联构成；所述第七三维反卷积单元由三维反卷积层、批量归一化层、Sigmoid激活函数层依次串联构成；所述第八至九三维反卷积单元由三维反卷积层、批量归一化层、Relu激活函数层依次串联构成；所述第十三三维反卷积单元由三维反卷积层、批量归一化层、Sigmoid激活函数层依次串联构成；

[0080] 作为优选，所述注意力机制为通道注意力机制；

[0081] 作为优选，所述第一至三三维卷积单元由三维卷积层、批量归一化层、LeakyRelu激活函数层、最大池化层依次串联构成；

[0082] 作为优选，所述第一至二全连接单元由全连接层、Relu激活函数层依次串联构成；

[0083] 所述基础重构网络工作流程如下：所述基础重构网络中生成器网络输入 224×224 大小图像，输出 $32 \times 32 \times 32$ 大小的三维矩阵，矩阵代表 32^3 尺寸模型空间，每个矩阵元素值表示 32^3 空间内单位体积属于目标模型体素的概率，取值范围为0至1；然后将所得三维矩阵输入优化器网络，输出优化后的 $32 \times 32 \times 32$ 三维矩阵；最后将三维矩阵每个元素经阈值分配网络所得的二值化阈值进行如式(1)二值化处理，输出重构三维模型；

[0084]
$$\text{Model} = \text{voxel}_{(i,j,k)} > t \quad (1)$$

[0085] 式中voxel表示三维矩阵元素表示的体素概率值，(i,j,k)表示 32^3 空间内单位体素坐标，t表示二值化阈值，Model为图像三维重构网络最终输出的三维模型；

[0086] 步骤3-1中，所述阈值计算网络如图3包括VGG13网络、一个全连接层、一个Softmax激活函数层。其中VGG13由十个二维卷积单元与三个全连接层组成，用于图像类别特征提取。全连接层用于特征降维。Sigmoid激活函数层用于特征概率化。VGG13网络的输入端接基础重构网络输入图像，输出端接全连接层的输入端；全连接层的输出端接Softmax激活函数层的输入端；Softmax激活函数层的输出端输出类别概率向量；最后该网络根据最大概率的类别赋予特定二值化阈值。

[0087] 作为优选，步骤3-2模型训练中融合损失函数见式(4)，融合损失为内容损失 D_{loss} 与类别损失 C_{loss} 相加构成。

[0088]
$$D_{\text{loss}} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (2)$$

[0089] 其中X定义为重构模型，Y定义为目标模型，基于二者相交区域与相并区域比值，含义为生成模型与目标模型间相似程度，损失值越小即生成模型越接近于目标模型。

[0090]
$$C_{\text{loss}} = \text{Bce}(D(X), y) \quad (3)$$

[0091] 其中D为模型类别预测网络，输出结果为预测类别，y为真实类别，Bce为二元交叉熵函数。

[0092] 如图4所述模型类别预测网络结构由四个三维卷积单元，三个全连接单元依次串联实现；四个三维卷积单元均由三维卷积层、批量归一化层、LeakyRelu激活函数层依次串

联构成;三个全连接单元中前两个均由全连接层、Relu激活函数层依次串联构成,最后一个由全连接层、Softmax激活函数层串联依次构成;其工作流程是获取重构网络所输出的模型,通过四个三维卷积单元进行下采样并展开采样特征,然后将三维模型的采样特征输入全连接单元,全连接单元输出模型类别的概率向量,从而预测输入模型类别。

[0093] $\text{Total_loss} = \text{D_loss} + \text{C_loss}$ (4)

[0094] 为了验证本发明的有效性,将其多种现有重构方法如3D-R2N2、Matryoshka、Pix2Vox-A等进行比较,实验基于Shapenet数据集选择13种不同类别三维模型进行训练和测试,以IoU(intersection over union)值作为评价指标,该指标表示如式(5)所示,其中定义X为重构模型,Y为目标模型,计算为二者相交区域除以相并区域,结果取值范围为0至1,用以评价网络重构模型与目标模型之间的拟合程度,IoU值越高表示重构效果越好。

[0095]
$$IoU = \frac{|X \cap Y|}{|X \cup Y|} \quad (5)$$

[0096] 实验结果如表1所示,本发明网络模型在多项类别上均达到出色的重构效果,全类别平均IoU为0.670,相比于Pix2Vox提升IoU,0.009并高于其他模型方法。

[0097] 表1本发明网络模型方法与现有方法重构结果比较

Category	Num	3D-R2N2	OGN[28]	OccNet[29]	Matryoshka[30]	IM-Net[31]	Pix2Vox-A	本发明
airplane	810	0.513	0.587	0.532	0.647	0.702	0.684	0.705
bench	364	0.421	0.481	0.597	0.577	0.564	0.616	0.622
cabinet	315	0.716	0.729	0.674	0.776	0.680	0.792	0.794
car	1501	0.798	0.828	0.671	0.850	0.756	0.854	0.863
chair	1357	0.466	0.483	0.583	0.547	0.644	0.567	0.574
display	220	0.468	0.502	0.651	0.532	0.585	0.537	0.537
lamp	465	0.371	0.398	0.474	0.408	0.433	0.443	0.450
speaker	325	0.662	0.637	0.655	0.701	0.683	0.714	0.714
rifle	475	0.544	0.593	0.656	0.616	0.723	0.615	0.641
sofa	635	0.628	0.646	0.669	0.681	0.694	0.709	0.715
table	1703	0.513	0.536	0.659	0.573	0.621	0.601	0.607
telephone	211	0.661	0.702	0.794	0.756	0.762	0.776	0.789
watercraft	389	0.513	0.632	0.579	0.591	0.607	0.594	0.604
Overall	8770	0.560	0.596	0.626	0.635	0.659	0.661	0.670

[0098]

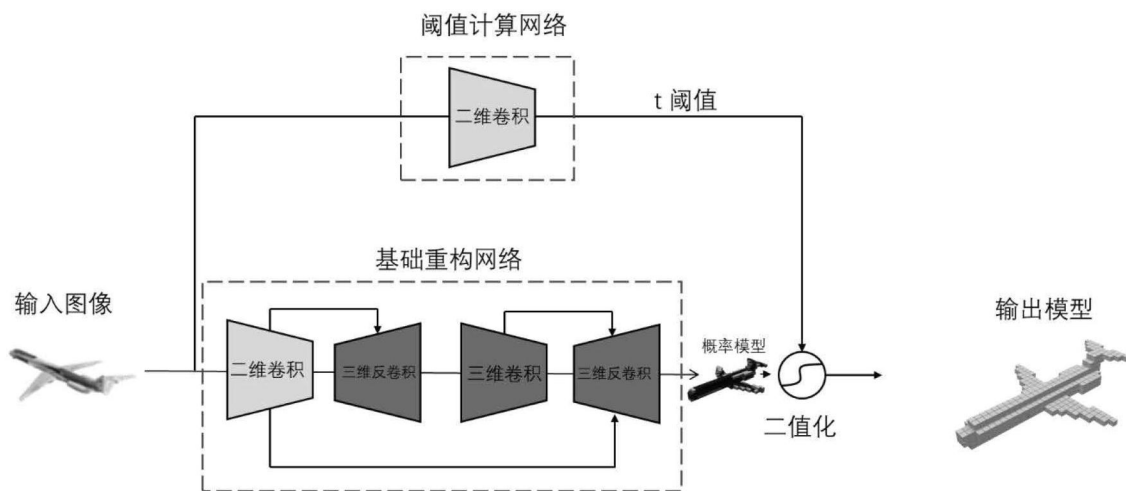


图1

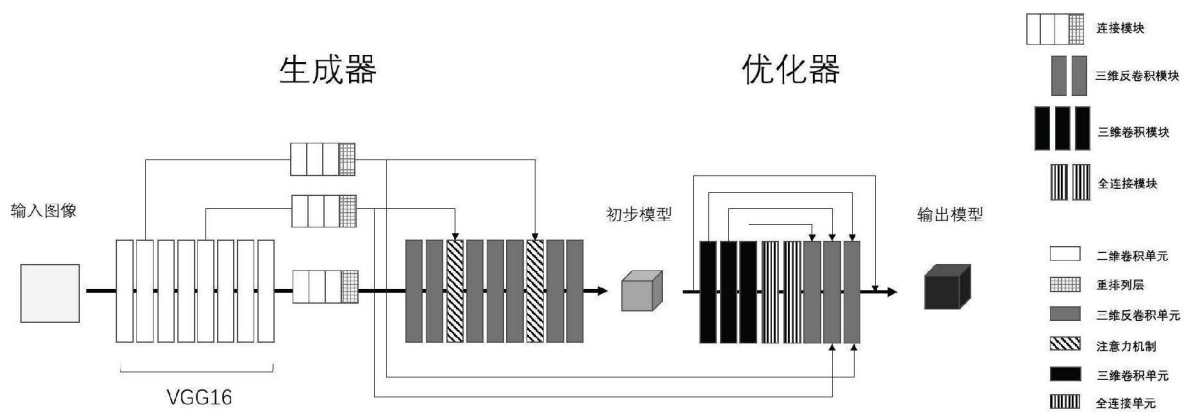


图2

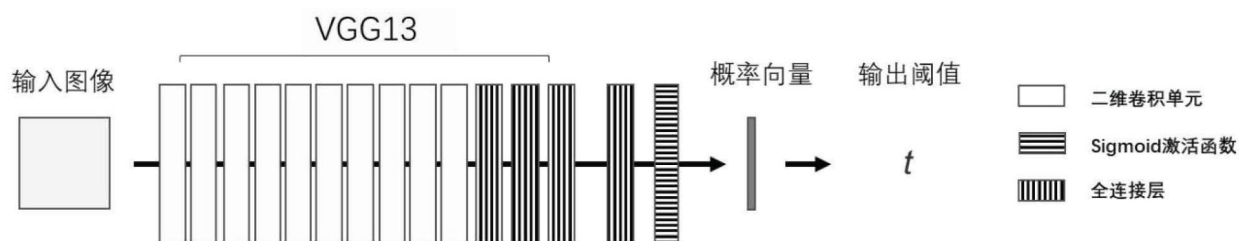


图3

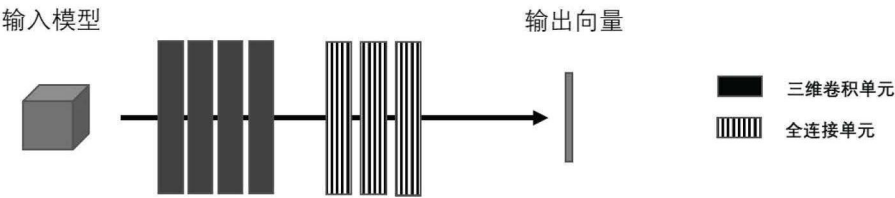


图4