文章编号　1004-924X(2024)10-1552-15

GCN引导模型视点的光学遥感道路提取网络

刘光辉[1,2*]，单　哲[1,2]，杨塬海[1,2]，王　恒[1,2]，孟月波[1,2]，徐胜军[1,2]
(1. 西安建筑科技大学 信息与控制工程学院，陕西 西安 710055；
2. 西安市建筑制造智动化技术重点实验室，陕西 西安 710055)

摘要：在光学遥感图像中，道路易受遮挡物、铺装材料以及周围环境等多重因素的影响，导致其特征模糊不清。然而，现有道路提取方法即使增强其特征感知能力，仍在特征模糊区域存在大量误判。为解决上述问题，本文提出基于GCN引导模型视点的道路提取网络(RGGVNet)。RGGVNet采用编解码结构，并设计基于GCN的视点引导模块(GVPG)在编解码器的连接处反复引导模型视点，从而增强对特征模糊区域的关注。GVPG利用GCN信息传播过程具有平均特征权重的特性，将特征图中不同区域道路显著性水平作为拉普拉斯矩阵，参与到GCN信息传播从而实现引导模型视点。同时，提出密集引导视点策略(DGVS)，采用密集连接的方式将编码器、GVPG和解码器相互连接，确保有效引导模型视点的同时缓解优化困难。在解码阶段设计多分辨率特征融合(MRFF)模块，最小化不同尺度道路特征在特征融合和上采样过程中的信息偏移和损失。在两个公开遥感道路数据集中，本文方法*IoU*分别达到65.84%和69.36%，*F1-score*分别达到79.40%和81.90%。从定量和定性两方面实验结果可以看出，本文所提方法性能优于其他主流方法。

关　键　词：光学遥感图像；道路提取；深度神经网络；图卷积网络
中图分类号：TP394.1；TH691.9　　**文献标识码**：A　　**doi**：10.37188/OPE.20243210.1552

Optical remote sensing road extraction network based on GCN guided model viewpoint

LIU Guanghui[1,2*], SHAN Zhe[1,2], YANG Yuanhai[1,2], WANG Heng[1,2],
MENG Yuebo[1,2], XU Shengjun[1,2]

(1. *College of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an* 710055, *China*;
2. *Xi'an Key Laboratory of Intelligent Technology for Building and Manufacturing, Xi'an* 710055, *China*)
* *Corresponding author, E-mail: guanghuil*@163.*com*

Abstract: In optical remote sensing images, roads are easily affected by multiple factors such as obstructions, pavement materials, and surrounding environments, resulting in blurred features. However, even if existing road extraction methods enhance their feature perception capabilities, they still suffer from a large number of misjudgments in feature-blurred areas. To address the above issues, this paper proposed the road extraction network based on GCN guided model viewpoint (RGGVNet). RGGVNet adopted the encoder-decoder structure and designed a GCN based viewpoint guidance module (GVPG) to repeatedly guide the model viewpoint at the connection of the encoder and decoder, thereby enhancing attention to fea-

收稿日期：2023-11-14；**修订日期**：2024-01-04.
基金项目：陕西省重点研发计划项目(No. 2021SF-429)；陕西省自然科学基础研究计划项目(No. 2023-JC-YB-532)

ture blurred areas. GVPG took advantage of the fact that the GCN information propagation process had the characteristic of average feature weight, used the road salience levels in different areas as a Laplacian matrix, and participated in GCN information propagation to realize the guidance model perspective. At the same time, a dense guidance viewpoint strategy (DGVS) was proposed, which uses dense connections to connect the encoder, GVPG module, and decoder to each other to ensure effective guidance of model viewpoints while alleviating optimization difficulties. In the decoding stage, a multi-resolution feature fusion module (MRFF) was designed to minimize the information offset and loss of road features of different scales in the feature fusion and upsampling process. In two public remote sensing road datasets, the *IoU* of our method reached 65.84% and 69.36%, respectively, and the *F1-score* reached 79.40% and 81.90%, respectively. It can be seen from the quantitative and qualitative experimental results that the performance of our method is superior to other mainstream methods.

Key words: optical remote sensing images; road extraction; deep neural network; graph convolution network

1　引　言

随着遥感技术的高速发展，海量的光学遥感图像被卫星摄像机或机载摄像机采集并存储。遥感道路提取算法利用这类图像为智能交通、自然灾害预警、数字城市、土地利用勘测等下游任务提供关键技术[1-2]。由于其在现实应用中的重要性，如何精准、完整、自动化地从高分辨遥感图像中提取出道路，已然成为学者们研究的热点。

遥感道路提取方法主要利用计算机视觉技术感知遥感图像中的特征，并判断每个像素是否为道路。通过分析过去近二十年提出的遥感道路算法，可根据是否使用深度神经网络将其分为两类：基于传统机器学习的提取方法和基于深度学习的提取方法。

在遥感道路提取方法的早期研究阶段，研究者们通常利用传统计算机视觉和机器学习方法挖掘图像中特定的特征来识别道路。这类方法大体可以被划分为：基于形态特征和基于手工特征。基于形态特征的道路提取方法依赖于道路在图像中特定的形态特征来识别道路，如亮度、颜色、纹理和几何形状等。当使用形态特征时，可能会引入优选的形状偏差。为了解决这一问题，Valero等[3]提出了一种基于方向数学形态学运算符的新方法用于道路提取，该运算符可以适应直线和稍微弯曲的结构，提取结果优于使用旋转矩形结构元素的标准方法。随后Chaudhuri等[4]提出了一种半自动化的道路检测方法，该方法利用了道路片段特性而开发定制的运算符可以更准确地提取道路片段。Grinias等[5]进一步探索了无监督的道路提取，通过结合马尔可夫随机场模型和随机森林方法对道路可能性进行聚类。上述方法可以有效地获取道路形状特征并实现道路提取，但面对遮挡、光线和对比度变化这类情况时鲁棒性较差。基于手工特征的方法允许研究人员根据具体情况选择和设计特定的特征，以更好地捕捉道路的多样性形状、颜色和纹理特征。针对遥感道路图像中存在的脊状或带状线性特征，Shao等[6]提出通过像素的灰度值和与线性特征的位置关系对像素进行区分，从而改善道路细节的处理并减少处理时间。Poullis等[7]同时捕获三种几何类型的信息：表面、曲线和交汇处，用于同时切割和分类多级遥感图像特征。Maboudi等[8]提出了一种多阶段的道路提取方法，该方法集成了遥感图像的结构、光谱、纹理以及上下文信息，并设计了一种新颖的线性指数用于区分细长路段与其他对象。基于传统机器学习的方法作为道路提取算法的基石，取得了显著的成功。然而，这些方法仍存在大量的局限性：首先，依赖于手工设计和选择的特征，在提取道路特征时需要领域专家的经验和知识。其次，对数据质量敏感，在特征模糊和背景复杂的遥感图像中基本无法正确提取道路。最后，模型泛化能力有限，通常仅能在单一场景提取。

得益于卷积神经网络(CNN)强大的特征提取能力，诸如FCN[9]，U-Net[10]，LinkNet[11]，Deep-

LabV3+[12]等经典的语义分割网络在遥感道路提取中取得了出色的性能。大量研究者在经典模型上改进以进一步适应遥感道路提取任务，其中特征模糊是影响道路提取精度的关键因素。特征模糊是指道路与周围特征高度相似或受如树木、车辆、建筑物阴影等其他物体遮挡，导致这些区域道路特征难以与其他特征区分。Buslaev等[13]基于残差的思想[14]改进原始U-Net编码部分，提升模型对道路特征与相似特征的区分能力。为解决道路特征被阴影和树木遮挡的问题，Xin等[15]结合密集连接[16]和U-net的优势，加强不同尺度道路特征信息的融合。Wu等[17]构建密集全局空间金字塔模块(DGSPP)，对上下文语义信息进行感知和聚合，以减轻阴影对道路特征的影响。Wang等[18]基于非局部(Non-local)的思想提出NL-LinkNet，将局部运算通过计算所有像素权重的方式转化为对道路图像全局和特征空间相关性建模，以改善被遮挡区域道路提取结果。Zhang等[19]融合道路局部和全局信息并设计门控细化单元，逐步将道路特征从复杂的背景中提取出。Jie等[20]提出远程上下文信息感知模块，通过获取特征被遮挡区域道路的远程上下文信息的方式推断被遮挡区域的道路。上述文献通过残差连接、多尺度信息融合、捕获上下文信息等方式提升了在遮挡区域算法的精度。近些年，注意力机制[21]和ViT[22]在计算机视觉领域大放异彩。因其优越的全局建模能力，引起了道路提取算法研究学者们的广泛关注。Wan等[23]引入双注意力机制，增强道路在全局空间中的定位，有效地推理被遮挡部分道路。Dai等[24]和Xu等[25]基于注意力机制分别设计可变形注意力模块和门控自注意力模块，以捕获特定形状和远程的道路特征信息。文献[26-28]从不同角度改进ViT基础模型，增强模型对全局空间中道路复杂特征和结构的感知。

综上所述，研究者们通过不同方法增强模型对道路模糊特征的辨别能力，极大提升了在特征模糊区域的提取精度。本文进一步分析发现遥感图像中道路特征往往同时存在显著和模糊区域。首先，受铺装材料的影响不同类型道路特征具有显著差异，例如农村非铺装道路和城市铺装道路在特征上完全不同，农村道路极易与周围环境混淆，其特征显著性明显低于城市铺装道路。其次，同一类型道路不同区域特征显著性差别较大，例如城市道路受建筑物阴影、行道树、车辆等遮挡尤为严重，同一路段被遮挡区域道路特征显著性较低。然而，现有基于CNN的道路提取算法对于这种情况，通常倾向于关注特征显著区域，即模型视点聚焦于特征明显的道路。模型视点即模型学习时的关注点。CNN利用卷积核扫描输入图像并通过监督标签更新卷积核权重。在扫描图像过程中，模型自然而然地被与监督标签相关的特征所吸引，导致其更为关注这些区域。由于卷积运算局部感知和参数共享的特性，视点区域特征对于模型学习过程的影响程度更大[29-31]。虽然模型以高效的方式学习图像特征，但视点过度集中容易导致对特征模糊区域的关注不足。因此，单从增强模型对道路特征感知能力的角度出发具有一定的局限性，在设计方法时更需要考虑如何平均对不同特征显著性水平道路的关注度。

基于此，本文提出基于GCN引导模型视点的道路提取网络(Road Extraction Network Based on GCN Guided Model Viewpoint, RGGVNet)，该网络采用编解码结构，并通过在编解码器连接处反复引导模型视点，平衡对不同特征显著性水平区域的关注。在编码部分采用ConvNeXt[32]作为主干网络，提取出初始多尺度道路特征。为有效引导模型视点，增强对特征模糊区域的关注，设计基于GCN的视点引导模块(GCN Based Viewpoint Guidance Module, GVPG)和密集引导视点策略(Dense Guidance Viewpoint Strategy, DGVS)，旨在通过GCN信息传播过程平均不同区域的特征权重。在解码阶段考虑到多尺度道路特征在融合和上采样过程中的信息偏移和损失，设计多分辨率特征融合模块(Multi-Resolution Feature Fusion Module, MRFF)实现精确完整地恢复特征分辨率，并输出提取结果。

2 基于GCN引导模型视点的道路提取网络

RGGVNet整体结构如图1所示。首先，将

原始遥感图像输入编码器(E)提取多分辨率道路特征 $\boldsymbol{X}^l$，$l\in(1,2,3,4)$。然后，将其分别输入GVPG模块并输出 $\boldsymbol{Y}^l$，内部连接如下侧小图所示。最后，将 $\boldsymbol{Y}^l$ 输入解码器(D)中，先通过MRFF中进行特征融合和上采样，而后分类器(C)输出提取结果。

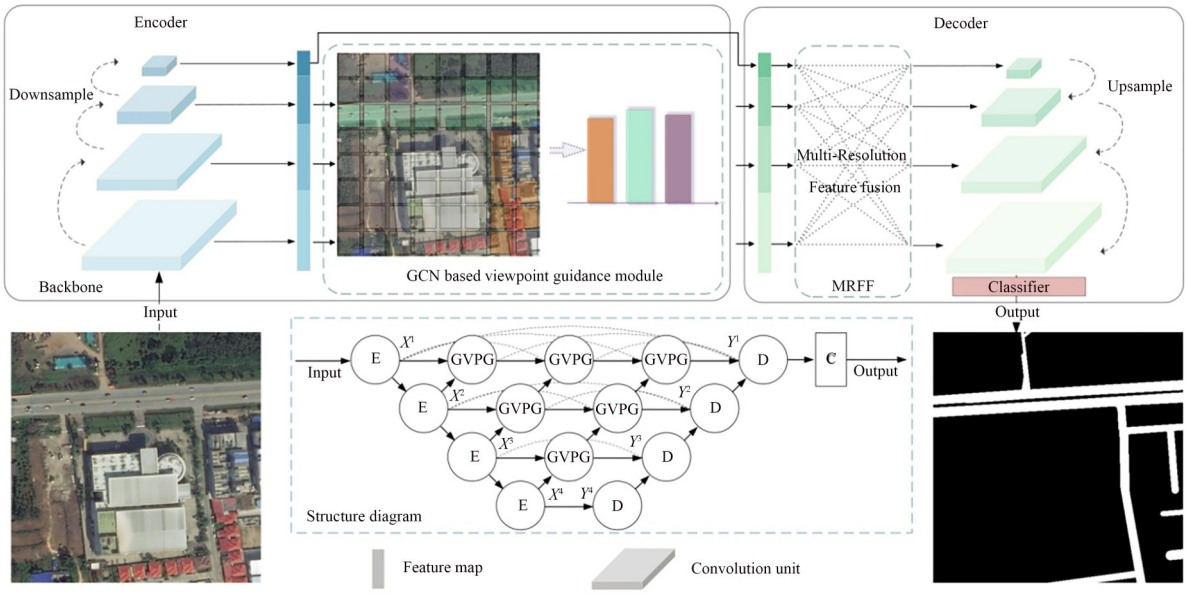

图1　RGGVNet网络结构图

Fig. 1　Overall structure of the RGGVNet

2.1　编码器

在语义分割模型中，编码器是提取图像特征的关键组件。在编码过程中能否提取到丰富的语义信息和精准的位置信息，直接影响分割的效果。ConvNeXt近期被大量计算机视觉算法使用提取图像初始特征并在下游任务中取得了优异的性能，例如图像分类[33]、目标检测[34]、语义分割[35]等。通过对网络结构、激活函数、卷积核等方面的改进，ConvNeXt具有强大的特征提取能力，能够有效捕获遥感图像中复杂的特征，因此本文使用其作为编码器主干网络。该网络具有四个特征提取阶段，每个阶段特征空间尺寸减半，而通道尺寸翻倍。一般认为，空间对于图像的位置信息，而通道对于图像的语义信息。与分类任务不同，道路提取中道路特征的位置信息和语义信息都非常重要。基于此，为丰富不同尺度道路信息，本文将其每一阶段提取出特征图都进行保留，并输入后续模块。

ConvNeXt具体特征提取过程如2图所示，首先将原始遥感图像 $\boldsymbol{I}\in\boldsymbol{R}^{h\times w\times 3}$ 输入预提取模块将图像分辨率直接下采样到四分之一，并将通道数升为 c(可调节的超参数，根据原论文选择96)。随后的四个阶段均通过堆叠不同层数的ConvNeXt块提取特征，每一阶段图像分辨率减半，通道数升至两倍。ConvNeXt块中对特征进行三次卷积，一次层归一化，以及一次GELU激活函数，其中每一阶段的最后一层卷积对原始特征进行下采样和升维。

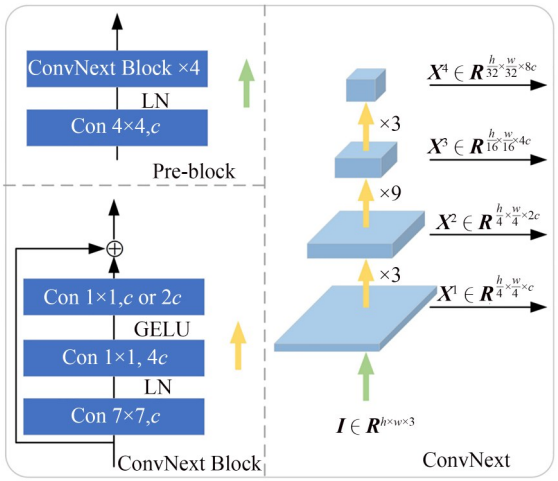

图2　ConvNeXt网络结构图

Fig. 2　Overall structure of the ConvNeXt

2.2　GCN引导模型视点

2.2.1　图卷积神经网络

对于给定的图 $G=(V,E)$ 及其邻接矩阵 A 和度矩阵 D，多层图卷积网络中逐层传播的规则如下所示[36]：

$$H^{(n+1)}=\sigma(LH^nW^n), \tag{1}$$

其中：H^n 表示第 n 层顶点特征，W^n 表示第 n 层可学习的权重矩阵，$L=I+D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 是对称归一化拉普拉斯矩阵代表了不同节点的相互关系，σ 是非线性激活函数。

GCN的传播过程本质上来说是对节点自身以及邻接节点的特征进行平均。GCN在其传播过程中代表节点相互关系的拉普拉斯矩阵参与信息的聚合和传递，可以更好地学习到节点之间的关系和图的全局特征。

2.2.2　基于GCN的视点引导模块

GCN的信息传播过程由于拉普拉斯矩阵的参加，具有平均不同节点特征权重的特性。因此，可以利用这一特性，对二维图像构建拉普拉斯矩阵，平均模型对不同区域关注，从而引导模型视点。实现这一方案的关键在于如何合理地对图像构建拉普拉斯矩阵，目前常用方法是将图像像素或区域视作为节点建立全连接的图[37][38]，并对其构建拉普拉斯矩阵。然而，这类方法将所有节点均与周围节点相互关联，并未考虑到节点之间是否具有相互关系。本文认为更为合理的是以道路特征显著性水平作为衡量节点相关性的依据并构建拉普拉斯矩阵，实现平等对待道路特征模糊和显著的区域。基于上述思想提出基于GCN的视点引导模块。

如图3右侧所示，基于GCN的视点引导模块首先对特征图划分区域并计算区域信息的相似度作为相似矩阵，该矩阵描述不同区域之间是否具有相似关系的同时表征了道路特征显著性水平。然后，将相似矩阵作为GCN中的拉普拉斯矩阵，实现以道路特征显著性水平衡量不同区域是否具有相似关系。最后，由于相似矩阵的参与，在图推理过程中引导模型对不同区域信息的感知和传播，完成对道路特征权重的重分配。

具体实现过程参见图3左侧。首先，采用ViT中提出的区域编码方法将特征 $X^l\in R^{H\times W\times C}$ 划分为 $\frac{HW}{hw}$ 个区域，每个区域包含 $h\times w\times c$ 个特征元素。编码后的特征为 $X_p\in R^{\left(\frac{HW}{hw}\times hwc\right)}$，其中 H,W,C 代表原始特征尺寸和通道数，h,w,c 代表编码后特征尺寸和通道数。

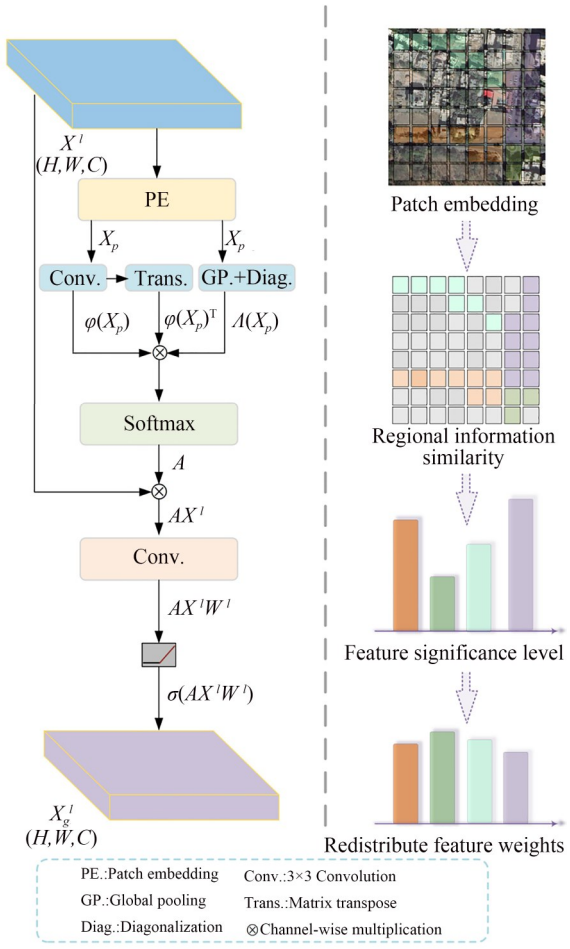

图3　GVPG模块示意图

Fig. 3　Illustration of the GVPG module

然后，对编码后的特征 X_p 采用卷积 $\phi(X_p)$ 提取出 $\frac{HW}{hw}$ 个区域的道路特征并进行矩阵转置 $\phi(X_p)^T$。同时，类似于GCN中的自连接操作，对 X_p 进行全局池化(GP.)和对角化(Diag.)，生成区域自连接矩阵 $\Lambda(X_p)\in R^{\frac{HW}{hw}\times\frac{HW}{hw}}$。如式(2)所示，相乘上述三个矩阵并通过 $Softmax$ 计算出每一区域特征信息与其余区域特征信息的相似程度，作为相似矩阵 A。相似矩阵一方面反映了不同区域之间道路深层特征的差异；另一方面，道路特征显著性较低区域与其他区域相似度较小，

该矩阵同时表征了道路特征显著性水平，实现了通过特征显著性水平衡量区域相互关系的目的。因此，将其作为GCN信息传播中的拉普拉斯矩阵是合理的。

$$\boldsymbol{A}=Softmax(\phi(\boldsymbol{X})_i\Lambda(\boldsymbol{X})\phi(\boldsymbol{X})_j^{\mathrm{T}}). \tag{2}$$

最后，按照多层图卷积网络的逐层传播规则，将模块学习过程表述为式(3)。拉普拉斯矩阵参与图推理过程引导模型对不同区域的道路特征信息进行聚合和传播，平均不同区域的特征权重。通过多轮迭代自适应地更新相似矩阵，使模型具有平等关注不同特征显著性水平区域的能力。

$$\boldsymbol{X}_g^l=\sigma(\boldsymbol{A}\boldsymbol{X}^l\boldsymbol{W}^l), \tag{3}$$

其中：$\boldsymbol{A}$是构建的拉普拉斯矩阵；$\boldsymbol{W}^l$是可学习的权重矩阵，等效于对特征进行一层卷积；σ是非线性激活函数，本文采用ReLU函数。

2.2.3　密集引导视点策略

遥感图像空间尺寸一般都比较大，而GVPG模块的引导范围具有一定限度，需要多次引导视点，即堆叠多层GVPG模块。然而，当GCN堆叠多层时容易造成过平滑现象[39-40]，造成优化困难。因此，本文基于密集连接的思想[16]提出密集引导视点策略。该策略可以具体化为公式(4)，即对每个GVPG模块的输出都与后续模块的输入相连接，形成一种密集的连接结构，具体操作如图4所示。

$$\boldsymbol{Y}=f_1(\boldsymbol{X}),f_2([\boldsymbol{X},f_1(\boldsymbol{X})]),f_3([\boldsymbol{X},f_1(\boldsymbol{X}),f_2([\boldsymbol{X},f_1(\boldsymbol{X})])]),\cdots. \tag{4}$$

密集引导视点策略引入模型中主要具有以下三方面优势：首先，由于直接的特征复用缓解了神经网络中梯度消失的问题，避免了在多次利用GCN进行信息传播时所造成的过平滑现象。其次，后续的模块可以直接访问早期特征，增强了特征的重用性，提升了网络性能。最后，在训练过程中该操作不仅不会带来额外的计算开销，同时加速了网络的优化速度。

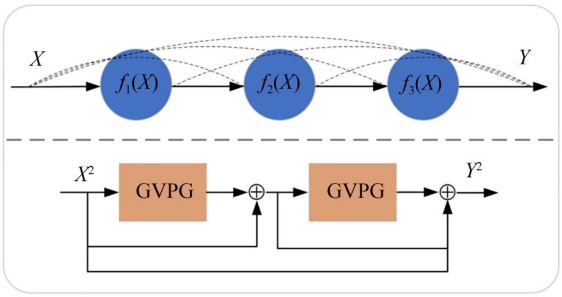

图4　密集引导视点策略示意图

Fig. 4　Illustration of the dense guidance viewpoint strategy

2.3　解码器

编码部分完成对道路特征的精准建模，输出四个不同分辨率特征表征$\boldsymbol{Y}^l\in\boldsymbol{R}^{H\times W\times C}$，$l\in(1,2,3,4)$。对于解码部分的设计本文考虑以下两个方面，结构如图5所示。

(1)编码阶段特征的分辨率和维度不同，输出的不同特征表征之间存在一定的位置和语义信息偏移。为解决该问题，解码部分设计多分辨率特征融合模块(MRFF)，采用特征连接和跨分辨率特征交互的操作方式，将每一特征表征都与其他特征表征进行融合，达到多分辨率特征信息充分融合的目的。为了尽可能学习到不同分辨率特征之间的信息偏移，相较于常规的跳跃连接[10]，该过程在采样前后均添加大卷积核改善对采样前后上下文信息的感知。

(2)解码部分通常需要对低分辨率特征上采样恢复高分辨率特征表征以输出道路提取结果。

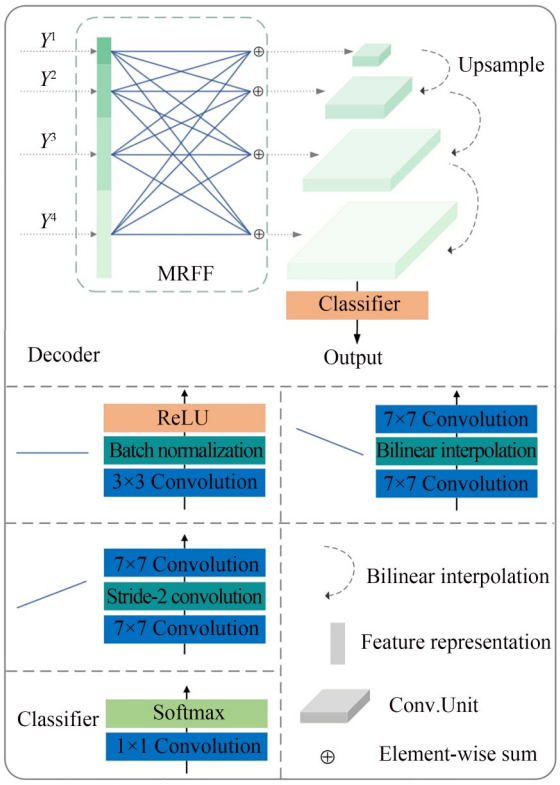

图5　解码器网络结构图

Fig. 5　Overall structure of the decoder

然而，若直接将低分辨率上采样到高分辨率会导致大量有效信息丢失。为最小化上采样过程中有效信息丢失，本文提出采用渐进式上采样策略。具体而言，渐进式上采样策略可以表述为公式(5)的迭代过程，其中$l\in(2,3,4)$。相较于直接将不同分辨率特征表征上采样到统一尺寸后融合的方式，该策略逐分辨率进行采样融合，最大程度避免不同分辨率间偏移以及采样倍数过大导致的信息损失。

$$\boldsymbol{Y}^{l}=Upsampling(\boldsymbol{Y}^{l}\oplus\boldsymbol{Y}^{l-1}). \tag{5}$$

最后，对于充分融合后的高分辨率表征，只需简单的卷积和*Softmax*设计的分类器，匹配每一像素类别实现对遥感图像道路的精准提取。

2.4 损失函数

遥感图像道路提取属于二分类密集预测问题，通常此类问题多选择交叉熵损失函数。本文分别统计了Massachusetts[41]和DeepGlobe[42]。遥感道路数据集中道路占比情况，分别为4.239%和2.864%，相比于建筑物、农田、树木等占比较少。若仅使用交叉熵损失，易造成模型弱化对道路类型前景的关注，导致学习过程陷入损失函数的局部极小值。考虑到Dice loss[43]在训练过程对每一类别同等重视，本文在交叉熵损失函数的基础上补充Dice loss缓解类不平衡问题，组成交叉熵-Dice loss遥感主辅损失函数，本文损失函数L_{seg}定义为：

$$L_{\text{seg}}=\alpha\cdot L_{\text{Dice}}+(1-\alpha)\cdot L_{\text{CE}}, \tag{6}$$

$$L_{\text{CE}}=\frac{1}{N}\sum_{i=1}^{N}y_i\log p_i+\left(1-y_i\right)\log\left(1-p_i\right), \tag{7}$$

$$L_{\text{Dice}}=1-\frac{\sum_{i=1}^{N}y_i\,p_i+\varepsilon}{\sum_{i=1}^{N}y_i+p_i+\varepsilon}, \tag{8}$$

其中：y_i代表第i个像素的真值，$y_i=1$表示该像素为道路，$y_i=0$表示该像素为背景；p_i代表第i个像素的预测概率$p_i\in\{0,1\}$；N是像素总数；$\varepsilon=1\times10^{-5}$是用于平滑损失的梯度因子；$\alpha$是调整交叉熵损失和Dice loss比例的超参数，根据工程经验一般选择0.4。

3 实验及结果分析

3.1 实验数据

本文选择Massachusetts遥感道路数据集和DeepGlobe遥感道路数据集进行实验。这两个数据集是遥感道路提取任务中最常用的数据集，均由真实世界的高分辨率光学遥感图像组成。

Massachusetts遥感道路数据集由1171张大小为1 500×1 500覆盖美国马萨诸塞州超过2 600 km^{2}的遥感图像组成。该数据集覆盖了广泛的城市、郊区和农村地区，并通过从OpenStreetMap项目中栅格化获得的道路中心线生成的标签。所有图像重新缩放到1 m/pixel的分辨率，其中包含1 108张训练图像，14张验证图像和49张测试图像。为统一尺寸，在原本1 500×1 500尺寸的数据底部和右侧填充像素到1 536×1 536，然后采用滑动窗口的方式裁剪出9个512×512尺寸的图像。

DeepGlobe遥感道路数据集地面分辨率为0.5 m/pixel，由8 570张大小1 024×1 024覆盖泰国、印度尼西亚、印度总面积2 200 km^{2}的遥感图像组成。数据集中不同样本的城市形态学、照明情况、道路密度及街道网络结构均有显著差异。一般情况下，对6 226张遥感图像及其标注图像划分为4 696张图像组成的训练集和1 530张图像组成的测试集。最后对原本尺寸1 024×1 024的图像采用滑动窗口的方式裁剪为4个512×512尺寸的图像。

经过裁剪后，实验数据集的样本图像如图6所示，上部分为Massachusetts遥感道路数据集，下半部分为DeepGlobe遥感道路数据集。

图6　数据集样本及其对应标签

Fig. 6　Sample images and labels of datasets

3.2　评价指标

为了验证本文提出算法的有效性，本文采用四个语义分割中常用的评价指标来评价不同算法性能：精准度（*Precision*）表示正确预测为道路与所有预测为道路的比例；召回率（*Recall*）表示预测为道路与实际道路的比例；F1分数（*F1-score*）是精准度和召回率的调和平均数，用于平衡准确率和召回率；交并比（Intersection Over Union，*IoU*）表示提取结果和真实标签的重叠面积和总面积之间的比率，各指标计算如下：

$$Precision=\frac{TP}{TP+FP}, \tag{9}$$

$$Recall=\frac{TP}{TP+FN}, \tag{10}$$

$$IoU=\frac{TP}{TP+FN+FP}, \tag{11}$$

$$F1\text{-}score=\frac{2\times precision\times recall}{precision+recall}, \tag{12}$$

其中：*TP*为道路类别预测正确的像素点总数；*FP*为道路类别预测错误的像素点总数；*FN*为背景类别预测错误的像素点总数。

3.3　实验设置

本文所进行的实验均使用NVIDIA GeForce GTX 3090（24 G）显卡和Pytorch深度学习框架构建、训练和测试模型。为保持实验公平性和一致性，本文与大部分对比模型保持一致的数据集划分方式和实验参数设置，采用AdamW作为优化器，批量大小设置为8、学习轮次为200、初始学习率5×10^{-4}。此外，采用CosineAnnealingLR学习率策略调整学习率，调整周期为50轮，最小学习率设置为1×10^{-6}。

为增强模型的泛化能力，训练时使用标准的数据增强，包括随机旋转、随机翻转、随机高斯模糊、随机掩码和随机对比度调整策略增强训练数据。

3.4　对比实验

3.4.1　*Massachusetts*遥感道路数据集对比实验

在Massachusetts遥感道路数据集本文对比算法包括，SegNet[44]，PSPNet[45]，CDG[46]，DA-RoadNet[23]，CADUNet[47]，DDU-Net[48]等主流道路提取算法。定量评价结果如表1所示。定量分析指标包括：*Precision*，*Recall*，*F1-score*和*IoU*。定量实验结果表明，本文所提算法优于其他主流算法，在所有对比方法中，由于RGGVNet引导模型视点关注特征模糊区域，在*Recall*，*F1-score*和*IoU*（81.93%，79.40%，65.84%）三个指标上具有明显优势。在*Recall*指标中，本文算法比最佳对比算法高出1.30%，较高的*Recall*意味着在道路预测中较少的漏检，正是预测任务中希望出现*Precision*和*Recall*是一对矛盾的性能指标，与DDU-Net相比，尽管RGGVNet的*Precision*更低，但是*Recall*远超DDU-Net。重要的是，在能更好反映算法性能综合指标*F1-score*中，本文所提算法与最佳对比算法相比高出1.21%。在*IoU*指标中，RGGVNet高于最佳对比算法1.65%，*IoU*是预测掩码与真实掩码的交并比，直接反映提取结果的质量。

表1　Massachusetts遥感道路数据集对比实验结果

Tab. 1　Results of comparative experiments on the Massachusetts road dataset

Methods	Precision/%	Recall/%	F1-score/%	IoU/%
FCN	69.16	70.14	69.65	53.43
SegNet	75.76	78.26	76.99	61.21
PSPNet	76.36	78.16	77.24	61.37
U-Net	76.12	79.01	77.54	62.52
LinkNet	81.50	80.63	77.07	62.70
DeeplabV3+	74.40	78.85	76.56	61.02
CDG	81.41	71.80	76.10	61.90
DA-RoadNet	79.16	77.25	78.19	61.90
CADUNet	79.45	76.55	77.89	64.19
DDU-Net	**82.54**	73.99	78.03	63.98
RGGVNet	77.02	**81.93**	**79.40**	**65.84**

每种算法对Massachusetts遥感道路数据集的定性评价结果如图7所示（彩图见期刊电子版）。可视化结果表明，本文所提算法明显优于其他对比算法。例如，在第一行红色区域，同时存在特征显著的主干道和特征模糊的辅道，由于GVPG模块的引导仅有本文所提算法能完整预测到主干道和辅道；第二行红色区域，周围环境特征与道路特征难以分辨其他，其他方法均在此处产生大量误判。RGGVNet对道路特征具有更强的感知能力，因此提取结果优于其他对比算法；第三行红色区域，道路特征受树木和阴影遮

挡严重导致特征模糊，其他方法均未完整地提取出道路，本文所提方法能更好处理这种情况；第四行红色区域，道路特征被树木几乎完整遮挡，GVPG 模块引导对特征模糊的关注，成功预测出被严重遮挡的道路；第五行红色区域，道路特征显著性水平明显低于其他区域，RGGVNet 模型视点平均对不同区域的关注，提取路网的完整性优于其他方法。

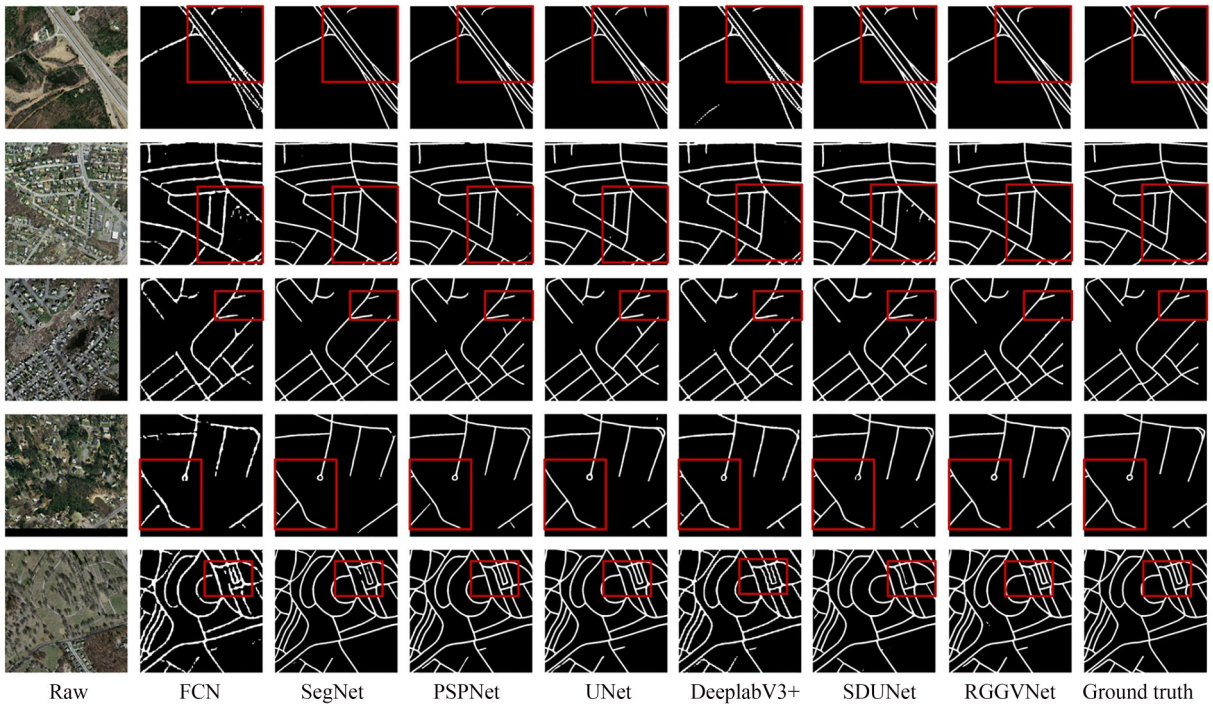

图 7　Massachusetts 遥感道路数据集不同算法可视化结果

Fig. 7　Visualized results of different algorithms on the Massachusetts road dataset

3.4.2　DeepGlobe 遥感道路数据集对比实验

在 DeepGlobe 遥感道路数据集本文对比方法包括 RADANet[24]，BDTNet[26]，DCS-TransUperNet[27]，D-LinkNet[49]，SDUNet[50]，CoANet[51] 和 RENA[52] 等主流遥感道路提取算法。

每种算法对 DeepGlobe 遥感道路数据集的定量评价结果如表 2 所示。定量分析指标包括：*Precision*，*Recall*，*F1-score* 和 *IoU*。定量实验结果表明，本文所提算法优于其他主流算法，与 Massachusetts 遥感道路数据集的实验结果相似，在 *Recall*，*F1-score* 和 *IoU*（84.37%，81.90%，69.36%）三个指标上具有明显优势。特别是在 *Recall* 指标上，本文所提算法比最佳对比方法高出 3.45%，对于预测任务来说是非常重要的，意味着本文方法对应特征模糊区域更加敏感。对于可以综合反映算法性能的 *F1-score* 和直接反映提取结果的 *IoU* 指标中，本文算法与最佳对比方法相比，分别提高 0.68% 和 0.99%。

每种算法对 DeepGlobe 遥感道路数据集的定性评价结果如图 8 所示（彩图见期刊电子版）。可视化结果表明，本文所提算法明显优于其他对比算法。例如，在第一行红色区域，同时存在特征显著的主路和特征模糊的小路道路，通过 GDVP 模块可以有效引导模型对小路的关注，实现完整提取道路；第二行红色区域，道路与周围环境容易混淆并且部分区域被树木遮挡，其他方法在此处误判情况严重，而 RGGVNet 在处理这种特征模糊情况时提取结果明显优于其他对比算法；第三行红色区域与第二行情况类似，进一步验证了本文所提方法的有效性；第四行红色区非铺装的道路与周围环境几乎融为一体，仅有 RGGVNet 能成功预测出所有的道路；第五行红色区域，道路上存在大量车辆和树木的遮挡，本所提方法效果明显优于其他方法。

表 2　DeepGlobe遥感道路数据集对比实验结果

Tab. 2　Results of comparative experiments on the DeepGlobe road dataset　　(%)

Methods	Precision	Recall	F1-score	IoU
FCN	71.38	74.74	73.02	57.51
SegNet	69.49	80.01	73.22	60.43
PSPNet	67.64	80.92	73.69	59.82
U－Net	81.50	80.40	80.95	67.99
LinkNet	77.45	79.98	78.68	64.86
DeeplabV3＋	68.72	80.48	74.14	61.02
D-LinkNet	—	—	—	64.12
RADANet	—	—	73.67	58.58
BDTNet	**84.18**	76.77	80.30	67.09
DCS-TransUperNet	82.44	78.43	80.39	65.36
SDUNet	78.40	74.20	79.40	66.80
CoANet	—	—	81.22	68.37
RENA	78.40	77.00	76.40	63.10
RGGVNet	79.58	**84.37**	**81.90**	**69.36**

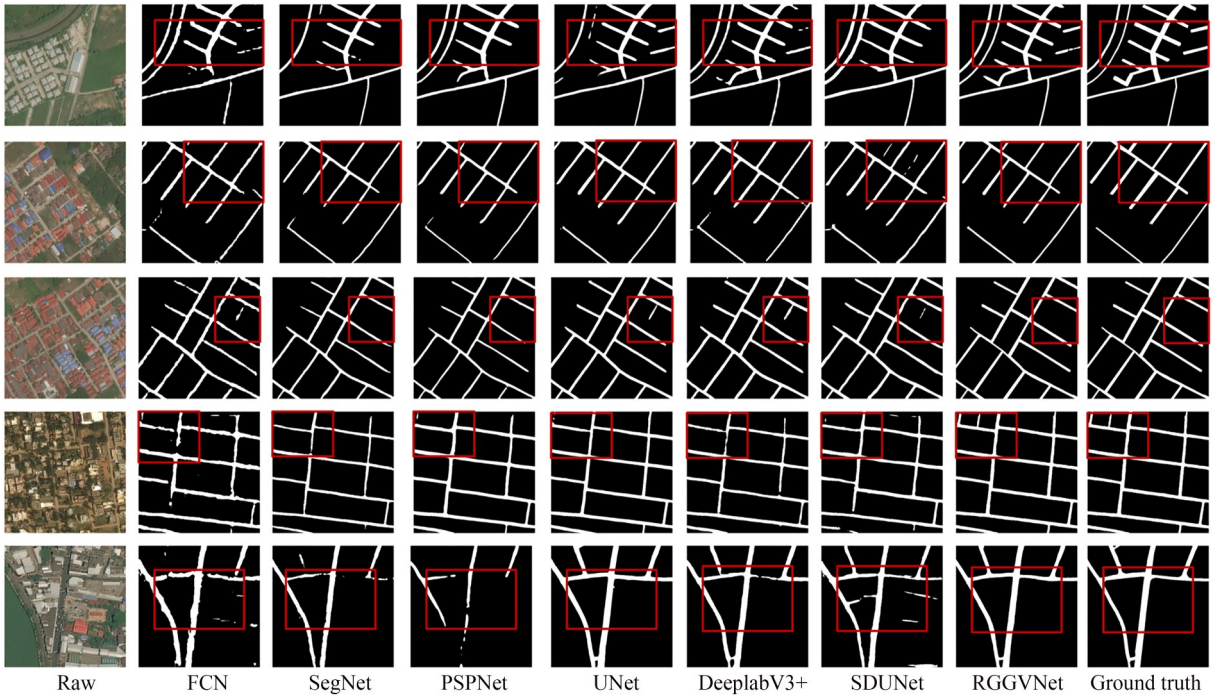

图 8　DeepGlobe遥感道路数据集不同算法可视化结果

Fig. 8　Visualized results of different algorithms on the DeepGlobe road dataset

3.5　消融实验

为了验证本文所提模型各模块和损失函数的有效性，选择在Massachusetts数据集上展开消融实验。实验采用ConvNeXt-T作为编码器以及去除MRFF模块的解码器作为基线模型(Baseline)，对比两个最具代表性的评价指标*IoU*和*F1-score*，并通过参数量(Params.)讨论模型复杂度。不同变种模型各种相关参数设置

和训练策略均保持一致，实验结果如表3所示。

表3　消融实验结果

Tab. 3　Results of ablation experiment

Method	*F1-score*/%	*IoU*/%	*Params.* ($\times 10^6$)
Baseline	75.89	61.54	29.663
Baseline-M	76.87	62.79	30.017
Baseline-ML	77.01	63.12	30.017
Baseline-MG	78.88	65.55	33.582
Baseline-MGS	79.17	65.70	33.582
Baseline-MGSL	**79.40**	**65.84**	33.582

Baseline-M：即在解码器部分添加本文设计的MRFF模块。由于MRFF模块消除特征之间语义和位置信息的偏移并最小化上采样过程中信息损失，对比基线模型*F1-score*和*IoU*分别提升0.98%和1.25%。同时，MRFF并未设计过多的卷积单元，参数量仅上涨0.354×10^6。

Baseline-ML：即在解码器部分添加本文设计的MRFF模块并添加Dice loss进行监督网络训练。在模型学习过程中，Dice loss引导模型对前景目标的关注，*F1-score*和*IoU*对比Baseline-M分别提升0.14%和0.33%。

Baseline-MG：在Baseline-M的基础上加入GDVP模块。实验结果表明，该模块在Baseline-M的基础上对*F1-score*和*IoU*分别提升2.01%和2.76%，并引入3.565×10^6的参数量。GDVP引导模型关注道路特征显著性水平较低区域，有效增强模型对特征信息的感知能力，显著提升了道路提取结果。与此同时，计算区域信息相似度有一定程度的计算开销，因此该模块引入少量参数。

Baseline-MGS：在Baseline-MG的基础上采用密集引导视点策略进行连接GDVP模块，该策略进一步提升*F1-score*和*IoU*。与此同时，该策略仅进行额外的连接，并不会带来参数量的提升。

Baseline-MGLS：代表在基线模型的基础上加入本文所提的所有方法。该算法在消融实验中取得最好结果，验证了本文所提方法的有效性。Baseline-MGLS对比基线模型在F1-score和IoU指标上，分别提升3.51%和4.30%，显著提升了评价指标的结果，意味着本文所提算法具有更优越的提取效果。

为验证不同损失函数比例α对实验结果的影响，本文选择在Massachusetts数据集上展开实验。实验对比两个最具代表性的评价指标*IoU*和*F1-score*，不同变种模型各种相关参数设置和训练策略均保持一致，实验结果如表4所示。

表4　损失函数超参数实验结果

Tab. 4　Experimental results of loss function hyperparametric

α	*F1-score*/%	*IoU*/%
0	79.17	65.70
0.2	79.33	65.80
0.4	**79.40**	**65.84**
0.6	79.13	65.81
0.8	78.97	65.74
1.0	78.63	65.48

可以看出当$\alpha=0.4$时，*F1-score*和*IoU*最高，*F1-score*和*IoU*均表现出先上升后下降的趋势。由于Dice loss直接计算分割结果与真实标签的交并比作为损失监督网络训练，有利于网络对前景目标的关注。然而，当α超过一定程度时，比例过大的Dice loss会导致梯度下降变得不稳定，直接影响F1-score和IoU。因此，当$\alpha=0.4$时Dice loss可以显著提升算法结果。

3.6　模型泛化性实验

为进一步验证模型泛化性，使用本文所提方法对WorldView-3号高分辨率遥感光学卫星收集到的图像进行了广泛测试。测试模型首先在DeepGlobe数据集训练到最佳结果，而后对收集到的图像进行测试。测试图像均未出现在训练集中，并且测试区域与训练区域具有较大差异。其中训练数据集主要覆盖泰国、印度尼西亚、印度等东南亚国家，而测试区域包括上海、拉斯维加斯、巴黎这三个城市的部分图像。部分实验结果如图9所示，由于测试数据并无标注标签，无法计算具体精度，因此仅进行定

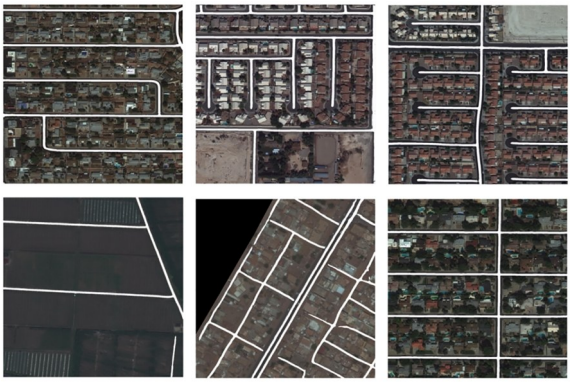

图9　在更广泛城市的测试结果

Fig. 9　Test results in wider cities

性分析。

可视化结果可以看出，本文方法具有良好的泛化性和鲁棒性。对于大部分区域道路都能取得精准的提取结果，特别是对于城市街区这种道路，几乎不会出现误判。

4　结　论

本文重新审视在光学遥感中现有道路提取方法在特征模糊区域难以正确提取道路这一问题，并基于GCN引导模型视点的思想提出RGGVNet。该网络核心模块GVPG在编码器和解码器连接处反复引导模型视点，实现对特征显著性水平不同的道路平等对待。为确保有效引导模型视点和道路特征信息的偏移和损失，进一步提出了密集引导视点策略和MRFF模块。

在Massachusetts和DeepGlobe遥感道路数据集实验的定量和定性分析表明，本文所提算法在遥感道路提取任务中具有明显优势，*IoU*分别达到65.84%和69.36%，*F1-score*分别达到79.40%和81.90%。进一步设计消融实讨论和验证了本文所提方法和模块有效性，在*F1-score*和*IoU*指标上分别提升3.51%和4.30%。

参考文献：

[1] 贾建鑫，孙海彬，蒋长辉，等. 多源遥感数据的道路提取技术研究现状及展望[J]. 光学 精密工程，2021，29(2)：430-442.
JIA J X，SUN H B，JIANG C H，*et al*. Road extraction technology based on multi-source remote sensing：[J]. *Opt. Precision Eng.*，2021，29(2)：430-442.(in Chinese)

[2] 姚凯旋，曹飞龙. 基于多输入密集连接神经网络的遥感图像时空融合算法[J]. 模式识别与人工智能，2019，32(5)：429-435.
YAO K X，CAO F L. Spatial-temporal fusion algorithm for remote sensing images based on multi-input dense connected neural network[J]. *Pattern Recognition and Artificial Intelligence*，2019，32(5)：429-435.(in Chinese)

[3] VALERO S，CHANUSSOT J，BENEDIKTSSON J A，*et al*. Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images[J]. *Pattern Recognition Letters*，2010，31(10)：1120-1127.

[4] CHAUDHURI D，KUSHWAHA N K，SAMAL A. Semi-automated road detection from high resolution satellite images by directional morphological enhancement and segmentation techniques[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*，2012，5(5)：1538-1544.

[5] GRINIAS I，PANAGIOTAKIS C，TZIRITAS G. MRF-based segmentation and unsupervised classification for building and road detection in peri-urban areas of high-resolution satellite images[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*，2016，122：145-166.

[6] SHAO Y Z，GUO B X，HU X Y，*et al*. Application of a fast linear feature detector to road extraction from remotely sensed imagery[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*，2011，4(3)：626-631.

[7] POULLIS C. Tensor-Cuts：a simultaneous multi-type feature extractor and classifier and its application to road extraction from satellite images[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*，2014，95：93-108.

[8] MABOUDI M，AMINI J，HAHN M，*et al*. Road network extraction from VHR satellite images using context aware object feature integration and tensor voting[J]. *Remote Sensing*，2016，8(8)：637.

[9] SHELHAMER E，LONG J，DARRELL T. Fully Convolutional Networks for Semantic Segmentation[C]. *IEEE Transactions on Pattern Analysis and*

Machine Intelligence. IEEE, 2017: 640-651.

[10] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]. *International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer*, 2015: 234-241.

[11] CHAURASIA A, CULURCIELLO E. LinkNet: exploiting encoder representations for efficient semantic segmentation[C]. 2017 *IEEE Visual Communications and Image Processing* (*VCIP*). *St. Petersburg*, *FL*, *USA*. IEEE, 2017: 1-4.

[12] CHEN L C, ZHU Y K, PAPANDREOU G, *et al*. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]. *Computer Vision-ECCV* 2018: 15*th European Conference*, *Munich*, *Germany*, 8-14, 2018, *Proceedings*, *Part VII*. ACM, 2018: 833 - 851.

[13] BUSLAEV A, SEFERBEKOV S, IGLOVIKOV V, *et al*. Fully convolutional network for automatic road extraction from satellite imagery[C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (*CVPRW*). *Salt Lake City*, *UT*, *USA*. IEEE, 2018: 197-1973.

[14] HE K M, ZHANG X Y, REN S Q, *et al*. Deep residual learning for image recognition[C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*). *Las Vegas*, *NV*, *USA*. IEEE, 2016: 770-778.

[15] XIN J, ZHANG X C, ZHANG Z Q, *et al*. Road extraction of high-resolution remote sensing images derived from DenseUNet[J]. *Remote Sensing*, 2019, 11(21): 2499.

[16] HUANG G, LIU Z, VAN DER MAATEN L, *et al*. Densely connected convolutional networks[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*). *Honolulu*, *HI*, *USA*. IEEE, 2017: 2261-2269.

[17] WU Q Q, LUO F, WU P H, *et al*. Automatic road extraction from high-resolution remote sensing images using a method based on densely connected spatial feature-enhanced pyramid[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 3-17.

[18] WANG Y, SEO J, JEON T. NL-LinkNet: toward lighter but more accurate road extraction with nonlocal operations[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 3000105.

[19] ZHANG Z X, SUN X, LIU Y X. GMR-net: road-extraction network based on fusion of local and global information[J]. *Remote Sensing*, 2022, 14(21): 5476.

[20] JIE Y S, HE H Y, XING K, *et al*. MECA-net: a MultiScale feature encoding and long-range context-aware network for road extraction from remote sensing images[J]. *Remote Sensing*, 2022, 14(21): 5342.

[21] VASWANI A, SHAZEER N, PARMAR N, *et al*. Attention is all you need[C]. *Advances in Neural Information Processing Systems*, *California*, *USA*, 2017, 30: 6000 - 6010.

[22] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, *et al*. *An Image is Worth* 16*x*16 *Words: Transformers for Image Recognition at Scale*[EB/OL]. 2020: *arXiv*: 2010.11929. http://arxiv.org/abs/2010.11929

[23] WAN J, XIE Z, XU Y Y, *et al*. DA-RoadNet: a dual-attention network for road extraction from high resolution satellite imagery[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 6302-6315.

[24] DAI L, ZHANG G Y, ZHANG R T. RADANet: road augmented deformable attention network for road extraction from complex high-resolution remote-sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5602213.

[25] XU Q X, LONG C, YU L, *et al*. Road extraction with satellite images and partial road maps[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 4501214.

[26] LUO L, WANG J X, CHEN S B, *et al*. BDT-Net: road extraction by Bi-direction transformer from remote sensing images[J]. *IEEE Geoscience and Remote Sensing Letters*, 1998, 19: 2505605.

[27] ZHANG Z, MIAO C L, LIU C A, *et al*. DCS-TransUperNet: road segmentation network based on CSwin transformer with dual resolution[J]. *Applied Sciences*, 2022, 12(7): 3511.

[28] LIU X Z, WANG Z Y, WAN J T, *et al*. RoadFormer: road extraction using a swin transformer combined with a spatial and channel separable convolution[J]. *Remote Sensing*, 2023, 15(4): 1049.

[29] SIMONYAN K, VEDALDI A, ZISSERMAN A. Deep inside convolutional networks: Visualis-

ing image classification models and saliency maps [J]. *2nd International Conference on Learning Representations*, ICLR 2014-Workshop Track Proceedings, 2014: 1-8.

[30] ZEILER M D, FERGUS R. *Visualizing and Understanding Convolutional Networks* [M]. Computer Vision-ECCV 2014. Cham: Springer International Publishing, 2014: 818-833.

[31] SELVARAJU R R, COGSWELL M, DAS A, *et al*. Grad-CAM: visual explanations from deep networks via gradient-based localization [J]. *International Journal of Computer Vision*, 2020, 128 (2): 336-359.

[32] LIU Z, MAO H Z, WU C Y, *et al*. A convnet for the 2020s [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. *New Orleans*, *LA*, *USA*. IEEE, 2022: 11966-11976.

[33] LI Z H, GU T C, LI B, *et al*. ConvNeXt-based fine-grained image classification and bilinear attention mechanism model [J]. *Applied Sciences*, 2022, 12(18): 9016.

[34] ZHOU J J, ZHANG B H, YUAN X L, *et al*. YOLO-CIR: the network based on YOLO and ConvNeXt for infrared object detection [J]. *Infrared Physics and Technology*, 2023, 131: 104703.

[35] HAN Z M, JIAN M W, WANG G G. ConvUNeXt: an efficient convolution neural network for medical image segmentation [J]. *Knowledge-Based Systems*, 2022, 253: 109512.

[36] KIPF T N, WELLING M. *Semi-Supervised Classification with Graph Convolutional Networks* [EB/OL]. 2016: *arXiv*: 1609.02907. http://arxiv.org/abs/1609.02907

[37] LI Y, GUPTA A. Beyond grids: learning graph representations for visual recognition [C]. *Proceedings of the* 32*nd International Conference on Neural Information Processing Systems. December* 3-8, 2018, *Montréal*, *Canada*. ACM, 2018: 9245-9255.

[38] LIU Q H, KAMPFFMEYER M, JENSSEN R, *et al*. Self-constructing graph convolutional networks for semantic labeling [C]. *IGARSS* 2020 - 2020 *IEEE International Geoscience and Remote Sensing Symposium*. *Waikoloa*, *HI*, *USA*. IEEE, 2020: 1801-1804.

[39] LI Q M, HAN Z C, WU X M. Deeper insights into graph convolutional networks for semi-supervised learning [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018, 32(1): 1.

[40] XU K, LI C T, TIAN Y L, *et al*. *Representation Learning on Graphs with Jumping Knowledge Networks* [EB/OL]. 2018: *arXiv*: 1806.03536. http://arxiv.org/abs/1806.03536

[41] MNIH V. *Machine Learning for Aerial Image Labeling* [M]. Canada: University of Toronto, 2013.

[42] DEMIR I, KOPERSKI K, LINDENBAUM D, *et al*. DeepGlobe 2018: a challenge to parse the earth through satellite images [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. *Salt Lake City*, *UT*, *USA*. IEEE, 2018: 172-17209.

[43] MILLETARI F, NAVAB N, AHMADI S A. V-Net: fully convolutional neural networks for volumetric medical image segmentation [C]. 2016 *Fourth International Conference on 3D Vision (3DV)*. *Stanford*, *CA*, *USA*. IEEE, 2016: 565-571.

[44] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.

[45] ZHAO H S, SHI J P, QI X J, *et al*. Pyramid scene parsing network [C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. *Honolulu*, *HI*, *USA*. IEEE, 2017: 6230-6239.

[46] WANG S, YANG H, WU Q Q, *et al*. An improved method for road extraction from high-resolution remote-sensing images that enhances boundary information [J]. *Sensors*, 2020, 20(7): 2064.

[47] LI J, LIU Y, ZHANG Y D, *et al*. Cascaded attention DenseUNet (CADUNet) for road extraction from very-high-resolution images [J]. *ISPRS International Journal of Geo-Information*, 2021, 10(5): 329.

[48] WANG Y, PENG Y X, LI W, *et al*. DDU-net: dual-decoder-U-net for road extraction using high-resolution remote sensing images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 4412612.

[49] ZHOU L C, ZHANG C, WU M. D-LinkNet:

LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (*CVPRW*). *Salt Lake City*, *UT*, *USA*. IEEE, 2018: 192-1924.

[50] YANG M X, YUAN Y, LIU G C. SDUNet: road extraction via spatial enhanced and densely connected UNet[J]. *Pattern Recognition*, 2022, 126: 108549.

[51] MEI J, LI R J, GAO W, *et al*. CoANet: connectivity attention network for road extraction from satellite imagery [J]. *IEEE Transactions on Image Processing*, 2021, 30: 8540-8552.

[52] SHAO S W, XIAO L X, LIN L P, *et al*. Road extraction convolutional neural network with embedded attention mechanism for remote sensing imagery[J]. *Remote Sensing*, 2022, 14(9): 2061.

作者简介：

刘光辉(1976—),男,副教授,西安建筑科技大学信息与控制工程学院硕士生导师,2016年于西安建筑科技大学获得工学博士学位,主要从事计算机视觉感知与理解、人工智能与智能化系统、建筑智能化技术方面的研究。E-mail: guanghuail@163.com

孟月波(1979—),女,陕西西安人,教授,西安建筑科技大学信息与控制工程学院硕士生导师,2014年于西安交通大学获得工学博士学位,主要从事计算机视觉感知与理解、人工智能与智能化系统、建筑智能化技术方面的研究。E-mail: mengyuebo@163.com